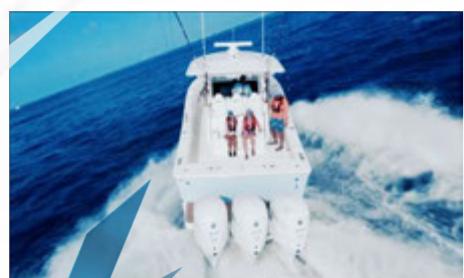


# YAMAHA MOTOR TECHNICAL REVIEW



**巻頭言**

人間研究とイノベーションの重要性	5
設楽 元文	

**製品紹介**

2025年新型「TRACER9/GT/GT+」	7
尾上 知輝	
"Ténéré 700" Development Target & Objectives	12
Stefano Galimberti Takamasa Tokitoh Paolo Barlaam Marco Rossini Kei Sugimoto Naoya Shiraishi Yoshitomo Akimoto	
スポーツボート フラッグシップ「295XD」の開発	19
三浦 宏信 杉本 亮平	
電動操船システム「新型 HARMO」の開発	25
前島 将樹 松井 隆浩 武田 謙太郎 郷家 孝之 梶原 天翔 池ヶ谷 祐次	
Wireless Joystick および MOB の開発	29
鷹野 直樹 徳田 光太 Andrew Artusa	
5人乗りゴルフカー開発	34
加治屋 晋一 勝田 悠馬 宮内 泰寛 櫻木 伊織 安西 賢汰 澤村 充 四俵 海太 池田 桂一郎	
車いす用電動化ユニット「JWG-1」の開発	38
水谷 浩幸 松苗 徹 岩城 泰宏 水野 正光 大西 慎太郎 井口 裕貴 杉原 優 斎藤 純平 八木 啓明 笹岡 俊樹 林 桂太郎 糸賀 信行	
新世代高速ワイヤボンダ「UTC-RZ1」	43
富山 俊彦 石山 義人	

**技術紹介**

ゴルフカー向け内製リチウムイオン電池パック開発	48
長澤 佑樹 宮崎 敏典 香田 昌伸 安藤 義崇汰 竹内 祐哉 宮田 明典 程 再冉 三輪 千夏 笹本 侑紀 藤原 一樹 田中 洋充	
リスクアセスメントの質向上を狙った VR 活用	
VR 活用による設備導入前リスクアセスメント精度向上	53
上川 和輝 曽根 沙季	
SMT フロア無人化に向けた自動化ソリューション技術	57
天内 真	
クラウドで実現する迅速簡便なシステム検証環境の構築	60
落合 寛彰 吉成 雅通 人見 瑛一 浜田 善夫	
社会共創による新価値創造「Town eMotion」Vol. 2	
—まちなか R&D における GSM 活用と循環車共創の実践と展望—	64
住友 剛 神原 瑞穂 杉浦 知志	
高出力応答が求められる小型モビリティへの適用を目指した FCV の電力マネジメントシステムの開発	71
鈴木 聖也 中田 暢大	
板鍛造工法による新「NMAX」「YECVT」部品の一体成形開発	76
田丸 翔吾 北郷 博成	

## 技術論文

マリン用水素直噴エンジンと水素ポートシステムの開発	81
鬼形 朋良 平本 成実 弘田 雅士 能勢 幸憲 Hien H. Nguyen	
船外機のクランク打音予測のためのジャーナル荷重予測手法の開発	90
井上 祐人 高比良 拓人 藤田 英之 小林 光司 近藤 拓 高橋 徹	
FRP プール50年の実績と構造的特徴に関する報告	98
松井 健良 内山 仁平 菊地 秀和 喜田 晃爾 糸永 汐里 萩野谷 健吾	
Development of CO <sub>2</sub> emission reduction technology for sport motorcycles	105
牧田 直希 鳥越 昌樹 高橋 俊彦 高瀬 浩規	
Human model on multi-body dynamics simulation of Motorcycle	118
植木 幹人 高山 照浩 矢部 昇	
A Study of Path Prediction Algorithm for Camera-Based Rider Assistance System	126
長谷川 翔馬 武田 拓己 尾上 太郎 品川 晃徳	
沖永良部島の脱炭素・持続可能なモビリティ社会実現へ向けた実証研究	136
和田 朋智毅 川口 雅弘 中村 俊之 福田 晋平 稲波 純一 山本 俊行 桑原 昌広 乾 大樹	
[NEDO 助成事業]	
ポスト5G向けチップオンウエハダイレクト接合3D 積層統合技術開発	146
菊地 広 菱沼 隼 李 瑾 永田 憲雅 尾川 謙一	
DiffMamba: Leveraging Mamba for Effective Fusion of Noise and Conditional Features in Diffusion Models for Skin Lesion Segmentation	161
Amit Shakya Shruti Phutke Chetan Gupta Rupesh Kumar Lalit Sharma Chetan Arora	
Two-Stage, Global-Local Approach for Cell Nuclei Segmentation in Histopathology Images	173
Amit Shakya Revat Saharan Chetan Gupta Rupesh Kumar Lalit Sharma Srivatsava Naidu Subrahmanyam Murala Chetan Arora	
Do not look so locally to fish skins: Improved YOLOv7 for fish disease detection with Transformers	182
Shruti Phutke Amit Shakya Chetan Gupta Rupesh Kumar Tsuyoshi Kuroda Lalit Sharma	
OSCMamba: Omni-directional Selective Scan Convolution Mamba for Medical Image Classification	192
Shruti Phutke Amit Shakya Chetan Gupta Rupesh Kumar Lalit Sharma	
Automatic Early Classification of Cassava Leaf Disease with Ensemble of Lightweight Models	204
Kaushik Vishwakarma Udaybhan Rathore Rupesh Kumar Amit Shakya Shruti Phutke Chetan Gupta Lalit Sharma	
A Novel Two-Stage Deep Learning Method for Enhanced Cell Nuclei Segmentation	214
Chetan Gupta Rupesh Kumar Amit Shakya Shruti Phutke Lalit Sharma	

**Foreword**

The Importance of Human Research and Innovation	5
Motofumi Shitara	

**New Products**

2025 New Model "TRACER9/GT/GT+"	7
Tomoki Onoue	
"Ténéré 700" Development Target & Objectives	12
Stefano Galimberti Takamasa Tokitoh Paolo Barlaam Marco Rossini Kei Sugimoto Naoya Shiraishi Yoshitomo Akimoto	
Development of the Sports Boat Flagship "295XD"	19
Hironobu Miura Ryohei Sugimoto	
Development of the "New HARMO" Electric Boat Maneuvering System	25
Masaki Maejima Takahiro Matsui Kentaro Takeda Takayuki Gouke Tensho Kajihara Yuji Ikegaya	
Development of the Wireless Joystick and MOB detection system	29
Naoju Takano Kota Tokuda Andrew Artusa	
Development of a Five-Passenger Golf Car	34
Shinichi Kajiyama Yuuma Katsuda Yasuhiro Miyauchi Iori Sakuragi Kenta Anzai Mitsuru Sawamura Kaita Shidawara Keiichiro Ikeda	
Development of the Electric Drive Wheelchair Unit "JWG-1"	38
Hiroyuki Mizutani Toru Matsunae Yasuhiro Iwaki Masamitsu Mizuno Shintaro Ohnishi Hirotaka Iguchi Yu Sugihara Hiroaki Yagi Kohei Saito Toshiki Sasaoka Keitaro Hayashi Nobuyuki Itoga	
Next-Generation High-Speed Wire Bonder "UTC-RZ1"	43
Toshihiko Toyama Yoshito Ishiyama	

**New Technologies**

In-house Development of Lithium-ion Battery Packs for Golf Cars	48
Yuki Nagasawa Toshinori Miyazaki Yoshinobu Koda Rasuta Ando Yuya Takeuchi Akinori Miyata Zairan Cheng Chinatsu Miwa Yuki Sasamoto Kazuki Fujiwara Hiromichi Tanaka	
Use of VR to Improve the Quality of Risk Assessment	
Enhancing the Accuracy of Pre-Installation Risk Assessment through VR Utilization	53
Kazuki Kamikawa Saki Sone	
Automation Solution Technology Toward an Unmanned SMT Floor	57
Shin Amanai	
Rapid and simple setup of system evaluation environment by the Cloud	60
Hiroaki Ochiai Masamichi Yoshinari Eiichi Hitomi Yoshio Hamada	
Creating New Value through Social Co-Creation: "Town eMotion" Vol. 2	
—Practical Applications and Future Outlook of GSM Utilization and Collaborative Development of Circular Vehicles in Urban R&D—	64
Takeshi Sumitomo Mizuho Sakakibara Satoshi Sugiura	
Development of a Power Management System for Fuel Cell Vehicles (FCVs) Aimed at Application in Small Mobility Vehicles Requiring High Power Response	71
Masaya Suzuki Nobuhiro Nakata	
Integrated forming development of new "NMAX" and "YECVT" components using the plate forging method	76
Shogo Tamaru Hironari Hongo	

**Technical Papers**

Development of a Hydrogen Direct-Injection Engine and Hydrogen Boat System for Marine Applications	81
Tomoyoshi Onikata Narumi Hiramoto Masashi Hirota Yukinori Nose Hien H. Nguyen	
Developing a Journal Load Prediction Method for Crank Knock in Outboard Motors	90
Yuto Inoue Takuto Takahira Hideyuki Fujita Koji Kobayashi Taku Kondoh Toru Takahashi	
Report on 50 years of FRP pool achievements and structural characteristics	98
Takayoshi Matsui Jimpei Uchiyama Hidekazu Kikuchi Koji Kida Shiori Itonaga Kengo Haginoya	
Development of CO2 emission reduction technology for sport motorcycles	105
Naoki Makita Masaki Torigoshi Toshihiko Takahashi Hiroki Takase	
Human model on multi-body dynamics simulation of Motorcycle	118
Motohito Ueki Akihiro Takayama Noboru Yabe	
A Study of Path Prediction Algorithm for Camera-Based Rider Assistance System	126
Shoma Hasegawa Takumi Takeda Taro Onoue Akinori Shinagawa	
Verification Study toward Realizing a Decarbonized and Sustainable Mobility Society on Okinoerabu Island	136
Tomoki Wada Masahiro Kawaguchi Toshiyuki Nakamura Shimpei Fukuda Junichi Inami Toshiyuki Yamamoto Masahiro Kuwahara Daiki Inui	
Development of Chip-on-Wafer Direct Bonding 3D Stacking Technology for Post-5G	146
Hiroshi Kikuchi Hayato Hishinuma Jin Li Norimasa Nagata Kenichi Ogawa	
DiffMamba: Leveraging Mamba for Effective Fusion of Noise and Conditional Features in Diffusion Models for Skin Lesion Segmentation	161
Amit Shakya Shruti Phutke Chetan Gupta Rupesh Kumar Lalit Sharma Chetan Arora	
Two-Stage, Global-Local Approach for Cell Nuclei Segmentation in Histopathology Images	173
Amit Shakya Revat Saharan Chetan Gupta Rupesh Kumar Lalit Sharma Srivatsava Naidu Subrahmanyam Murala Chetan Arora	
Do not look so locally to fish skins: Improved YOLOv7 for fish disease detection with Transformers	182
Shruti Phutke Amit Shakya Chetan Gupta Rupesh Kumar Tsuyoshi Kuroda Lalit Sharma	
OSCMamba: Omni-directional Selective Scan Convolution Mamba for Medical Image Classification	192
Shruti Phutke Amit Shakya Chetan Gupta Rupesh Kumar Lalit Sharma	
Automatic Early Classification of Cassava Leaf Disease with Ensemble of Lightweight Models	204
Kaushik Vishwakarma Udaybhan Rathore Rupesh Kumar Amit Shakya Shruti Phutke Chetan Gupta Lalit Sharma	
A Novel Two-Stage Deep Learning Method for Enhanced Cell Nuclei Segmentation	214
Chetan Gupta Rupesh Kumar Amit Shakya Shruti Phutke Lalit Sharma	

This year marks the milestone of Yamaha Motor's 70th anniversary. Over these seven decades, we have built and solidified the Yamaha brand by delivering *Kando*\* to customers through our products and services. And we must continue to enhance our presence through the brand that our customers have shaped together with us.

In doing so, we must always be mindful of the unique style of Yamaha that defines us.

We have defined the pursuit of Yamaha's unique style as our guiding principle in *Monozukuri* (engineering, manufacturing and marketing), expressed through five characters: 発 (*Hatsu*), 悅 (*Etsu*), 信 (*Shin*), 魅 (*Mi*), and 結 (*Ketsu*) - meaning Innovation, Excitement, Confidence, Emotion, and Ties.

When we focus on new products born from fresh ideas, and on the technological fields that underpin them, we can see a history of challenging innovation through unique concepts—surpassing their era with what had never been achieved before. In other words, we have consistently pursued “*Hatsu*” = Innovation. Furthermore, to refine our brand value, we have set forth our long-term vision: “ART for Human Possibilities — Let's strive for greater happiness.” What we seek is happiness that comes from enriched lifestyles and enriched hearts, beyond technology itself. Since the COVID-19 pandemic, we have felt through rising demand that people everywhere in the world are craving this kind of happiness, born from both lifestyle and spiritual richness, more strongly than ever.

To bring this long-term vision to life, Yamaha Motor's new core technologies are Energy Management, Intelligent Systems, and Software Services. The cornerstone of value creation that supports them all is Human Research. In this digital age, when so much work can be replaced by AI, an opportunity lies before us — as a company that has always emphasized Human-Centric *Monozukuri*. Considering what we have observed since the pandemic, we find that when we engage in dialogue with investors on this point, it resonates strongly.

What we should be especially mindful of now are those abilities unique to human beings: understanding emotions, creativity, ethics, flexibility, and the ability to learn through experience. By engaging the five senses in various experiences, humans deepen knowledge and understanding. They connect this knowledge and understanding to flexible thinking and ideas grounded in emotion, intuition, and lived experience. What matters most is the deep understanding and emotional bonds that come only from direct experience. This applies not only to products, but also to watching rugby, soccer, motorsports and more. Yet what we aim to appeal to in people's sensibilities is not the virtual, but the real world.

It has now been more than thirty years since the collapse of Japan's economic bubble, during which time the country's economic stagnation has been repeatedly pointed out. Much of this is due to Japanese companies being unable to adequately respond to the structural changes in society and the economy during that period — namely globalization, digitalization, declining birthrate and aging population, and changes in values.

For us, a company that has embraced the corporate purpose of creating *Kando* and advanced human-centric *Monozukuri*, an opportunity is at hand. Understanding that we must respond effectively to these structural changes in society and the economy, our mission is to continuously innovate through both technology and business.

\**Kando* is a Japanese word for the simultaneous feelings of deep satisfaction and intense excitement that we experience when we encounter something of exceptional value

今年はヤマハ発動機にとって、節目の創立70周年にあたります。この70年の歩みの中で、当社は製品やサービスを通じてお客様に感動を届けることで、ブランドとしてのヤマハを確立してきました。そして私たちはこれからもお客様により形成されてきたブランドとしての存在感を高め続ける必要があります。

その際に常に意識すべきは、ヤマハとしての「らしさ」です。

そして私たちはこの「らしさ」を、「発・悦・信・魅・結」の五つの漢字でモノづくりの道標として定義してきました。

新しい発想による商品や、その礎となる技術領域にフォーカスしてみると、独創的なコンセプトで革新に挑むこと、時代の先をいく革新的なこと、今まで誰も成し得なかったこと、言い換えると「発」=イノベーションに挑戦してきた歴史を確認することができます。

さらに当社は、ブランドとしての価値に磨きをかけるため、「ART for Human Possibilities 人はもっと幸せになれる」を長期ビジョンとして掲げています。追い求めるのは、技術の先にある生活の豊かさ、心の豊かさからくる幸せです。私たちはコロナ禍以降、世界中の人々が特に以前よりも増してこの生活と心の豊かさからくる幸せを欲していることを、需要の増加によって実感することができました。

長期ビジョンを具現化するための当社の新たなコア技術は、エネルギー・マネジメント、知能化、ソフトウェアサービスであり、これらを支える価値創造の礎が「人間研究」です。デジタルの時代、多くの仕事がAIに置き換え可能になる中で、人間中心のモノづくりを重視してきた当社にとって、チャンスが訪れています。コロナ禍以降の現象を踏まえ、投資家の皆さんとこの点について対話すると共感を得ることができます。

私たちが今改めて意識すべきは人間ならではの能力である、感情の理解、創造性、倫理観、臨機応変な対応力、そして体験を通じた学習の重要性です。人間は五感を通じて様々な体験をすることで、知識や理解を深め、感情や直感、経験に基づく柔軟な思考や発想に結びつけることができるのです。重視すべきは、実際に体験することで得られる深い理解や感情的な結びつきです。これは商品だけでなく、ラグビーやサッカー、モータースポーツなどの観戦からも得られますが、私たちが感性に訴えるのは、バーチャルではなくリアルの世界です。

バブル崩壊後30年、日本経済の低迷が言われて久しいです。この間に起こった社会経済の構造変化、すなわち、グローバル化、デジタル化、少子高齢化、価値観の変化というものに、日本企業がうまく対応できなかったことは大きいと思います。

感動創造を企業目的に掲げ、人間中心のモノづくりを進めてきた当社にとって、チャンスが訪れています。そしてこの社会経済の構造変化にうまく対応しなければならないことを理解したうえで、技術と事業によってイノベーションを起こし続けていくことが、私たちの使命です。

### ■著者



設楽 元文  
Motofumi Shitara  
代表取締役社長

# 2025年新型「TRACER9/GT/GT+」

## 2025 New Model "TRACER9/GT/GT+"

尾上 知輝



## Abstract

Since the launch of the first-generation model in 2015, the "TRACER series" has continued to deliver new value to the sport-touring category. By achieving a high-level balance between the dynamic performance of a sport bike and the practicality for versatile use, it has evolved through four generations with various refinements, establishing a unique position as a model that meets a wide range of riding needs. With the introduction of the fifth-generation new model, this paper presents the key features of the product.

## 1 はじめに

2015年に登場した初代モデル以降、「TRACERシリーズ」はスポーツツーリングカテゴリーに新たな価値を提供してきた。スポーツバイクとしての高い運動性能と、多目的に使用可能な実用性を高次元で両立させることにより、大小さまざまな改良を重ねながら四世代にわたり進化し、幅広い用途に対応するモデルとして、独自の地位を確立した。

今回、第五世代の新型モデルを導入するにあたり、本稿ではその製品特長を紹介する。

## 2 商品の企画

モデルコンセプトには、"The matured Multirole fighter of the motorcycle"を掲げ、「TRACER9」、「TRACER9 Y-AMT」、「TRACER9 GT/「TRACER9 GT Y-AMT」、「TRACER9 GT+」の全5種類へバリエーションを拡大した。シリーズのエントリー

グレードから、ヤマハが有する先進技術を集結させたフラッグシップモデルまでを網羅し、幅広いニーズに応える多目的スポーツツーラーとして、さらなる熟成と進化を遂げた。

平日も週末も、街中もワインディングも問わず、いかに気軽に走り出せるかというキャラクターを "7days pleasure" と定義し、真のオールラウンダーとしての理想を追求したモデルである。

## 3 開発内容

多数の新機能を搭載しつつ、モーターサイクルとしての基本部分の商品性向上にも注力し、全バリエーションにおいて競争力を高めた。それぞれの内容について以下に示す。

### 3-1. トルクフルな CP3エンジン

888cm<sup>3</sup>・水冷・4ストローク・DOHC(ダブル・オーバーヘッド・カムシャフト)・3気筒・4バルブのCP3(クロスプレーン・コンセプトの3気筒)エンジンを搭載する。78.0mm × 62.0mmのボア×

ストローク、圧縮比11.5:1により、トルクフルで余裕のあるパワーを発揮する(93Nm/87.5kw)。きめ細かな制御によって、低回転域から高回転域まで扱いやすいリニアな出力特性を実現した。

### 3-2. 車体開発について

当社独自の製法によるCF(コントロールド・フィリング)アルミダイキャストフレームを採用した。メイン部分は現行モデルを踏襲しつつ、ヘッドパイプおよびクロスチューブの剛性最適化を図った。内部構造の見直しにより前後で500gの軽量化を達成し、接地感と乗り心地の向上を狙って専用開発したタイヤ(ブリヂストン製「BATTAX SPORT TOURING T32」)と組み合わせ、ハンドリング性能を高めた。

さらに、ハンドル切れ角の拡大(32°→35°)により最小回転半径は2.9mとなり、ネイキッドモデルを上回る小回り性能を実現し、日常の取り回し性を向上させた。シート形状の見直しにより足つき性と快適性を改善し、ロングツーリングにも快適に対応できる性能を備えた。リヤフレームは50mm延長し、タンデムライダーの居住性も確保した。単独走行でもタンデムでも快適に楽しめる車体構成となっている。

ツアラーでありながら俊敏性を兼ね備える「TRACER」は、そのスポーツ性をフロントマスクのスリムさおよびシャープさで表現する。今回の新型においてもそのイメージを崩すことなく、ウインドプロテクションの向上に注力した。快適性向上のため、風洞実験およびCFD解析を繰り返し実施した。並行してライディングポジションも見直し、ハンドルのしごり角などを変更した。よりリラックスした上体姿勢により、長時間・長距離走行時の疲労軽減を図った(図1、図2)。



図1 Function Area



図2 Air Flow イメージ

### 3-3. 代表的な“Newフィーチャー”

当モデルのフィーチャーマップを図3に示し、その中の一部を以下に紹介する(図3)。



図3 フィーチャーマップ

### 3-3-1. アダプティブクルーズコントロール+Y-AMT

先代モデルで当社として初採用したアダプティブクルーズコントロール(以下 ACC)に、CP3エンジン搭載モデルとしては2024年「MT-09」に初搭載された「Y-AMT<sup>[1]</sup>」を組み合わせて搭載する。ACCは、前方に搭載したミリ波レーダーが状況に応じて定速巡航・減速・加速の制御を自動で行う。ACC作動中は、定速走行時のギア選択に加え、車速の変化に応じてY-AMTが自動的に変速を行う。一般的な四輪のAT車と同様に、追従走行中のシフトアップ・シフトダウン、あるいはその維持をシステムに委ねることができる、二輪車として世界初<sup>※</sup>の制御である。これにより、両手両足の操作負荷を軽減し、利便性を向上させた。さらに、定速走行時には変速頻度を抑えた専用シフトプログラムを用意し、急減速時には素早いシフトダウンによって減速感を高めるなど、落ち着いた快適な移動を実現する。

※当社調べ(2025年3月時点)

### 3-3-2. ブラインドスポットディテクション

ミリ波レーダーは車体前方だけでなく後方にも搭載した。後方から接近する車両や死角(ブラインドスポット)に存在する車両を検知し、ミラー内に表示するブラインドスポットディテクション(以下 BSD)機能<sup>※</sup>を新たに採用した。後方レーダーは周囲の走行状況を可視化する(図4)。

※本機能は運転支援を目的としたものであり、安全確認はライダー自身が行うことを前提としている。



図4 BSD インジケータ

### 3-3-3. フォワードコリジョンワーニング

車両前方に搭載されたミリ波レーダーにより先行車を検知し、衝突リスクが高まった際に2段階で警告を行う運転支援システム「FCW(フォワードコリジョンワーニング)<sup>※</sup>」を採用した(図5)。

ライダーの入力を確認後、前モデルから搭載されているレーダー連携ブレーキ制御が作動する。

Level. 1:ディスプレイ表示

先行車との接近により衝突の可能性がある場合、ディスプレ

イ表示によって警告し、減速や回避を促す。

Level. 2:ブレーキパルス

ディスプレイ表示のみではライダーの認知が遅れる可能性を考慮し、リヤブレーキを瞬間に加圧することで危険を知らせる。

※本機能は衝突を回避するものではなく、あくまで運転支援を目的としている。

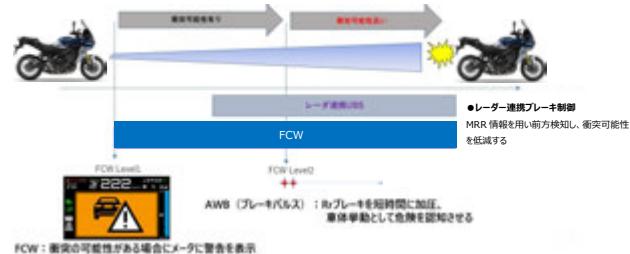


図5 FCW 作動イメージ

### 3-3-4. マトリクス LED ヘッドライト

複数のLo/Hiビーム用LEDと車体上部に搭載されたカメラが連携し、周囲の交通状況を自動で判断する。対向車や先行車の眩惑を抑制するため、照射エリアを自動的に調整する。これにより、ライダーの夜間走行時の負担軽減および安全運転支援に寄与する。さらに、コーナリング時には車体の傾きに応じて配光パターンを自動調整し、旋回時の良好な視認性に寄与する(図6)<sup>[2][3]</sup>。



図6 ヘッドライト点灯例

### 3-3-5. ヴィーカルホールドコントロール

坂道での停止時や発進時にブレーキ操作をアシストする「ヴィーカルホールドコントロール(VHC)」(図7)を、当社として初めて採用した。本機能はブレーキ圧を保持し、停車時の車体挙動を安定させる制御である。“ADVANCED”と“MANUAL”的2パターンを設定し、“ADVANCED”では車体姿勢の検知により路面が坂道かどうかを判定し、停車状態を維持するために必要なブレーキ入力を保持する。“MANUAL”では、ライダーが意図

的にブレーキレバーへ強く入力することで作動し、停車状態を維持することができる。

○坂道での停止時や坂道からの発進時にブレーキ操作を補助し、極低速での安定性向上を狙う。

VHC : Vehicle Hold Control



図7 VHC 作動イメージ

### 3-3-6. タイヤプレッシャーモニタリングセンサ

「TRACER9 GT+」には、タイヤ内の空気圧を計測する“TPMS（タイヤプレッシャーモニタリングシステム）”（図8）を標準装備した。計測された数値は車載の受信機で受信の上メーター内に表示され、走行中でも確認可能である（図9）。空気圧の異常を早期に把握できることで、パンクやタイヤ損傷によるリスクを低減する。「TRACER9 GT」にはオプション設定とした。



図8 TPMS 部品構成



図9 空気圧表示例

### 3-3-7. ローメンテナンスチェーン

摩耗性に優れ、低摩擦のDLCコーティングを施したローラー部を備える“ローメンテナンスチェーン”（DID 社製）（図10）を当社として初めて採用した。これにより、優れた耐久性とデザイン性に加え、チェーンの伸びを抑制し、摺動部の抵抗を軽減することで押し歩き時の軽快さを実現した。



図10 チェーンイメージ

### 3-3-8. キーレスロッキングシステム

スマートキーを採用する「TRACER9 GT」／「TRACER9 GT Y-AMT」および「TRACER9 GT+」には、トップケースとサイドケースのロックおよびアンロックを一括操作可能な“キーレスロッキングシステム”を、当社として初めて搭載した。

スマートキーからの入力を車体側の“BCM（ボディコントロールモジュール）”が受信することで、メカニカルキーを使用することなく車体側のスイッチを押すだけでケースの脱着および開閉が可能となった。さらに、ケース脱着時のカプラーやケーブルの抜き差しを不要とするためロック機構に対しワイヤレス給電を実現し、利便性を高めた（図11）。



図11 ワイヤレス給電機

## 4 おわりに

新型「TRACER9シリーズ」は、日常の移動から週末のツーリングまで、お客様の“モーターサイクルライフ”を飛躍的に広げるモデルとして進化を遂げた。5種類のラインナップにより、幅広い層のニーズに応えるとともに、“感動”を伴う走行体験をお客さまのもとへ届けられることを願う。

## ■参考文献

- [1] 技術紹介;進化した自動化トランスミッション「Y-AMT」の開発 [https://global.yamaha-motor.com/jp/design\\_technology/technical/presentation/pdf/browse/59gs02.pdf](https://global.yamaha-motor.com/jp/design_technology/technical/presentation/pdf/browse/59gs02.pdf)
- [2] マトリクスLEDヘッドライト 機能紹介 [https://youtu.be/Lqkzi4uWFn4?si=MzwwvERibXO39k\\_3](https://youtu.be/Lqkzi4uWFn4?si=MzwwvERibXO39k_3)
- [3] マトリクスLEDヘッドライト紹介記事 夜間走行時の安心感Up↗ 疲労感をDown↘ マトリクスLEDヘッドライトの凄さを開発者が語ります! - ヤマハ バイク ブログ | ヤマハ発動機株式会社 <https://www.yamaha-motor.co.jp/mc/yamaha-motor-life/2025/08/20250829-001.html>
- [4] 機能・装備:TRACER9 GT+ Y-AMT - バイク・スクーター | ヤマハ発動機 <https://www.yamaha-motor.co.jp/mc/lineup/tracer9gt-plus/feature.html>

## ■著者



尾上 知輝

Tomoki Onoue

モーターサイクル車両開発本部

MC 車両開発統括部

SV 開発部

# “Ténéré 700” Development Target & Objectives

Stefano Galimberti Takamasa Tokitoh Paolo Barlaam Marco Rossini

Kei Sugimoto Naoya Shiraishi Yoshitomo Akimoto



## Abstract

The “Ténéré 700”, launched in 2019, has been highly acclaimed for its unique value proposition — offering true off-road performance combined with excellent on-road versatility. For the MY25 model, development focused on enhancing rider control, off-road capability, and electronic integration, while staying true to the core values of the Ténéré family. Key technical objectives included achieving EU5+ emissions compliance, improving suspension performance, and introducing rider-selectable power modes.

These advancements further strengthen Yamaha’s leadership in the sub-1,000 cm<sup>3</sup> On-Off segment.

## 1 INTRODUCTION

We updated the well-received “Ténéré 700” model by equipping it with the EU5+ compliant “CP2” liquid-cooled 689 cm<sup>3</sup> engine, while maintaining its performance and adopting a lightweight chassis.

As a global model, it is scheduled for production and release in various countries.

Building upon the foundation of the “MY 2019 Ténéré 700”, the MY 2025 model was developed under the concept of “The Real Adventure Motorcycle for a Global Market”, aiming to achieve a high-level balance between off-road capability, on-road comfort, and electronic control features, while preserving the core values of the “Ténéré” spirit.

## 2 DEVELOPMENT TARGET

At the beginning of the development of the “Ténéré 700 MY 2025”, we revisited the reasons behind the global success of the “MY 2019 Ténéré 700” and identified the core features that resonated with riders worldwide.

The concept of “Renewal through Continuity” guided our approach: to evolve the model while preserving its essential character.

Key attributes such as compactness, lightweight design, riding ergonomics, suspension performance, and off-road capability were retained, while targeted improvements were made based on rider feedback and internal evaluations.

To enhance the riding experience, development efforts focused on the following areas:

- Riding position
- Suspension setting (front fork and rear suspension)
- YCCT (Yamaha Chip Controlled Throttle) ECU mapping
- ABS (Anti-lock Brake System) with 3 selectable modes
- TCS (Traction Control System) with ON/OFF settings

These improvements were achieved through collaborative testing and validation among YMRE (Yamaha Motor R&D Europe), YMC (Yamaha Motor Co., Ltd.), YMENV (Yamaha Motor Europe NV), and selected suppliers.

The goal was to maintain the intuitive and confidence-inspiring ride of the original model, while adapting to evolving rider expectations and regulatory requirements.

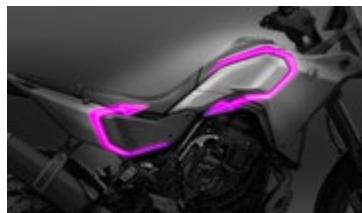


Fig. 1 Functional Styling and Design Update – Vehicle Overview: "Ténéré 700" (Model Year 2025)

C-Chamfer shape approach (Fig. 1):

- Inside this area: Shape and surface focused on rider movement
- Outside this area: Shaped for a good air flow management

Development activities were primarily conducted at YMRE in Italy in collaboration with YMC and YMENV.

Joint tests were conducted by YMRE, YMC and YMENV during this period (Fig. 2).



Fig. 2 Development Test Structure and Collaboration Framework

## 3 DEVELOPMENT CONTENT

### 3-1. Chassis

The basic chassis structure (Fig. 3) was inherited from the previous model, with two changes.

A larger muffler bracket pipe with a reinforcement patch was adopted to increase rigidity of the structure.

Additionally, a passing-through type carrier mounting boss was implemented to achieve a better load distribution.



Fig. 3 Chassis changes: Muffler Bracket and Carrier Boss

### 3-2. Suspension

Suspension performance was significantly improved to enhance off-road capability while maintaining comfort and stability during on-road riding.

The front suspension (Fig. 4) features a Ø43 mm upside-down fork manufactured by KYB, with a stroke length of 210 mm.

It includes spring preload adjustment, as well as compression and rebound damping adjustability.

Air bleeding is also supported to facilitate maintenance.

The rear suspension (Fig. 5) adopts a link-type configuration with 200 mm of wheel travel and 101 mm of suspension stroke, contributing to improved load absorption.

A lightweight and rigid full-aluminium swingarm is employed, paired with a KYB aluminium cylinder unit featuring a piggyback tank.

Compression and rebound damping are adjustable, and a convenient preload adjustment knob is provided for a quicker setup.

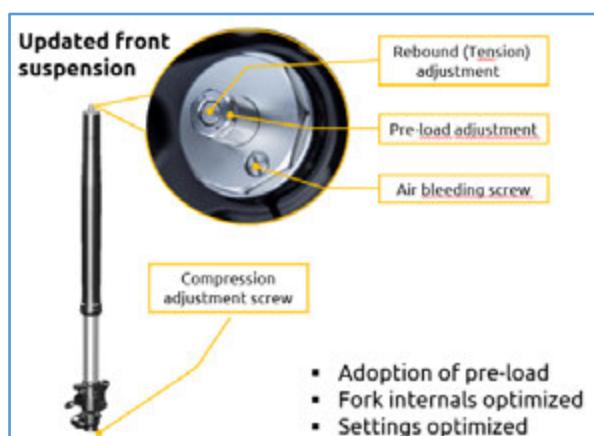


Fig. 4 Front Suspension: Configuration and Adjustment Features

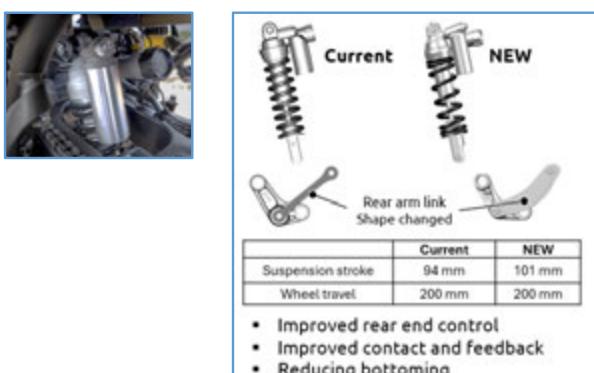


Fig. 5 Rear Suspension: Configuration and Adjustment Features

### 3-3. Wheel/Brake

The brake system and wheel configuration were carried over from the previous model, featuring a 21-inch front wheel and an 18-inch rear wheel.

This setup was retained to maintain all the excellent characteristics of maneuverability, stability, and overall vehicle dynamics across a wide range of riding conditions.

### 3-4. Off-road Oriented Features

To enhance off-road performance and rider operability, several dedicated features were newly adopted.

A one-piece rally seat (Fig. 6) combined with a redesigned fuel tank and cap improve rider mobility and vehicle integration.

The ABS (Anti-lock Brake System) is switchable across three modes, allowing preferable braking control in off-road conditions (Fig. 7).

A USB Type-C port was added to improve connectivity and convenience (Fig. 7).

An aluminium skid plate (Fig. 8) ensures protection against obstacles, with a ground clearance of 240 mm.

The knuckle visors (Fig. 9) were redesigned to improve handling.

Wide foldable foot pegs (Fig. 10) with removable rubber inserts allow adaptation to various terrain conditions.

The clutch cover (Fig. 11) was revised to improve rider leg movement, and the side stand switch was repositioned.



Fig. 6 Seat: One Piece Rally Seat and New Fuel Tank Design

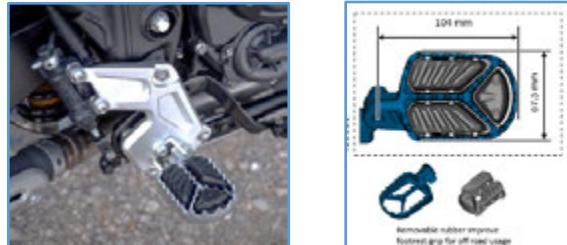


Fig. 10 Wide Foldable Foot Pegs with Removable Rubber

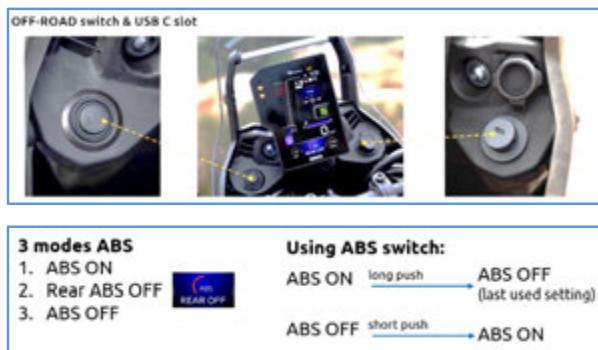


Fig. 7 Cockpit area: Switchable ABS Function and USB Type-C Port



Fig. 8 Aluminium Skid Plate and Ground Clearance



Fig. 9 Updated Knuckle Visor

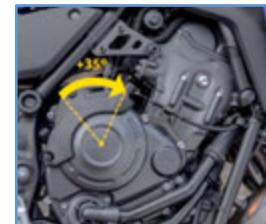


Fig. 11 Revised Clutch Cover Design

### 3-5. Engine

The "CP2" engine, a 689 cm<sup>3</sup> liquid-cooled parallel twin-cylinder unit, has been refined to comply with the latest emission regulations while maintaining its peak output of 54 kW at 9,000 rpm (Fig. 12). Torque delivery has been optimized across the entire rpm range, with specific improvements at low rpm achieved through intake duct redesign (Fig. 13), contributing to improved handling in off-road situations. Transmission gear shapes have also been revised (Fig. 14) to enhance shift feel and rider feedback.

This engine incorporates Yamaha Chip Controlled Throttle (YCC-T) technology (Fig. 15), allowing riders to select between two riding modes: SPORT (PWR1) and EXPLORER (PWR2). The SPORT (PWR1) mode is designed for riders who prefer a sharp and responsive throttle feel in all riding conditions. It preserves the snappy and exciting throttle response characteristic available in previous models, while maintaining the renowned smoothness and ease of the "Ténéré" engine family.

On the other hand, the EXPLORER (PWR2) mode is suitable for riders who enjoy a more laid-back riding style, regardless of their experience level. This mode also

provides practical support when riding on wet roads or off-road surfaces, where precise control of engine response is required, such as on loose stones, slippery ground, or muddy terrain.

The development and refinement of these power modes represents several months of joint effort by YMRE, YMENV, and YMC, with the cooperation of all involved parties enabling the achievement of the targeted performance.

The Traction Control System (TCS) offers two selectable modes via dedicated handlebar switches: TCS OFF and TCS ON. Activating TCS helps/assist vehicle control on wet or slippery off-road surfaces. The system operates smoothly and is not overly intrusive, ensuring that engine torque reduction is gradual and predictable. TCS intervenes more frequently when the rider selects the PWR-1 SPORT mode, and less so when PWR-2 EXPLORE is chosen. During development, particular attention was paid to ensuring that TCS operation would not cause any unexpected sensations or discomfort during riding.

Additionally, the optional quick shifter (Fig. 16) supports both upshift and downshift operations, providing even smoother gear changes.

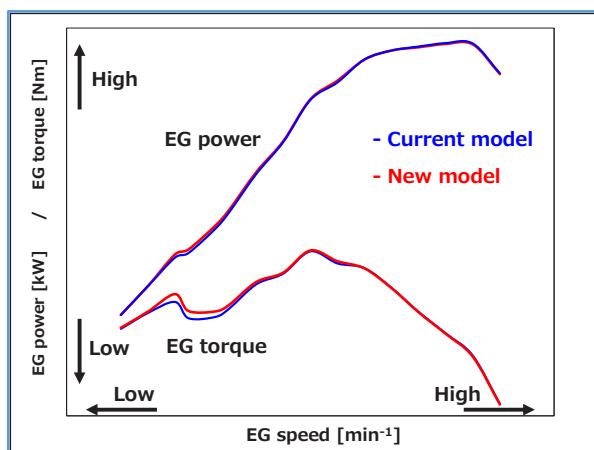


Fig. 12 “CP2” Engine: Output Characteristics and Emission Compliance EURO5+

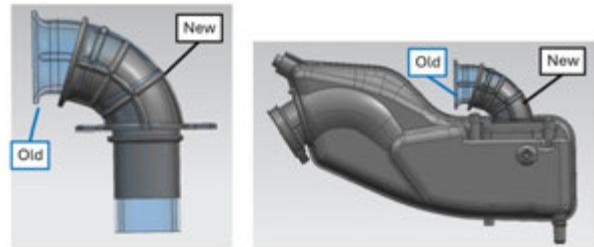


Fig. 13 Intake: Torque Improvement by Optimized Duct Shape

The number of convex dogs and concave dogs in the gears has increased from five to six (1<sup>st</sup> to 3<sup>rd</sup> gears), to reduce shocks.

The dog angle has also been changed (4<sup>th</sup> to 6<sup>th</sup> gears) to reduce vehicle behavior in response to throttle ON/OFF.

These revisions enable smoother gear shifts.

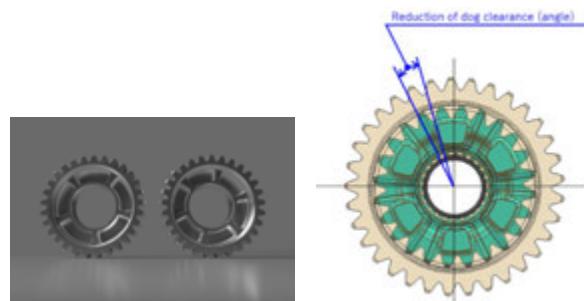


Fig. 14 Improved Shift Feel through Transmission Gear Redesign

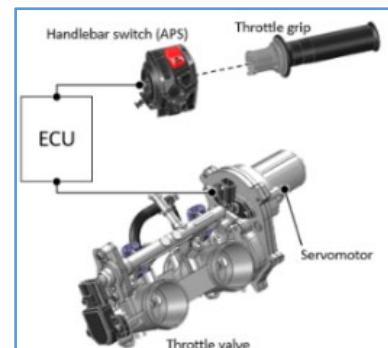


Fig. 15 Riding Mode Selection via YCC-T



Fig. 16 Quick Shifter with Up/Down Functionality

### 3-6. EL Components

To reinforce the off-road and adventure image of the vehicle, electrical components were newly designed in reference to the “DAKAR” competition model.

The headlamp (Fig. 17) adopts a distinctive layout consisting of four LED modules mounted on robust aluminum brackets, achieving both high visibility and a rugged design identity. This configuration enhances the vehicle’s presence in urban and off-road environments.

The switch handles (Fig. 18) and Human-Machine Interface (HMI) were redesigned to improve operability. A newly introduced joystick enables intuitive navigation of display menus and supports a wide range of functions with enhanced tactile feedback.

A vertically oriented 6.3-inch meter (Fig. 19) was adopted, emphasizing simplicity, visibility, and reliability. The meter allows independent adjustment of key riding parameters via the joystick, including:

- ABS: ON/Rear OFF/OFF
- TCS (Traction Control System): ON/OFF
- Engine Power Map: SPORT/EXPLORE

In addition to core riding functions, the meter supports smartphone connectivity via Bluetooth. Riders can view turn-by-turn navigation, answer calls, control music playback, and check real-time weather forecasts along the planned route.

These features contribute to a more personalized and versatile riding experience, suitable for both adventure and daily use.



Fig. 17 Four-LED Headlamp Configuration and Visibility



- ✓ Newly designed 5-way joystick
- ✓ Intuitive operation
- ✓ Auto-cancellation flashers
- ✓ Engine Mode selection – 2 available
  - ✓ SPORT (PWR-1)
  - ✓ EXPLORE (PWR-2)

Fig. 18 New Switch Handle and HMI Interface



Fig. 19 Meter, 6.3-Inch Vertical Meter and Connectivity Features

## 4 CONCLUSION

After a rigorous development process, the project met all technical and performance targets. Final specifications include a vehicle weight of 208 kg (with full fuel and fluids) and 189 kg dry, with a front/rear weight distribution of 49%/51%.

The seat height is 875 mm, with a narrow profile to improve ground reach.

The riding position remains the same as previous model, intuitive and natural. The rider can appreciate the revised design, seat, and fuel tank shape, which allow more freedom to move fwd. All package is contributing to agile handling and ease of control across a wide range of conditions.

The engine and chassis deliver a harmonious response, providing clear feedback and a controllable, satisfying riding experience.

A European press launch was held in Agafay, Morocco, where journalists praised the model's controllability on various surfaces and its distinctive styling.

The market—including both media and customers—recognized Yamaha's unique approach of maintaining a lightweight, functional, and off-road-oriented design philosophy.

This achievement was made possible through close collaboration across the entire value chain, including product planning, design, engineering, testing, business and cost planning, quality assurance, procurement, manufacturing, marketing, sales, and service, and reflects a shared commitment to delivering a product that meets the expectations of riders around the world.

The development of the new “Ténéré 700” not only reaffirmed the model's core strengths, but also provided valuable insights into rider needs, design priorities, and system integration.

These insights are expected to inform and support future product development and refinement, particularly in the areas of off-road capability, user interface, and platform versatility, contributing to the continued evolution of Yamaha's adventure and off-road strategy.

### ■著者



Stefano Galimberti

YMRE<sup>1)</sup>

PL



Takamasa Tokitoh

YMRE<sup>1)</sup>

BD Engineering  
PC



Paolo Barlaam

YMRE<sup>1)</sup>

Testing  
PC



Marco Rossini

YMRE<sup>1)</sup>

EL Engineering  
PC



Kei Sugimoto

YMC<sup>2)</sup>

Powertrain Development Center  
Product Development Section  
2nd PT Design & Development  
Division



Naoya Shiraishi

YMC<sup>2)</sup>

Powertrain Development Center  
Product Development Section  
2nd PT Testing Division



Yoshitomo Akimoto

YMC<sup>2)</sup>

Powertrain Development Center  
Product Development Section  
2nd PT Testing Division

1)Yamaha Motor Research & Development Europe s.r.l.

2)Yamaha Motor Corporation

三浦 宏信 杉本 亮平



## Abstract

Yamaha Motor develops and markets sports boats (SB) equipped with water jet propulsion systems, which generate thrust by drawing in water from beneath the hull and expelling it at high pressure toward the rear. These sports boats are classified into two categories according to their intended use.

The first is the "Family Fun (FF) model" (Figure 1), designed primarily for marine activities such as wake surfing and towing, as well as for enjoying leisure and relaxation on the water. The second is the "Family Sport Hybrid (FSH) model" (Figure 2), a center-console type boat designed with fishing applications in mind. Currently, our company offers a lineup of four size ranges — 19 ft, 22 ft, 25 ft, and 27 ft — for both FF and FSH models, and we proudly hold the No. 1 market share in the U.S. jet boat segment.

In order to further expand our product lineup and business scale — and to strengthen our presence as a premium brand within the Family Fun category — Yamaha Motor developed the 29-foot class flagship model "295XD," the largest sports boat ever produced by our company, for market introduction.

## 1 はじめに

当社は、船底から水を吸引し、後方へ高圧で噴出することで推進力を得るウォータージェット推進機を搭載したスポーツボート(以下SB)を展開している。

このSBは用途に応じて2つのカテゴリに分類される。ひとつは、ウェイクサーフィンやトーリングなどのマリンアクティビティ、水上でのリラックスタイムを楽しむことを主目的とする「Family Fun(以下FF)モデル」(図1)である。もうひとつは、釣り用途を意識したセンターコンソール艇である「Family Sport Hybrid(以下FSH)モデル」(図2)である。現在、当社は19ft/22ft/25ft/27ftの4サイズレンジで、各種FFおよびFSHモデルを展開しており、アメリカのジェットボート市場においてシェア1位を誇っている。

さらにラインナップの拡充と事業規模の拡大を図るとともに、FFカテゴリにおけるプレミアムブランドとしての存在感を一層

高めるため、当社最大となる29ftクラスのフラッグシップモデル「295XD」の市場導入に向けた開発を行った。



図1 Family Fun モデル



図2 Family Sport Hybrid モデル

## 2 開発の狙い

### 2-1. 企画コンセプト／目標

企画コンセプトは“Halo of YAMAHA”とし、フラッグシップモデルとしての位置づけに加え、ヤマハブランドを象徴する存在(Halo)となることが求められた。

その実現に向けて、以下の3項目を企画目標として設定した。

1. プレミアムな製品の実現
2. 次世代 CASE テクノロジーの搭載
3. ウェイクサーフィン機能の搭載

### 2-2. 技術目標

企画目標の達成に向けて、以下の3項目を技術目標として設定し、開発を進めた。

1. 船内レイアウトの刷新および新規フィーチャーの開発
  - ・デッキレイアウトの刷新
  - ・大型ハードトップの採用
  - ・収納式スイムシートの採用
  - ・NVH(Noise、Vibration、Harshness)の向上
  - ・大型モデルの運動性の向上
2. 最新操船アシストシステムの採用
  - ・「HELM MASTER EX」<sup>[1][2]</sup>の搭載
3. ウェイクシステムの進化

### 3-2. スタイリング／デッキレイアウト

エクステリアの特徴としては、船首に向かって伸び上がるガルネル(船体の最上部の縁)ライン、個性的な艇体側面の形状、さらに機能性を兼ね備えた大型ハードトップがあげられる(図3)。



図3 エクステリア

インテリアの主なポイントは、スタン(船尾)からバウ(船首)まで段差なく続くフロア(図4)、左舷側に通路を配置することで実現した広々としたバウエリア(図5)、そして快適な操船席スペース(図6)とスタンエリアである(図7)。



図4 インテリア



図5 バウエリア



図6 操船席

表1 主要諸元

全長(m)	9.0(バウローラー含む)
全長(m)	8.7(バウローラー除く)
全幅(m)	2.8
全高(m)	3.3
乾燥重量(kg)	約3,300(開発中の数値)
燃料タンク容量(L)	420
種類・気筒数・配列	並列4気筒 直接水冷却
総排気量(cm <sup>3</sup> )	1,812
エンジン出力(kw)	183kw(250PS)/7,500rpm × 2機掛け
ドライブ形式	φ 160mm シングルステージ軸流ポンプ ステンレス3翼インペラ
US 標準価格(25年)	\$215,749



図7 スタンエリア

### 3-3. 大型 FRP ハードトップ

デッキ広範囲をカバーする日除けエリアのため、大型 FRP (Fiber Reinforced Plastic) ハードトップを開発した(図8)。設計要件は表2に示す。

表2 ハードトップ設計要件

機能	日除け、雨除けエリア(「275SD」比 140%)
構造	外洋耐久試験に耐えること 航走中の変位が許容内であること
安定性	艇体への横安定性の影響が許容内であること
スタイリング	上部アルミ構造部材が露出していないこと ハードトップ全体を後部脚部のみで支えること



図8 ハードトップ

特に難易度が高かったのは構造要件である。外洋において、荒天状態の航走時に発生する重力角速度でも艇体強度を確認する事が求められる。外洋試験時には各方向に大きな重力加速度が発生することが過去の試験結果から想定されていた。そのため、重量物であるハードトップを支えるタワーの強度および剛性の確保は、本プロジェクトにおける重点課題の一つであった。

まず、機能要件を踏まえたうえで3D モデリングを行った。その際、アルミ製の上部構造物を FRP 製ハードトップの上下部品で覆い隠す構造とすることで(図9)、スタイリング要件も同時に満たすことができた。

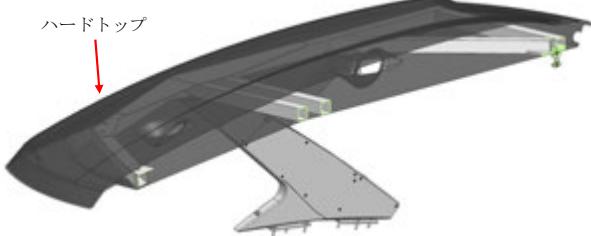


図9 ハードトップ断面

次に、3D データを用いて強度解析を、製造性、コストと軽量化のバランスを考慮しながら行い、試作品による静止荷重試験を経て外洋耐久試験に臨んだ。結果、外洋試験を一度でクリアすることができ、ハードトップの開発をスムーズに完了することができた。

性能要件については、構造部を含めて400kg 以上の上部構造を搭載することで艇体の重量重心が高くなり、その結果、動的な横安定性の低下が確認された。これに対し、船底滑走面の両側に外下がりのチャイン形状(リバースチャイン)(図10)を採用することで、船底から受ける動的な水圧を増大させ、ボートの左右方向のふらつき挙動を大幅に改善することができた。

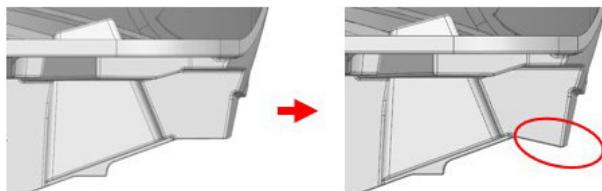


図10 リバースチャイン

市場を見渡すと、最高級モデルであっても、このような構造のハードトップは存在しておらず、一般的にはサイズが小さいもの、キャンバス生地を用いたソフトトップ、あるいはハウ两侧にもピラーを立てて支える構造のいずれかが採用されている。そのような中で、上記のすべての要件を満たすハードトップを開発できたことは、達成であると考える。

### 3-4. NVH

NVH は、顧客がボートに求める要件の中でも非常に重要な項目であり、“Halo of YAMAHA” にふさわしい高い静粛性の実現を目指して、各種の騒音低減対策を講じた。

特に騒音への寄与が大きいデッキ振動の抑制に注力し、主な加振源であるジェットポンプについてはラバーマウント化による構造見直しを実施した(図11)。これにより指示部のラバーマウントの減衰効果により振動や騒音を抑える効果を実現させた。さらに、デッキ各部の補強および積層構成の最適化を図る

ことで、構造的な振動伝達の低減を図った。また、音響解析を活用した排気構造の最適化を行い、排気管長の見直しとレゾネーターを追加し排気システムを刷新させた(図12)。特に巡航域において現行モデル比で大幅な騒音低減を達成した。以上により、フラッグシップモデルにふさわしい高い静肅性を実現することができた。

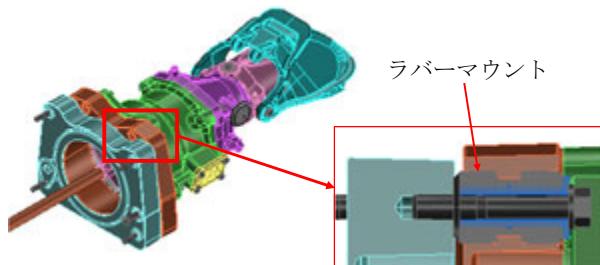


図11 ジェットポンプのラバーマウント構造

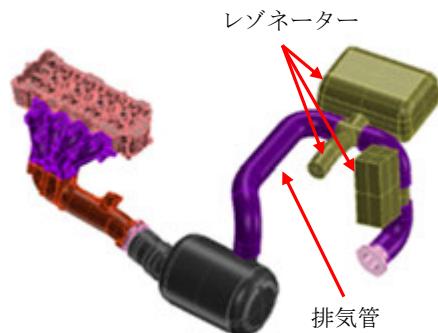


図12 新排気システム

### 3-5. 運動性能

本モデルは、現行機種「275SD」と比較して全長は約440mm(5%)の増加にとどまるが、乾燥重量は約650kg(25%)増加しており、搭載エンジンは同一であることから、加速性能および最高速力の低下が課題とされた。

これに対し、ウォータージェット推進の効率を最大限に引き出すため、船体のジェットダクト周辺の流体解析を実施し、ダクト形状およびインペラを最適化した結果、ダクト内の圧力損失が現行モデルに対し50%低減させることができた(図13)。これにより、同一エンジン出力でありながら、艇体サイズの拡大による最高速力の低下を最小限に抑え、プレミアムモデルにふさわしい加速性能(図14)を実現することができた。

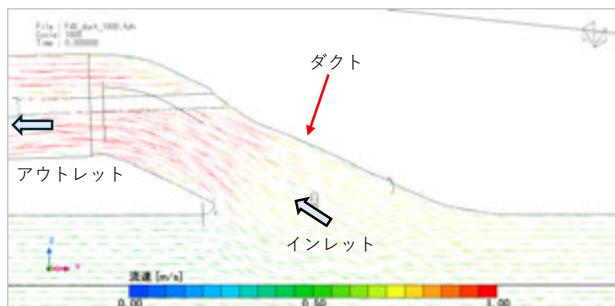


図13 ダクト流体解析

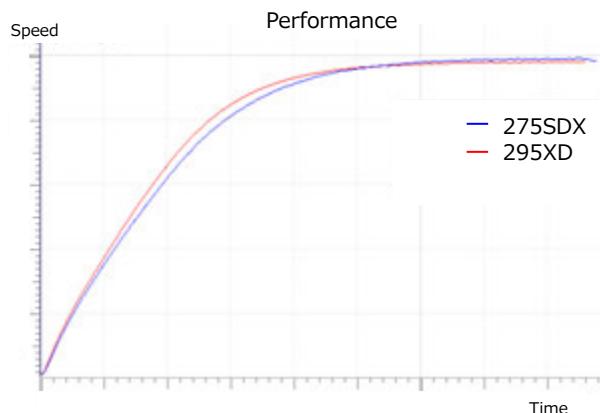


図14 加速性能

### 3-6. 「HELM MASTER EX」の搭載

離着岸は、ボート操船の中でも最も難易度の高い操作であり、テクノロジーによる革新が求められている。当社も操船アシストシステムの開発に注力しており、2023年には「DRIVE X」<sup>[1]</sup>、2024年には「HELM MASTER EX」<sup>[2]</sup>(以下、「HM EX」)を開発し、離着岸性および低速操船性を大幅に向上させた。本モデルでは「HM EX」を搭載することで、JOYSTICKによる直感的な操作性を実現した(図15、図16)。



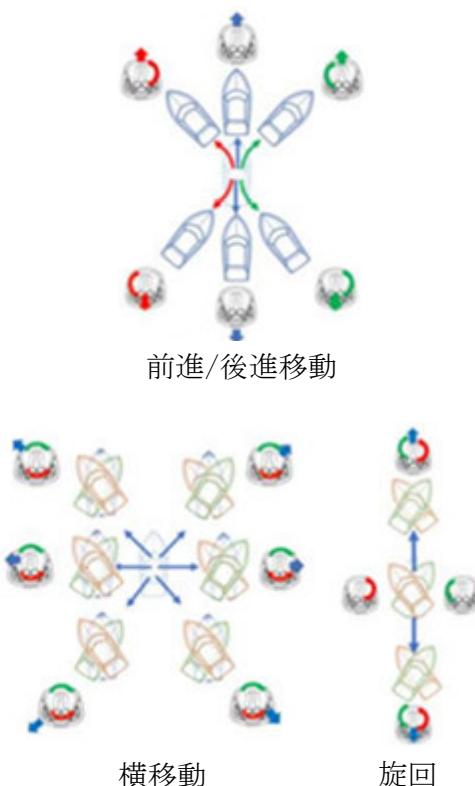


図16 「HM EX」の操船イメージ

検討にあたっては、2機掛けのエンジン間距離(以下、エンジンピッチ)がポイントとなった。着岸性能の主要因である横移動速度を最大化するには、エンジンピッチを大きくする必要がある。しかし、船底形状がV字であるため、エンジンピッチを大きくし過ぎるとデッキレイアウトの自由度が制限され、居住性や収納スペースを犠牲にしてしまう。そこでプロトタイプ艇を用いて検証を行ったところ、エンジンピッチを25ft モデルと同じ600mm としても制御のパラメーター調整で十分な横移動性能を確保できることが確認された。それにより、デッキレイアウト自由度を犠牲にすることなく、高い離着岸性と低速操船性を実現することができた。

### 3-7. ウェイクシステムの進化

ウェイクサーフィンは、1990年代後半から2000年代にかけてアメリカで本格的に普及したスポーツであり、ボートがつくる波(ウェイク)を利用してサーフボードで波乗りを楽しむものスポーツである。当社のラインナップでは25ft および22ft モデルにウェイクシステムを搭載しており、以下のコンポーネントで構成されている(表3)。

表3 ウェイクシステム構成

名称	機能
バラストシステム	船内のセンターと左右舷後方に標準搭載しているバラストバッグに水を注入することで大きなウェイクを発生させる
電動ウェイクブースター	船尾に取り付け、ウェイクの形を整える
スピードコントロール	一定のスピードで航走することでウェイク形状を安定させる

アメリカでは、ウェイクサーフィンに特化した高価格帯のボートも多数存在しており、本モデルにもウェイクシステムを搭載することが企画要求としてあげられた。

開発にあたっては、形成されるウェイクの評価指標(高さ、角度、パワーなど)を定めた上で、想定するユーザーにとって理想的なウェイクを定義した。その上で、各種コンポーネントの詳細検討を行った。具体的には、エンジンルーム内に最大限に搭載可能なバラストバッグ量(約300kg)、電動ウェイクブースター(図17)のフランジの形状と取付け角度を最適化、スピードコントロールシステムのチューニングを実施した。

これらをプロジェクト序盤から前倒しで進めたことで、各コンポーネントを高い完成度で作り込むことができた(図18)。その結果、目標である「初心者、中級者に適したウェイク」を実現することができ、より充実したマリンアクティビティ体験を可能とした。



図17 電動ウェイクブースター



図18 ウェイクイメージ

## 4 おわりに

「295XD」は、企画コンセプトを刷新し、開発を一から見直すことで、当社SBの象徴的存在、“Halo of YAMAHA”にふさわしい艇に仕上がった。今後も、さらなる進化と改良に取り組み、先進的なモノづくりを通じて、お客様に魅力ある商品を提供し続けていく。

### ■参考文献

- [1] 村山卓弥, 竹若誠人, 沼田裕貴: フラッグシップスポーツボート「275SDX」の開発, ヤマハ発動機技報, No. 57(2022)  
[https://global.yamaha-motor.com/jp/design\\_technology/technical/product/pdf/browse/57ss04.pdf](https://global.yamaha-motor.com/jp/design_technology/technical/product/pdf/browse/57ss04.pdf)
- [2] 三浦宏信, 高川 翔: スポーツボート「255 FSH SPORT H」への「Helm Master EX」搭載, ヤマハ発動機技報, No. 58(2023)  
[https://global.yamaha-motor.com/jp/design\\_technology/technical/product/pdf/browse/58ss06.pdf](https://global.yamaha-motor.com/jp/design_technology/technical/product/pdf/browse/58ss06.pdf)

### ■著者



三浦 宏信  
Hironobu Miura  
マリン事業本部  
開発統括部  
WV／SB開発部



杉本 亮平  
Ryohei Sugimoto  
Y.M.U.S.<sup>(1)</sup>

(1)YMUS: Yamaha Motor Corporation, U.S.A.

## 電動操船システム「新型 HARMO」の開発

Development of the "New HARMO" Electric Boat Maneuvering System

前島 将樹 松井 隆浩 武田 謙太郎 郷家 孝之 梶原 天翔 池ヶ谷 祐次



## Abstract

At Yamaha Motor Co., Ltd. (hereafter, "the Company"), initiatives toward carbon neutrality are advancing through electrification. In the marine field, one of the major challenges for electrification is cruising range, as both the weight and size of onboard batteries pose limitations.

To address this, the Company began developing products dedicated to low-speed applications, maximizing the inherent advantages of electric power—low-rpm, high-torque performance—while minimizing battery weight and size.

In 2021, Yamaha launched in Europe the first-generation "HARMO," an electric boat maneuvering system specialized for low speeds. It offers an unprecedented experience in which passengers can still hear the sounds of water and wind, feeling a sense of harmony with nature. Now, aiming for broader adoption of this technology, the Company has developed an evolved version—the new "HARMO."

## 1 はじめに

現在「ヤマハ発動機(以下当社)」において、カーボンニュートラルの実現に向けた電動化の取り組みが進められている。マリン分野における電動化の大きな課題は航続距離であり、船舶に搭載するバッテリの重量およびサイズに課題がある。そこで、電動の強みである低回転・高トルクを最大限に生かし、同時にバッテリの重量とサイズを最小限に抑えるべく、低速用途に限定した製品開発を開始した。2021年には、“乗船者が水の音や風の音を聞いて自然との一体感を感じられる”という、今まで体感したことがない価値提供が可能となる低速に特化した電動操船システム「HARMO」(以下初代 HARMO)を欧州で販売を開始した。今回、この「HARMO」のさらなる普及を目指し、進化させた「新型 HARMO」の開発を行った。

回頭性能”を高めるために、モータはリムドライブ方式とし、ステアリングは大舵角を可能とする一体型電動ステアリングを採用した。「新型 HARMO」として更なる進化として、「トータルチルト機能」「HELM MASTER EX」の定点保持やオートパイロットなどの快適操船機能、そしてマリン初となるジョイスティックのみシステムやジョイスティックホールド機能により付加価値を高めた製品を目指した。ターゲット市場は運河や河川などの低速遊覧(図1)であり、市場導入に先立ち、国内外での実証運航や試乗会を通じて開発仕様の妥当性の検証を行った。

## 2 開発の狙い

「初代 HARMO」の“低速を究める”の基本コンセプトを踏襲し、低速時の“圧倒的な静謐性”、“高効率”、“高推力”、“小回り”・



図1 ターゲット市場(低速用途)

### 3 商品の特徴

#### 3-1. 基本諸元

表1に電動推進機本体の主要諸元を示す。電動推進機ユニットの駆動モータは48V バッテリを使用し、システム駆動には12V バッテリを採用している。

表1 主要諸元

モデル名	H3.1B
最大出力	3.1kW
最大スラスト	F9.9船外機同等
舵角	左右70deg
ステアリングシステム	デジタルステアリング
重量	58kg
トランサム高さ	L(20inch) X(25inch)
カラー	シルバー、グレー

#### 3-2. 電動推進機本体

「新型 HARMO」では、従来船外機と同じブラケットクランプ取り付け構造とし、幅広いプレジャーボートへの搭載が可能となった。(図2)。本体カラーは、北米市場向けに人気のグレー色、欧州および国内市場向けには「HARMO」のイメージカラーであるシルバー色を設定した。

また、推進機の構成部品のモジュール化を行い、推進機本体のコンパクト化を図った。



図2 初代 HARMO(左)と新型 HARMO(右)

#### 3-3. 電動システム

電動推進機本体の内部構造を図3、4に示す。電動推進モータの基本構造は「初代 HARMO」から変更なく、インペラの外周部にステータおよびロータを配置したリムドライブ方式により高トルクと高静粛性を実現している。「新型 HARMO」では、このリムドライブモータの普及拡大に備えた設計改良と生産性改善をいくつか織り込んでいる。

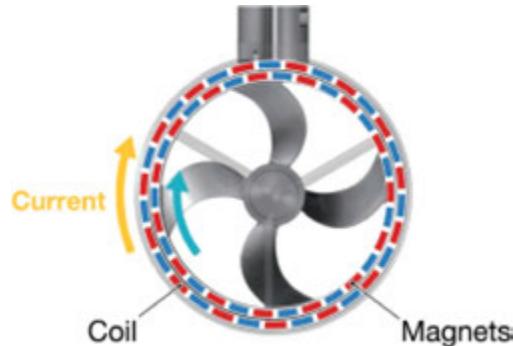


図3 リムドライブ方式

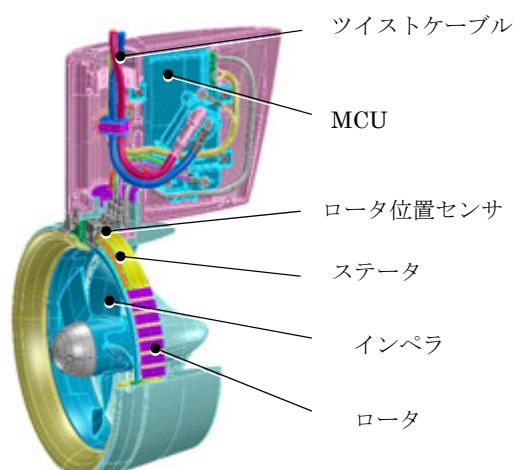


図4 リムドライブモータ、MCU の配置図

図5に示す通り、「初代 HARMO」ではステータのコイル間を接続するボードネクションと呼ばれる部品がリング状の形状であり、材料歩留まりの悪さや曲げ加工の多さが原価増の要因となっていた。一方、新モデルではモータの直並列パターンおよび巻き線仕様を見直すことで、対象部品の小型化を実現した。また、新モデルでは、メカ式電源リレーを FET(Field Effect Transistor)に変更し、MCU(Motor Control Unit)に内蔵した。

また、「新型 HARMO」では、ロータ位置検知ホールセンサがアルミ部材から受けるセンシング誤差を MCU 制御パラメータ補正により影響を抑えることでモータ効率を最大化した。そして、ツイストケーブルの採用により、転舵部の配線たるみスペースを廃止し、コンパクトなシルエットを実現した。また、水中で使用されるモータを樹脂で封止するインサート成形の量産性を高めるべく、有害な気泡の混入やガス焼けのない成形品質を目指した。図5に示すとおり、形状が複雑で微小な隙間が存在し、繊細な構成部品を有しているため、樹脂を隅々まで充填しつつ、射出成形圧による部品の変形を極限まで抑えることが品質確保上の最重要点だった。新モデルでは、成形サプライヤと当社の生産技術部門が協働し、熟練の技術と知識を最大限に

活用し、先端の樹脂流動解析を実際の成形射出圧力および金型温度と同定しながら実施し、最適な金型仕様と成形条件を決定し、これを成し遂げた。

また、ステータとロータのギャップ間への異物侵入に対する耐摩耗性を向上させるために、ステータの内円筒面およびロータの外円筒面にコーティングを施し、耐摩耗性を向上させた。コーティング材の選定にあたっては、社内採用品に限定せず、航空宇宙分野や建築構造物に採用されている材料まで視野を広げた。最終的に、樹脂への密着力と硬度に極めて優れるセラミックコーティングを採用した。

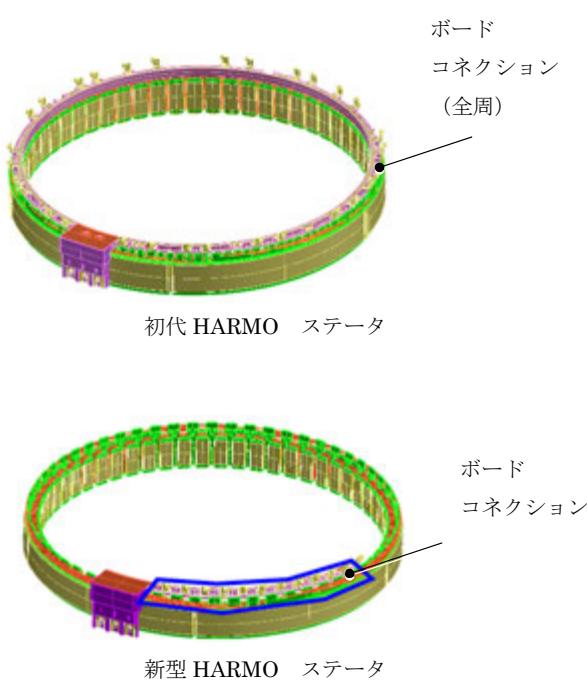


図5 ステータ構造の比較

### 3-4. 電動ステアリングシステム

#### ・電動ステアリングシステム

本モデルでは、システムの基本構成および±70degの大舵角は「初代 HARMO」を踏襲しつつ、静肅性ならびに整備性の向上を図った(図6)。

まず、操船システムの信号を受けたステアリングコントロールユニット(SCU)が、ステアリングモーターと減速機構を介して、リムドライブモータ部と一体化されたシャフトを回転させて舵切する点は初代同様である。一方、内部の減速機構をギヤボックス化し、舵切時の動き始めおよび停止時のモータ動作をチューニングすることで、ギア音の低減を図った。さらに、定期メンテナンス用のグリスニップル位置の変更や、「初代 HARMO」には搭載されていなかった「YDIS(YAMAHA DIAGNOSTIC SYSTEM2)」を追加し、整備性を高めた。

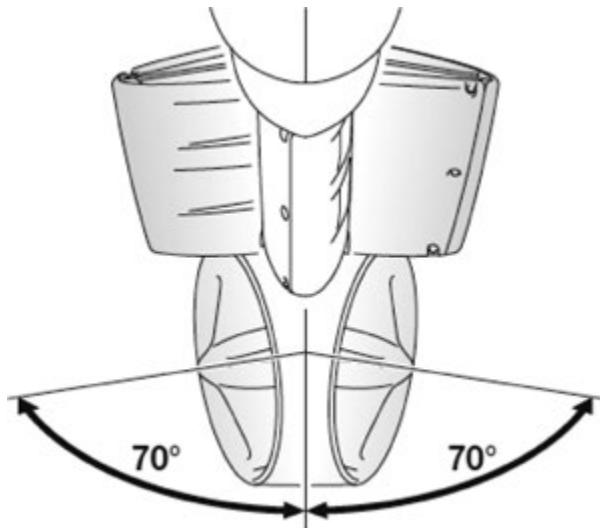


図6 電動ステアリング舵角度範囲

#### ・トータルチルト機能

大型船外機モデルのトータルチルト機能を「HARMO」向けに最適化した上で採用した。本機能を有効にすると、パワーチルトスイッチを素早く2回押すことで、スイッチを押し続けることなく、「HARMO」のチルトアップ／ダウン操作を自動で実行できる。本機能は作動時に周囲へ通知するためブザー音を鳴らすが、「HARMO」本来の静かな印象やパワーチルト動作の俊敏な印象を損なわないよう、ブザーの吹鳴パターンを変更している。また、艇体との干渉を防止するチルトリミッター機能も他モデル同様に採用しており、YDISにより設定可能である。

### 3-5. 操船システム

「新型 HARMO」では、マリン分野で初となるジョイスティックのみの操船システムを開発した。これにより、「初代 HARMO」で搭載が必須であったリモコンスロットルレバーやハンドルが不要となり、操船席周辺レイアウトをすっきりと構成することが可能となった(図7)。

#### RIGGING ACCESSORIES

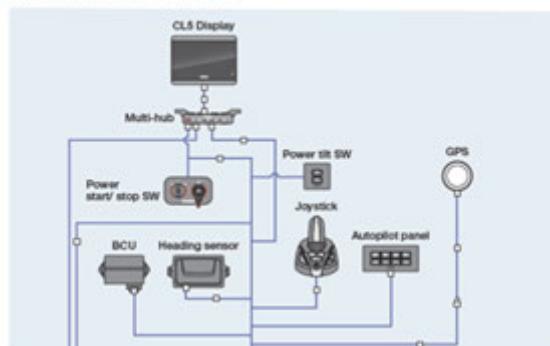


図7 ジョイスティックのみシステム構成

さらに、「HELM MASTER EX」の定点保持機能やオートパイロットなどの操船支援機能を追加し、より快適な操船性を実現した(図8)。また、ジョイスティックホールド機能については、「初代 HARMO」ではプレイヤーボートを前進させる際にジョイスティックを前方に倒し続ける必要があったが、「新型 HARMO」では自動車のオートクルーズ機能のように、設定したプロペラ回転数で航走することが可能となった。

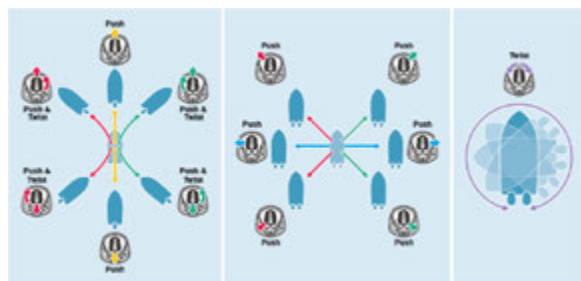


図8 ジョイスティック操作と船挙動

### 3-6. デザイン

“SMART&ELEGANCE”をデザインコンセプトとし、高い環境性能を実現する先進技術“ADVANCE”とともに、名称「HARMO」の由来であるヒト・環境・ボートとの調和“HARMONY”を意識した。これにより、クリーンかつエレガントなボートテールシリエットを通じて、ヤマハらしい上質さ“PREMIUM”を表現している(図9)。



図9 スタイリングプロポーション

### 3-7. 実証運航

開発試作機を用いて小樽運河や徳島ひょうたん島クルーズでの市場投入前の開発品質検証を実施した(図10)。実証運航終了品の分解調査し、開発仕様へフィードバックすることで、実使用環境に基づいた品質で、仕様を開発者一同が自信をもって開発を遂行することができた。

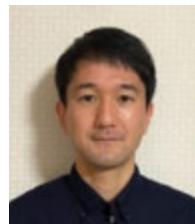


図10 小樽運河クルーズ実証運航

## 4 おわりに

「初代 HARMO」からのさらなる商品性向上と量産対応を目指し、関係部門が連携して開発に取り組んだ。設計・生産・評価の各工程において技術課題を乗り越え、実証運航による検証を経て製品完成度を高めることができた。その結果、「新型 HARMO」は生産立ち上げを達成し、マリン電動化の新たな価値を具現化する製品として結実した。“水上での静けさと快適さ”という体験価値をお客さまに届けるべく、今後も技術の深化に取り組んでいく。

### ■著者



前島 将樹  
Masaki Maejima  
マリン事業本部  
開発統括部  
先行開発部



松井 隆浩  
Takahiro Matsui  
マリン事業本部  
開発統括部  
PJT 設計部



武田 謙太郎  
Kentaro Takeda  
マリン事業本部  
開発統括部  
先行開発部



郷家 孝之  
Takayuki Gouke  
マリン事業本部  
開発統括部  
先行開発部



梶原 天翔  
Tensho Kajihara  
YEC  
システム開発部



池ヶ谷 祐次  
Yuji Ikegaya  
マリン事業本部  
開発統括部  
先行開発部



### Abstract

The “Helm Master EX” system has provided joystick steering functions that allow even inexperienced users to operate a boat easily, along with automatic position and heading control for holding a fixed point—ideal for fishing. However, on boats with structures that make it difficult to see the surroundings from the helm, or on larger vessels, even with the joystick it can be hard to judge the distance from the dock when mooring, which can cause anxiety about collisions. Additionally, even when using the station-keeping function, the captain must remain at the helm, which makes it difficult to fully enjoy activities such as fishing. The new “Wireless Joystick” combined with a Man Overboard (“MOB<sup>※1</sup>”) detection system forms the “Wireless Station”, which not only resolves these user concerns and dissatisfactions, but also frees the captain from the traditional constraint of a “fixed helm position.” By doing so, it provides a new sense of comfort, freedom, and exhilaration for boat operation.

※1 MOB: Abbreviation for *Man Overboard*, indicating a person falling overboard.

## 1 はじめに

「Helm Master EX」は、ボート操船に不慣れなユーザーでも容易に操作できるジョイスティック (Joystick) 操船機能や、釣りがしやすいよう自動位置・方位制御による定点保持機能を提供してきた。しかし、操船席から周囲を見渡しづらい構造のボートや大型艇では、ジョイスティックを用いても着桟時に桟橋との距離を把握しにくく、接触への不安を覚える場合がある。また、定点保持機能を使用していても船長は操船席に留まる必要が

あり、釣りなどのアクティビティを楽しみにくいという課題があった。本製品「ワイヤレスジョイスティック (Wireless Joystick)」と「落水検知システム (MOB<sup>※1</sup>検知システム)」を組み合わせた「ワイヤレスステーション (Wireless Station)」は、こうしたユーザーの不安や不満を解消するだけでなく、従来の“固定された操船席”という制約を取り払うことにより、船長に新たな快適性、解放感、爽快感を提供するものである。

※1 MOB: *Man overboard* の略で、船外への落水を示す。

## 2 開発の狙い

Wireless Station を構成する Wireless Joystick は、競合船外機メーカーに先駆けてヤマハ発動機株式会社(以下当社)が市場に投入した製品である。マリン業界全体では、すでに船外機向けの Wireless Joystick が販売されているが、本製品では他社にない多機能搭載による差別化を図った。具体的には、船外機のエンジン始動、パワートリム／チルト(PT/T)操作といった基本機能に加え、Joystick 操船、定点保持、オートパイロット機能などの操船機能、さらに錨の昇降、ポート照明の ON/OFF、釣果ポイントをマルチファンクションディスプレイ(MFD)に迅速に登録できるボタンなど、多彩な機能を盛り込んだ。これにより従来の操船席の様々な機能を集約した“ポータブル操船席”が完成した(図1)。

実際の使用環境を想定し、操作性、ポータブル性、長時間使用を特に重視した。①操作性：片手で全機能の操作が可能である。②ポータブル性：水中落下時の浮力を保持しつつ、可能な限り小型・軽量化を図った。③長時間使用：1日中気兼ねなく使用できるよう、リチウムイオン電池を採用し、有線／無線の2way 充電機能を搭載した。

また、さらなる安心感をユーザーに提供するため、MOB 検知システムを新規開発し、MOB タグとセットで初めて Wireless Joystick が使用可能となる仕組みを設計した。操船者が MOB タグを身に着けることにより、落水時には船外機の推進器(プロペラ)が自動停止する。

製品に対する機能要求、使用場所や使用方法に応じた洗練された設計により、機能から形状が導き出された“機能美”として、“Dynamic & Elegance”を体現した「Helm Master EX シリーズ」の統一された意匠コンセプトを踏襲することで、“ヤマハ 純正”を所有する悦びと安心感を提供する。



図1 ポータブル操船席コンセプト

## 3 商品の特徴

### 3-1. Wireless Joystick

Wireless Joystick の基本形状は、濡れた手での操作することを想定し、手のひらで物を握る際の力のかかり方を考慮した形状とした。筐体は滑りにくい材質で成形し、船上での安定した使用感を実現している。

操船者の手の大きさや利き手の違い、操作シチュエーションに左右されない中立的な操作感を提供するため、機能レイアウトを工夫した。各機能ボタンは仕切りで分け、各々が独立した形状を持たせることで、操船時に目視せざとも指先の感覚で必要な機能を探りやすくした(図2)。

情報表示画面は、操作時に手で隠れない位置に配置し、各操作モードに応じて必要な情報を適切に認識できるよう、表示内容をシンプルに構成した。高品質な LCD モジュールを採用することで、直射日光下での屋外視認性と低消費電力を両立させた。

本体は、材質・形状・構造において細部まで設計を施し、軽量化を図った。また、海上での使用において不意の落水を想定し、落水後も確実に浮上するよう設計することで、安心感を提供する。



図2 意匠形状および機能レイアウト

### 3-2. Wireless Station

Wireless Joystick の主な使用方法は、ポータブル操船席として船内を移動しながら操作することである(図3)。これにより、例えば着桟時にはボートの縁でロープを手繫り寄せる、フェンダーを出すなどの作業を行うが Wireless Joystick にてボート

の位置、方位の微調整をすることでそれら作業をより円滑に行うことができる。Wireless Joystick を片手で操作し、もう一方の手でロープを手繫り寄せる、フェンダーを出すなどの作業を円滑に行うことが可能となる。



図3 Wireless Joystick 使用シーン①

また、船首や船尾でも使用できるため、同乗者との会話や同乗者シートでの操船も可能である(図4)。特に船首での操船は視野が開け、風を感じながら操作できるため、操船に不慣れなビギナーでもボートを操る爽快感を容易に感じることができる。



図4 Wireless Joystick 使用シーン②

Wireless Joystick に織込んだ機能は以下の通りである(図5)。

- ・エンジン始動／停止
- ・船外機 PT/T 操作
- ・錨の昇降
- ・ボート照明の ON/OFF
- ・Joystick 操船
- ・Joystick Hold 機能<sup>※2</sup>
- ・オートパイロット機能
- ・定点保持機能
- ・ポイント登録機能
- ・バウスラスター連携 ON/OFF

※2 Joystick 操船時に Joystick から指を離しても、エンジンの進力を自動で保持する低速域のオートパイロット機能

#### Features of Wireless Controller



図5 Wireless Joystick 機能一覧

#### 3-3. 従来操船席の補助として

Wireless Joystick は、従来の操船席を補助する用途としても活用できる。

ボートが目的地まで長時間移動する際には、目標地点に向けて自動航行するオートパイロット機能を使用することが多い。その際、船長は操船席のシートに座るが、大型艇では操船の意思を入力するハンドルや Joystick などの操船機器とシートとの距離が離れていることがある。そのため、操船時には一度シートから立ち上がる必要があった。本製品を用いることで、スピードや進行方向の微調整入力を手元で行えるようになり、移動の手間を省き、快適な操船を実現する。

#### 3-4. クレードル

Wireless Joystick の未使用時における固定およびバッテリー充電を行うため、専用のクレードルを設計した。Wireless Joystick 本体およびクレードルは無線給電に対応しており、充電時に給電用電気接点が露出しない構造とすることで、海水や雨滴にさらされる環境下でも高い信頼性を確保している。充電

方式としては、補助的に自宅や船のキャビン内などで使用可能な USB-C 端子も搭載しており、ユーザーの好みに応じた柔軟な運用が可能である。

クレードルへの着脱においては片手操作を追求した。適切な位置に固定爪と可動爪を配置することで、固定時には本体を片手で保持したまま軽く押し付けるだけでロックが完了する。取り外し時には人差し指で可動爪を軽く動かすと、本体が自動でポップアップし、手の中に納まるため、スムーズな取り外しが可能となる(図6)。



図6 クレードル

クレードル本体も、固定・充電・取り外しといった機能に加え、船のコンソールにおける占有スペースの削減を考慮し、必要最低限の機能を必要最低限のサイズに収めたうえで、Wireless Joystick との一体感を意識した意匠とした。

### 3-5. MOB 検知システム

MOB 検知システムの主な用途は、Wireless Joystick 使用中に船長が落水した際、自動で船外機を停止させることである。MOB タグと呼ばれるセンサは、電波を発信しており、水没によってその電波が途絶え、ボート内の受信機に届かなくなることで落水と判定される(図7)。MOB タグは、腕時計型バンドとカラビナを付け替えることができ、容易に身に着けられるよう設計されている。船長用と乗客用で機能が異なるため、筐体色を分けることで誤認識のリスクを低減した。さらに、濡れた人体を傷つけにくいよう丸みを帯びた形状とし、船上での使用に最大限配慮した意匠とした(図8)。

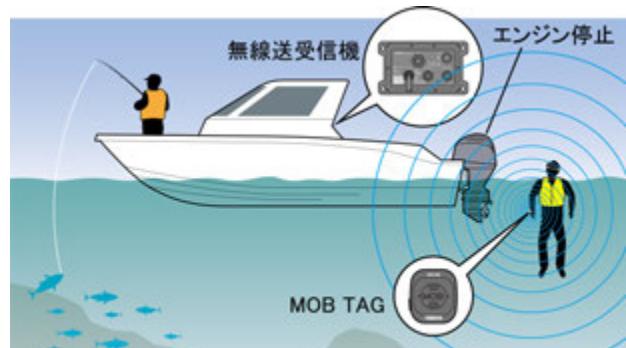


図7 MOB 検知方法



図8 MOB タグ

同乗者用の MOB タグも設定しており、落水時に船長用タグと同様のメカニズムで落水を判定する。ただし、同乗者用タグの場合はエンジンを停止せず、操船席のディスプレイや Wireless Joystick 画面に MOB 発生を表示することで、船長が同乗者の落水に迅速に気づけるよう配慮している。

同乗者用タグは最大7個まで同時接続が可能であり、同乗者落水時にエンジンを停止させないのは、レギュレーションに準じた仕様である。

さらに、船長タグが落水し、同乗者タグが通信状態にある場合には、エンジンを急停止・急減速させるのではなく、衝撃を緩和するために減速した後に停止させる“思いやり機能”も織り込んだ。

### 3-6. 安心して使ってもらうために

前述のとおり、さらなる安心感をユーザーに提供するため、MOB タグとセットで初めて Wireless Joystick が使用可能となる仕組みを織り込んだ。具体的には、以下のように BCU(ボートコントロールユニット)が無線送受信機を介して MOB タグおよび Wireless Joystick の状態を常時把握しており、MOB タグの電源が OFF の状態では、Wireless Joystick からの操作情報をアクチュエータに送信しない制御構造となっている(図9)。



■著者



鷹野 直樹

Naoju Takano

マリン事業本部

開発統括部

先行開発部



徳田 光太

Kota Tokuda

マリン事業本部

開発統括部

先行開発部



Andrew Artusa

YMUS<sup>1)</sup>

Connected Division

1) YMUS: Yamaha Motor Corporation, U.S.A.

## 4 おわりに

本製品は、競合他社に追従するのではなく、リードすることを目的として、チャレンジングな機能を盛り込んでいる。市場投入前に実施したポートショーでの発表やポートビルダー向けの試乗会において、実際に体験いただいた方々から「今までにない操船体験を味わうことができた。これはゲームチェンジャーになる!」といった非常に高い評価を得ることができた。そのため、私たちはこの“ポータブル操船席”がマリンアクティビティをより自由に快適にかつ簡単なものにし、これまで操船に自信が持てずポート利用を控えていたユーザー層にも普及していくものと確信している。今後も、難しいを簡単に、大変を楽に変えていける“ヤマハならでは”的な楽しいアイテムを創造していきたい。

# 5人乗りゴルフカー開発

## Development of a Five-Passenger Golf Car

加治屋 晋一 勝田 悠馬 宮内 泰寛 櫻木 伊織 安西 賢汰 澤村 充  
四俵 海太 池田 桂一郎



### Abstract

In 1975, "Yamaha Motor Co., Ltd." (hereinafter referred to as "the Company") launched its first golf car, the "YG-292". In 2025, the Company's golf car business will celebrate its 50th anniversary. In June of this year, the Company released two new five-passenger electric golf cars for the Japanese domestic market — the "G30Es" and "G31Eps" (see Table 1). Like the current lineup, both are available in electromagnetic induction type\* and manual-drive type versions. In addition to the quiet operation characteristic of electric vehicles, these models deliver powerful, smooth, and stable driving performance. Looking ahead to the next 50 years, these models carry forward the core concepts of the current generation — "a comfortable mobility space," "a sense of security in driving characteristics," and "outstanding operational reliability." At the same time, they achieve further evolution in driving performance and realize a higher-level balance between economic efficiency and environmental performance. The following sections outline the development objectives and key features of these models.

\* Automatic driving system that operates by following an induction wire embedded in the ground.

### 1 はじめに

1975年に「ヤマハ発動機株式会社」(以下当社)が初のゴルフカー「YG-292」を発売し、2025年にゴルフカー事業は50周年を迎える。

本年6月、当社は日本国内向けの5人乗り電動ゴルフカー「G30Es」および「G31Eps」を発売した(表1)。現行モデルと同様に、電磁誘導式\*および手動走行式の2機種をそろえ、電動ならではの静肅性に加え、力強く滑らかで安定感のある走行性を備える。

本モデルは、次なる50周年に向けて現行モデルの基本コンセプトである“快適な移動空間”、“安心感のある走行特性”、“優れた運行信頼性”を踏襲しつつ、走行性を進化させ、経済性と環境性を高次元で両立させた。以下に、開発の狙いと特徴を示す。

※埋設した誘導線に沿って走行する自動運転システム

表1 仕様諸元表

	電磁誘導式バッテリーモデル		マニュアル式バッテリーモデル			
	G30Es		G31Eps			
バッテリー容量	4kWh	6kWh	4kWh	6kWh		
型式	JB0 JB4 JB1 JB7					
駆動方式	ACモーター					
定格出力	5.0kW					
走行方式	電磁誘導		マニュアル			
乗車定員	5人					
バッテリータイプ	リチウムイオンバッテリー					
公称電圧	51.2V					
充電器方式	車載式定電流充電器					
登降坂性能	20度登降坂					
手動時走行速度	0~19km/h					
誘導時走行速度	通常8km/h 高速10km/h 低速6km/h 徐行 3km/h		—			
ブレーキ方式	油圧式4輪ディスク及びモーター回生ブレーキ					

## 2 開発の狙い

本製品は、基本コンセプトを踏襲しつつ、新設された LSM (Low Speed Mobility) 事業の起点となるモデルとして車両コンポーネントを刷新し、走行性の進化を図った。

また整備性向上、用途にあわせた駆動バッテリーの容量選択を達成するため、従来モデルでは1つの制御ユニットによる車両・走行・充電の集中制御としていたが、本モデルでは新たに分散協調型制御を採用した。

新設された車両制御ユニット(以下 VCU)とすべての制御コンポーネントは CAN 通信で接続され、集中制御している。各コンポーネントは、個別に搭載されたコントロールユニットにより、正確かつ迅速に動作する。主な利点は以下の通りである。

1. 各コンポーネントが個別に最適制御されることでシステム全体の制御性および効率が向上
2. 一部のシステム障害が全体に波及することを防ぎ、信頼性が向上
3. 機能向上や電子部品等の廃番対応などのアップデートをコンポーネント単位で実施可能となり、柔軟性が向上
4. 機能追加やバリエーションモデルへの対応において、拡張性が向上

また、外装およびフレームは従来モデルから流用しているが、刷新した制御系コンポーネントの搭載性および冷却性を確保するため、最適化を施している。



図1 VCU

### 3-2. 駆動バッテリー(内製リチウムイオンバッテリーパック)

新たに開発されたリチウムイオンバッテリーパック(図2)は、長期間使用されるゴルフカーの特性を考慮し、セルの正極材に長寿命かつ熱安定性に優れるリン酸鉄(LFP)を採用した。また、ゴルフ場の使用環境に応じて、4kWh および 6kWh の2種類から選択可能とした。

内蔵されたバッテリーマネジメントシステム(BMS)は、VCUとの組み合わせにより、バッテリーを適切に管理している。開発設計は当社が担当し、製造はヤマハモーター・パワープロダクツ株式会社が担い、四輪車の駆動用電池パックとして、当社初の完全内製化を実現した。

また整備性向上、交換性向上のため搭載されているリヤシート下のフレームを分割式とし、アクセスしやすいようにした。



図2 リチウムイオンバッテリー

## 3 新規設計コンポーネントの紹介

本モデルで刷新された主なコンポーネントは、以下の通りである。

### 3-1. VCU (Vehicle Control Unit)

車両全体を制御するコントロールユニット(図1)である。高性能マイコンを採用し、制御周期の高速化および誘導線検出回路の応答性改善を図った。また、従来アナログ回路で構成されていた誘導走行回路をソフトウェア処理に変更することで、誘導線の検出および追従性を向上させ、電磁誘導走行時の乗り心地を改善するとともに、コスト低減を実現した。これにより、複雑なカーブパス上でもより滑らかな走行が可能となった。

### 3-3. 車載充電器

上記駆動バッテリーと制御用12V バッテリーを同時に充電可能なマルチ充電器(図3)である。設定により、多様なバッテリーに対して適切な充電制御を行うよう設計されている。また、従来モデルと同等の効率を維持しつつ、パワースイッチング制御技術の進化により、スイッチング損失低減を図る Phase Shift Full Bridge 回路と交流の無効電力の最小化を図る Power Factor Correction 回路技術の導入により、力率を50%以上改善したこと、ゴルフ場の配電盤容量低減を可能とし、電気料金の削減に貢献する。



図3 車載充電器

#### 3-4. 駆動モーター

従来のDCモーターに替えて、速度およびトルクの制御精度に優れるACモーター(図4)を採用した。きめ細やかな回生制御とブラシレス構造により、車両性能を維持したまま、高効率化およびメンテナンスフリー化を実現した。



図4 駆動モーター

#### 3-5. MCU (Motor Control Unit)

前述のACモーターと組み合わせることで、力強く安定感のある走行と、きめ細やかな回生制御による高効率化を実現しているMCUである(図5)。また、フレーム下部に導風板を追加し、走行風をMCUの冷却に積極的に活用することで、従来の外装を流用しながら、刷新したコンポーネントのレイアウト(図6)を成立させている。



図5 MCU

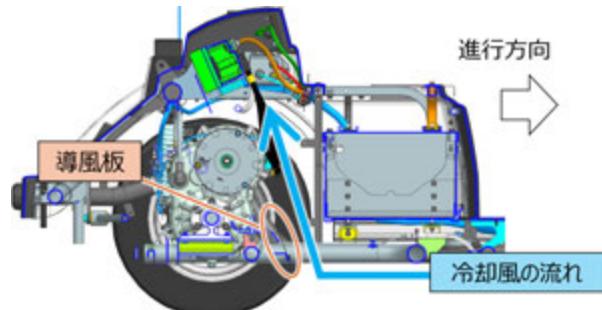


図6 冷却を考慮したレイアウト

#### 3-6. 油圧ユニット(ブレーキシステム)

新規開発の油圧ユニット(図7)はVCUおよびMCUとの協調制御により、回生ブレーキ制御と連動した滑らかで安定感のある制動を実現している。



図7 油圧ユニット

#### 3-7. アクセルペダル

新規開発のアクセルペダル(図8)は従来の金属製ペダルに対し、四輪車で使用実績のある樹脂製汎用モジュールを流用、アームとペダル部分のみを車体に合わせて最適設計した。また、スプリングには圧縮ばね2段構造を採用した。

これにより、コスト低減を図りつつ、スプリング、摺動部が樹脂ケースに内蔵される構造により給脂不要な、メンテナンスフリーを実現している。



図8 アクセルペダル

### 4 製品の特徴

前述の新コンポーネントを活用し、さらなる商品性の向上を実現した。

#### 4-1. 快適性向上

電磁誘導走行時において、新たに採用したACモーターとVCUおよびMCUの2つのコントローラの組み合わせにより、滑らかで快適な乗り心地(図9)を実現した。

VCUの高性能化に伴い、自動操舵もより精密な制御が可能となり、車両の挙動が抑えられ、搭乗者の疲労低減に寄与する。



図9 発進時の速度安定性

また、前方車両から発信される電波を検出して自動減速・停止する前方車両検知システム(図10)は、従来モデルに対し停止に必要な制動距離は同等のまま、停止直前の減速度を緩和することで特に停止時の不快な挙動を低減した。



図10 前方車両検知システム

#### 4-2. 経済性向上

ACモーターおよびMCUによる効率向上に加え、新設された駆動バッテリーとACモーター、VCU、MCUの組み合わせにより、最適な回生ブレーキ制御を実現し、消費電力を従来比30%低減<sup>※</sup>した。

これにより、バッテリー容量を従来モデルの8kWhから6kWhへ削減しながら、満充電1回あたりの航続ラウンド数を従来モデルと同等としている。さらに、重量増加を抑えることで、走行性の向上および消費電力の低下を実現し、コスト低減にも貢献している。

安定性と長寿命を兼ね備えたLFPリチウムイオンバッテリーは、新車購入後の7年長期保証を実現した。また、ブラシレス構造のACモーターと、制御バッテリーにMFバッテリーを採用することで、メンテナンスに要する手間を大幅に削減した。

※当社標準コースによる測定結果

#### 4-3. 環境性向上

新設されたコンポーネントおよびワイヤーハーネスには、鉛フリーはんだを採用し、環境負荷の低減に配慮した。また、リチウムイオンバッテリーは共同回収システムに参画することで、環境負荷への対応に加え、リサイクルにも取り組んでいる。

## 5 おわりに

ゴルフカーには、ゴルファーがプレーに専念できる快適な移動手段であること、ならびに10年以上使用されることが多いため、ゴルフ場にとって信頼性の高い車両であることが求められる。

また、リゾート地やテーマパークで使用されるランドカー、工場搬送車(SF-Ve)、グリーンスローモビリティ(GSM)などのベースモデルとして、多様な環境下で活用されている。

今回開発した新モデルは、新たに採用したコンポーネントおよび制御システムにより、走行性、制御性、信頼性、拡張性にすぐれた高いロバスト性を備えた車両であり、これらに応える製品である。

多様な場面において、利用者の期待を超える安心かつ快適な移動手段として、本モデルが貢献することを期待する。

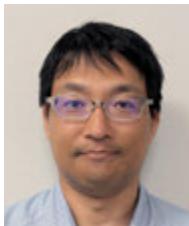
### ■著者



加治屋 晋一  
Shinichi Kajiyama  
LSM開発部



勝田 悠馬  
Yuuma Katsuda  
LSM開発部



宮内 泰寛  
Yasuhiro Miyauchi  
LSM開発部



櫻木 伊織  
Iori Sakuragi  
LSM開発部



安西 賢汰  
Kenta Anzai  
LSM開発部



澤村 充  
Mitsuru Sawamura  
モビリティシステム  
開発本部  
企画部



四俵 海太  
Kaita Shidawara  
モビリティシステム  
開発本部  
企画部



池田 桂一郎  
Keiichiro Ikeda  
モビリティシステム  
開発本部  
企画部

# 車いす用電動化ユニット「JWG-1」の開発

## Development of the Electric Drive Wheelchair Unit "JWG-1"

水谷 浩幸 松苗 徹 岩城 泰宏 水野 正光 大西 慎太郎 井口 裕貴 杉原 優  
齋藤 紘平 八木 啓明 笹岡 俊樹 林 桂太郎 糸賀 信行



### Abstract

Yamaha Motor Co., Ltd. (hereinafter "the Company") has applied its advanced control and drive technologies—developed through the creation of the electrically power-assisted bicycle "PAS"—to products that support welfare and an aging society. In 1995, the Company launched the joystick-type electric drive unit "JW-I" for wheelchairs, followed in 1996 by the world's first assist-type electric drive unit "JW-II." Since then, through full and minor model changes, these units have continuously maintained a leading share in Japan's market for simplified electric (manual-and-electric hybrid) wheelchairs.

The newly introduced "JWG-1" marks the first full model change in ten years for the joystick-type electric drive unit. This paper outlines the development overview and key features of the new model.

## 1 はじめに

ヤマハ発動機株式会社は、電動アシスト自転車「PAS」の開発を通じて培った高度な制御技術および駆動技術を活用し、福祉・高齢化社会に対応する製品のひとつとして、ジョイスティック型車いす用電動ユニット「JW-I」を1995年に発売した。翌1996年には世界初のアシスト型車いす用電動ユニット「JW-II」を発売し、その後フルモデルチェンジやマイナーチェンジを重ね、国内における簡易型(手動兼用型)電動車いす市場で常に高いシェアを維持している。

今回紹介する「JWG-1」は、10年ぶりのフルモデルチェンジとなるジョイスティック型電動化ユニットであり、本稿ではその開発概要と特長について紹介する。

## 2 商品の企画

「JWG-1」は、「もっと力強く、もっと操作しやすく、もっと自由に」をコンセプトに、「JWX-1 PLUS+」をフルモデルチェンジした新製品である。我々の商品の一番の特長である操縦性の良さを維持しつつ、国内外の幅広いユーザーに受け入れられるよう、製品の主要構成である“電動化ユニット性能”、“操作部”、“バッテリー・充電器”を大幅に改良した。

さらに、機能や性能のみならず、お客様に安心してご利用いただけるよう、強度や耐久性にもこだわり、使いやすさと信頼性の両立を目指して開発を行った。加えて、国際規格であるISO基準に準拠した“TÜV<sup>※</sup>マーク”的取得を進めるとともに、米国医療機器登録に向けて、FDA(米国食品医薬品局)認可も申請している。

※ TÜV(Technischer Überwachungs-Verein): ドイツで生まれ、行政から独立した第三者試験認証機関。製品の安全規格試験・認証業務を行う。

### 3 製品の特長

#### 3-1. 仕様概要

車いす用電動化ユニット「JWG-1」の諸元表を表1に示す。

表1 「JWG-1」の諸元表

駆動方式	後輪直接駆動
重量(バッテリー含まず)	16インチ:15.1kg、20インチ:15.6kg、 22インチ:15.9kg、24インチ:16.4kg
ユニット耐荷重 ※1	160kg
操舵方式	ジョイスティック操舵
駆動車輪径	16インチ(16×1.75)、20インチ(20×1 3/8)、 22インチ(22×1 3/8)、24インチ(24×1 3/8)
制御方式	マイクロコンピュータ制御
制動方式	電磁ブレーキ+モーター回生制動
駆動モーター (AC サーボモーター)	36V 150W ×2
手動/電動切替	16インチ:後方左右クラッチレバー切替式 20, 22, 24インチ:手元操作切替式
自走操作 最高速度 ※2 (5段階速度調節)	4.5km/h 仕様 16, 20, 22, 24インチ: 前進0.9~4.5km/h、後進0.5~2.3km/h 6.0km/h 仕様 16インチ: 前進1.1~5.7km/h、後進0.6~2.9km/h 6.0km/h 仕様 20, 22, 24インチ: 前進1.2~5.9km/h、後進0.6~3.0km/h
介助操作 最高速度 ※2 (5段階速度調節)	6.0km/h 仕様 16インチ: 前進1.1~5.7km/h、後進0.6~2.9km/h 6.0km/h 仕様 20, 22, 24インチ: 前進1.2~5.9km/h、後進0.6~3.0km/h
実用登坂角度 ※3	6°(勾配 約10%)
バッテリー (マイコン内蔵型)	(36V 6.45Ah)×1個 (5時間率)
バッテリー重量	2.4kg
充電器	定格出力42V 1.6A マイコン制御による自動充電
電動走行距離 ※4	16インチ:20km 20, 22, 24インチ:25km
タイヤバルブ	米式

●仕様、外観、および上記諸元は改良のため、予告なく変更することがあります。

※1:JWG-1ユニットの耐荷重は160kg(乗員、荷物、フレーム重量を含む)ですが、詳細はフレーム強度も含め車いすメーカーにご確認ください。

※2:最高速度は仕様諸元に基きます。

※3:転倒防止バーを次の条件で調整した場合の値です。・キャスターと後輪が地面に接している状態で転倒防止バー先端の補助輪は地面から35~50mmの高さであること。・後輪の車軸中心から転倒防止バー先端の補助輪の中心までの距離が245mm以上であること。車いすメーカーが設定する実用登坂角度が優先されますので、車いすメーカーにご確認ください。

※4:ISO 7176-4 :2008による測定方法(速度6km/h、バッテリー新品、気温18~25°C、乗員100kg)

#### 3-2. 駆動ユニット

新たに開発したモーターおよび遊星減速機により、駆動トルクは従来モデルの25.3Nmから50.1Nmへと向上した。耐荷重も従来モデル125kgから160kgへと強化した。動力性能および耐荷重の向上に加え、強度解析と実験を通じて構造を最適化することで(図1)、ユニット寸法および製品重量の増加を約5%に抑え、コンパクト化と軽量化を両立した。ホイールの耐荷重向上に伴い、軽量でありながら乗り心地と耐久性を兼ね備えた高分子複合材製ノーパンクタイヤを、メーカーオプションとして初めて採用した。

機能面では、車軸内部に新たなロック・解除機構を設けることで、ホイールの着脱操作性を向上させた。クラッチセンサーはマグネット検知式に変更し、クラッチ切り替え機構の検知精度を改善した。

走行制御については、速度追従性および耐荷重向上に伴う走行負荷範囲の拡大に対応するため、速度制御をPI制御、モーター制御を電流制御に変更した。これにより、坂道や高負荷路面での走行性能を向上させた。電動車いすは駆動左右輪の速度差によって操舵を行うため、駆動フィールと操舵フィールを別々に調整することが難しいが、熟練エンジニアのノウハウにより官能評価で高い点数を得たセッティングを実現した。さらに、走行性能の向上および電池の小型化により、電池単体での回生電力吸収が困難となったため、回生電力を抑制または消費させる“回生電力コントロール制御”を新たに導入した。

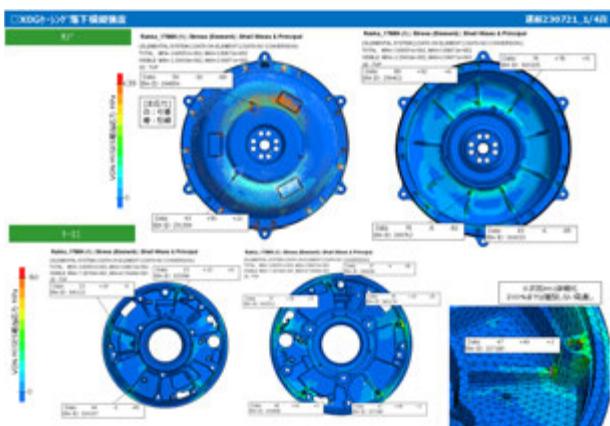


図1 強度解析結果

### 3-3. 自走操作部(搭乗者用)

自走操作部の主な変更点は、液晶画面およびデザインである。従来モデルでは画面が自走操作部の手前に配置されていたが、「JWG-1」では画面位置を前方に移し、角度を手前側に傾けることで視認性を向上させた(図2)。そして2.4インチのフルカラー液晶を採用することで表示部を鮮明化し、従来モデルでは対応できなかった警告内容の文字表示を可能とすることで、車両状態をより確認しやすくした。

さらに、電源および速度調整を従来のトグルスイッチからボタンスイッチに変更することで操作性を向上させるとともに、自走操作部のデザインも質感の高いものへと改良した。



図2 自走操作部

### 3-4. 介助操作部(介助者用)

介助操作部は、機能およびデザインの両面で大幅な変更を行った(図3)。画面表示は従来モデルのLEDバッテリー表示から、自走操作部と同様の2.4インチ液晶へと変更した。取付け位置は、従来モデルが右側ハンドルのみの対応であったのに対し、「JWG-1」では左側への取り付けも可能となり、さらに画面を縦横いずれの方向にも設置でき、向きに応じた表示変更が可能である。

最大の変更点は入力方式であり、従来の押しボタンによるオン・オフ入力から、磁気センサーによる速度調整可能なレバー入力方式へと変更した。操作性の向上に加え、緊急時にレバーを強く握り込んだ際には車両が停止する仕様としている。



図3 介助用操作部

### 3-5. バッテリー、充電器

バッテリーには、36V・6.45Ahのリチウムイオンバッテリーを採用した。エネルギー密度の高い新型セルを選定し、緻密なレイアウトおよび筐体設計を施すことで、従来品と比較して1kgの軽量化を達成した。これにより、エネルギー密度(Wh/kg)は約30%向上し、持ち運びの利便性が大幅に向上した(図4)。

連続走行距離は、20・22・24インチ仕様で25km、16インチ仕様で20kmであり、容量は使用者のニーズに十分応える仕様である。寿命末期の劣化バッテリーを使用し続けるリスクを回避するため、バッテリー監視制御を行い、使用者へ電池交換を通知するシステムを従来モデルと同様に採用した。

バッテリーの着脱構造も改良した。「JWG-1」では、持ち手の下にロック解除レバーを設けることで、持ち上げと同時にロック解除が可能となり、容易に脱着できる仕様とした(図5)。搭載位置は、さまざまな手動車いすに電動化ユニットを装着できるよう、ユニットに直接設置する固定式とバッグ収納式の2仕様を設定した(図6)。

充電器も全面的に改良した。持ち運びの利便性を高めるため、軽量・小型化を図り、「JWG-1」ではバッテリーを置いて充電するクレイドル部分を分離構造とすることで、充電器本体をよりコンパクトに設計した。さらに、クレイドルの代替としてオプション品である車載ケーブルを使用することで、車載充電にも対応可能な仕様とした(図7)。



図4 バッテリー

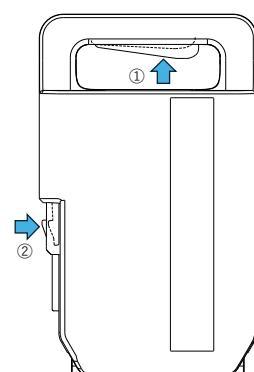


図5 バッテリーロック解除レバー



図6 バッテリー固定方式



図7 充電器

## 4 おわりに

1995年に「JW-I」が誕生してから、JW 事業 (Joy Wheel の略称であり、手動車いすを電動化するユニットの企画・開発・製造・販売を行う事業)は本年で30周年を迎える。この節目となる2025年にリリースした「JWG-1」は、各種アップデートにより、より幅広いユーザーに活用いただける製品となった。国内のみならず海外の使用者に対しても、本製品の良さを体感いただけだと確信している。今後の JW 事業のさらなる発展に向けて、新技術の獲得、新製品の開発、パートナーとの協業を推進し、“新たな感動”を世界中の方々に届けていく。

JWビジネス  
**30th ANNIVERSARY**  
—since 1995—

### 3-6. ディーラー向けアプリ

「JWG-1」では、Bluetooth 機能を標準装備し、ディーラー向けアプリの操作性向上を目的として改良を施した。従来モデルでも走行パラメータの調整が可能なアプリを設定していたが、本モデルでは速度、加速度、減速度、走行トルク、ジョイスティック操作に関して、より細かな調整が可能となるよう仕様を変更した。

さらに、従来モデルでは対応できなかった介助操作部に対しても、入力方式および応答性の調整が可能となるよう改良を加えた。故障診断サポート機能も強化し、電動ユニットおよびバッテリーの状態に関するエラー履歴、注意喚起履歴、故障診断情報などをアプリ上で即時に確認可能とし、迅速なサービス対応を実現した。

■著者



**水谷 浩幸**  
Hiroyuki Mizutani

LM 事業本部  
SPV 事業部  
プロダクト開発部



**松苗 徹**  
Toru Matsunae

LM 事業本部  
SPV 事業部  
プロダクト開発部



**岩城 泰宏**  
Yasuhiro Iwaki

LM 事業本部  
SPV 事業部  
プロダクト開発部



**水野 正光**  
Masamitsu Mizuno

LM 事業本部  
SPV 事業部  
プロダクト開発部



**大西 慎太郎**  
Shintaro Ohnishi

LM 事業本部  
SPV 事業部  
プロダクト開発部



**井口 裕貴**  
Hirotaka Iguchi

LM 事業本部  
SPV 事業部  
プロダクト開発部



**杉原 優**  
Yu Sugihara

LM 事業本部  
SPV 事業部  
プロダクト開発部



**八木 啓明**  
Hiroaki Yagi

LM 事業本部  
SPV 事業部  
プロダクト開発部



**齋藤 紘平**  
Kohei Saito

LM 事業本部  
SPV 事業部  
第2開発部



**笹岡 俊樹**  
Toshiki Sasaoka

LM 事業本部  
SPV 事業部  
第2開発部



**林 桂太郎**  
Keitaro Hayashi

MC 車両開発本部  
品質革新技術部



**糸賀 信行**  
Nobuyuki Itoga

ロボティクス事業部  
技術統括部  
SMT 第2開発部

## 新世代高速ワイヤボンダ「UTC-RZ1」

Next-Generation High-Speed Wire Bonder "UTC-RZ1"

富山 俊彦 石山 義人



## Abstract

The Assembly Solution Business Unit of Yamaha Robotics Co., Ltd. ("the Company") develops and supplies wire and die bonders. These bonders play a critical role in the assembly of semiconductor back-end process, and in recent years, there has been growing demand for production automation and equipment condition monitoring during manufacturing processes. This paper introduces the "UTC-RZ1" wire bonder, having developed in response to these needs.

The Company's history in wire bonder business dates back to 1959, when Shinkawa Ltd. was founded, focusing its business on the development, manufacturing, and sales of wire bonders. In 2019, it came under the umbrella of Yamaha Motor's Robotics Division, and in July 2025, it was integrated with other group companies to establish a new organization as Yamaha Robotics Co., Ltd.

A wire bonder is a machine that performs connection between the pads, which are the electrodes of a semiconductor integrated circuit (IC chip), and the leads of the package terminals on which the IC chip is mounted. The history of wire bonders spans about 60 years. In the early days, manual bonders were used, with operators performing positioning, bonding, and wire connection while viewing the work through a microscope.

In 1977, during the growth phase of global semiconductor manufacturing, the Company introduced the world's first fully automatic wire bonder. This innovation enabled operators to manage multiple machines simultaneously, thus greatly improving productivity.

From the 1980s onward, as semiconductor ICs had become more advanced for applications such as personal computers and mobile phones, the Company developed loop-forming technologies and bump-ball bonding processes to support finer-pitch IC chips (smaller electrode pads and increased number of electrodes) and multi-chip stacking, as illustrated in Figure 1.

In the 2000s, with the increasing capacity of compact memory cards and the rise of mobile devices such as smartphones, low-loop forming technology and overhang stacked-chip functionality were introduced. Meanwhile, following the 2008 global financial crisis, semiconductor manufacturing faced pressure to cut costs. Approximately 50%

of wire materials were replaced from gold to copper, and efforts were made to reduce frame costs. To accommodate those low-cost materials, the Company implemented multi-step bonding functions to improve bonding reliability. In the 2020s, the spread of remote work due to the COVID-19 pandemic, along with rising labor costs and difficulties in hiring manufacturing operators, drove demand for automation technologies that handle material transport and automated quality-inspection functions. In this regard, Yamaha wire bonders were equipped with maintenance monitoring and self-diagnostic functions, enabling labor-saving operation. Additionally, the spread of electric vehicles (EVs) has spurred the development of thick-wire bonding technology using copper wires of 50  $\mu\text{m}$  or larger for power semiconductor applications. Furthermore, with the growing application of generative AI, demand is increasing for wire bonding technologies supportive of next-generation devices. On the basis of this long history of semiconductor and wire bonder development, the Company now introduces its new "UTC-RZ1" wire bonder.

## 1 はじめに

ヤマハロボティクス株式会社アセンブリソリューション事業部（以下、当社）では1 Stop Smart Solution のコンセプトを基にしたワイヤボンダ、ダイボンダを世に送り出している。これらのボンダは半導体後工程での組立に関わる重要な工程であり、近年では生産の自動化や生産時の装置状態のモニタリングなどの要求が出ている。本稿ではそれらの要求に対応したワイヤボンダ「UTC-RZ1」を紹介する。

当社のワイヤボンダの歴史は株式会社新川として1959年の創立以来、ワイヤボンダの開発・製造・販売を主軸として事業を展開し、2019年にはヤマハ発動機ロボティクス事業部の傘下となった。2025年7月にグループ会社と統合し、「ヤマハロボティクス株式会社（以下当社）」として新たな体制を構築した。

ワイヤボンダは、半導体集積回路（以下 IC チップ）の電極であるパッドと、IC チップを搭載するパッケージ端子のリードとを結線する装置であり、その歴史は約60年におよぶ。黎明期には手動ボンダが用いられ、オペレータが顕微鏡を覗きながら位置決め・接合・結線を行っていた。

世界の半導体製造が成長期を迎えた1977年、当社は世界初の全自動ワイヤボンダを販売した。これにより、オペレータが複数台の装置を同時に操作可能となり、生産性が大幅に向上了。1980年代以降は、パソコンや携帯電話機などにおける半導体 IC の高度化に対応するため、図1のような IC チップのファインピッチ化（電極パッドの縮小および電極数の増加）やチップ積層化に対応するループ形成技術およびバンブボール接合プロセスを開発し、製品プロセスの多様化に対応してきた。

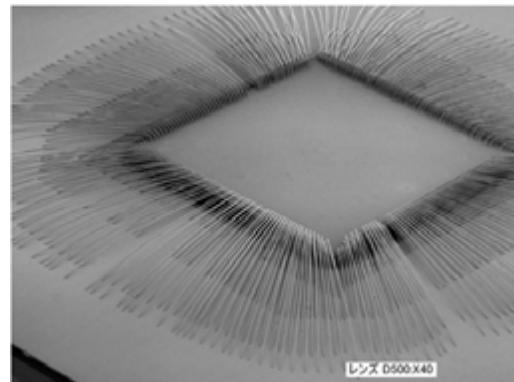


図1 ロジック向けワイヤボンドの例

2000年代には、小型メモリーカードの大容量化やスマートフォンなどのモバイルデバイスに対応した低ループ形成技術およびオーバーハング積層チップ対応機能が搭載された。一方、2008年のリーマンショック以降、半導体製造コストの削減が求められ、ワイヤ素材は金から銅へと約50%が切り替わり、フレームのコスト削減も進行した。これに伴い、低コストマテリアルに対応する接合技術として、マルチステップ機能（多段階接合機能）を導入した。

2020年代には、新型コロナウイルスの影響によるリモートワークの拡大や製造オペレータの入件費高騰・採用難といった課題を背景に、マテリアルハンドリングに対応した自動化技術や品質管理における自動チェック機能の需要が高まり、ワイヤボンダにはメンテナンスモニタや自己診断機能を搭載し、省人化に対応している。また、電気自動車の普及拡大に伴い、パワー半導体への対応技術として50um 以上の太径銅線によるワイヤボンディング技術を開発した。さらに、生成 AI の普及により、次世代デバイスに対応するワイヤボンディング技術のニーズも高まっている。以上の半導体とワイヤボンダの長い歴史を基に今回、当社の新型ワイヤボンダ「UTC-RZ1」を紹介する。

## 2 新世代高速ワイヤボンダの概要

半導体製造コスト削減の流れを受け、ランニングコストを抑え、より生産性を上げるために装置の小型化・高生産性に加え、SDGsに基づく環境対応への要求にも応えるべく、当社は小型化、省エネを目的とした新型ワイヤボンダ「UTC-RZ1」を新たに開発した。

フレームプッシャー(図2-1)の代わりに新たにフレーム引き込み爪を採用し従来機(UTC-5000)比で床設置面積を25%削減(図2-2)し、生産性を14%(\*当社標準サンプル比)向上させることで、工場における単位面積当たりの生産性を37%向上させた高生産性モデルである。図2-3は装置小型化による寸法比較とフレームプッシャーのイメージと工場レイアウト配置台数比較のイメージとなる。また、装置ランニングコストを下げるよう圧縮エアーの消費量を50%削減し、最大消費電力も12%削減することで、環境負荷の低減にも配慮した。



図2-1 フレームプッシャー

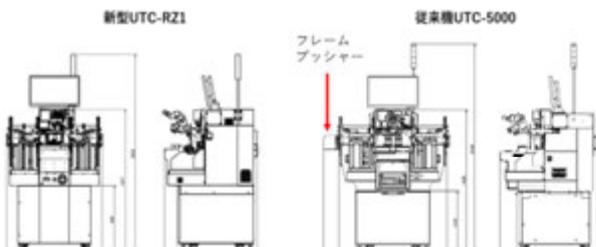


図2-2 装置寸法

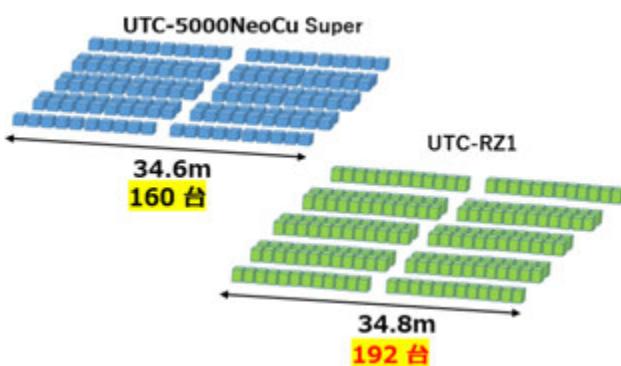


図2-3 フロア面積当たりの装置台数比較

自動化対応として、マテリアルハンドリング対応および図3の自動ツール交換機能“Capillary Auto Changer”をオプションで搭載可能し、装置のリモートオペレーションにも対応可能した。

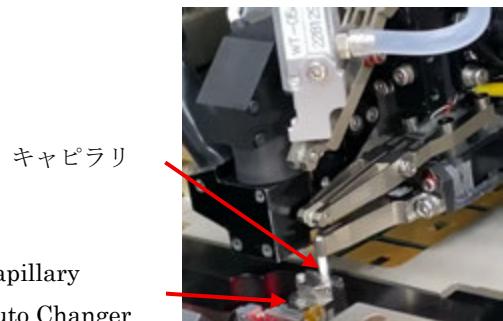


図3 キャピラリオートチェンジャー

生産品質向上のため、監視項目を従来の6項目から13項目に拡大したプロセスモニタプラス機能を新たに標準搭載し、より多くのボンディング状態の監視を可能とした(図4)。

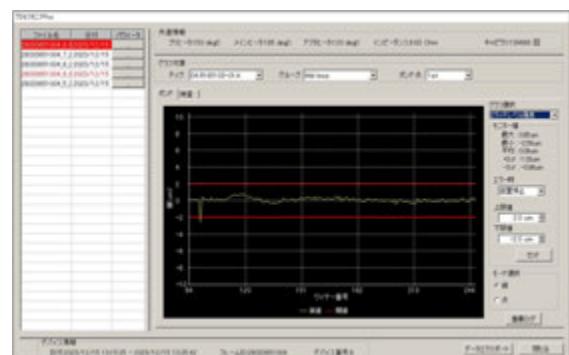


図4 プロセスモニタの画面

### 2-1. 装置内ユニットの技術紹介

XY軸リニアモータは、モータ推力を約30%向上させるとともに、ファンモータの冷却効率を改善することで発熱を抑制した。

Z軸モータは、従来のコア付き(マグネット駆動)からコアレス(コイル駆動)へと変更し、軸速度5%の高速化を実現した。加えて冷却構造を見直し、冷却用エアーの消費量を半減した。さらに、位置決め用スケールの汚染対策を施した。従来はホーンホルダを交換する際はボンドヘッドを外す必要があったが、新型ワイヤボンダはボンドヘッドを外さずホーンホルダの交換が可能な容易な構造とすることで、メンテナンス性の向上を図った(図5)。

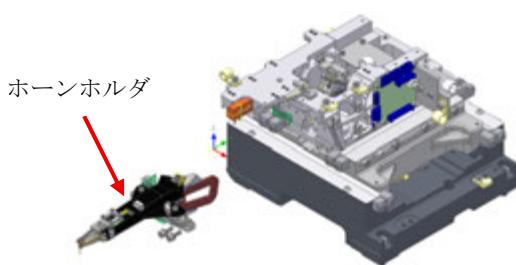


図5 ホーンホルダをボンドヘッドから外した図

ボンディングを行う超音波トランスデューサは、発振周波数をプログラム上で高周波と低周波に切り替え可能な“2周波対応”とし、さまざまなデバイスに対応できるようにした。また、キャピラリストップによる突き当て方式を採用し、“キャピラリ自動交換”にも対応可能な構造とした。さらに、発振出力を拡大することで、パワー半導体デバイスにも対応できるようにした。

電装系では、制御基板を刷新し、従来機とのソフトウェアアーキテクチャの互換性を維持しながら、機能・性能の向上およびソフトメンテナンス性の改善を実現した。

搬送系においては、ヒートステージ上の熱分布の安定化と製品切り替え時の交換部品点数の削減を図るとともに、新たに“フレーム引き込み爪”を採用し、装置幅寸法の縮小を実現した。

## 2-2. 装置実装技術紹介

ループ形成は、直感的に設定しやすい“UI(ユーザーインターフェース)”と、3軸同期制御によってループ円弧軌跡のズレを最小限に抑え、ループ高さのばらつきを低減する“SimLoopモード”を搭載した(図6)。

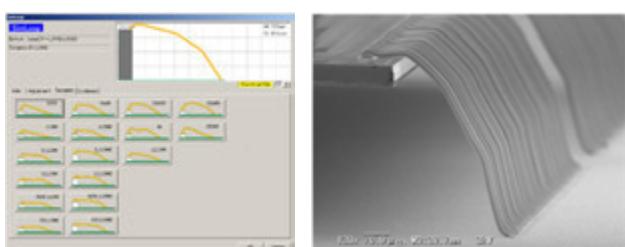


図6 SimLoop の UI とワイヤボンド例

また、テールカットにおいては、従来機能を集約するとともに、テールカットにおけるワイヤーサギング(曲がり)課題に対応するために新たに“NEXモード”を搭載し、ワイヤボンドの品質を向上させた(図7)。

Tail cut mode	Number of sagging wire (n = 630wires)	Photo
STD (SLOW 100%)	13.1% (83/630wire)	
NEX (SEQ2)	0.5% (3/630wire)	

図7 NEX モードのワイヤーサギング発生率

## 2-3. 装置ソフト技術紹介

UIは従来機の構成を踏襲し、操作におけるユーザーフレンドリー化を図った。また、各ユーザーごとに多数分かれていたソフトプログラムを集約することで、ソフトウェアのリリースまでの時間を短縮可能とした。

## 2-4. 装置開発課題

今回の装置リリースにおいて、開発で最も苦労した点は、高速化を実現するための“振動対策”である。

高速化に伴いツール先端の振動が増加する中、ベースプレートやツールを保持するボンドヘッド周辺の構造物に対し、カメラ配置の見直し、重量の低減、ボンドヘッドとYテーブル一体化における剛性(図8-1)の改善にも注力した。加えて、X軸、Y軸のモータの制御に対しても最適化を施すことで、さらなる高速化を実現した(図8-2)。

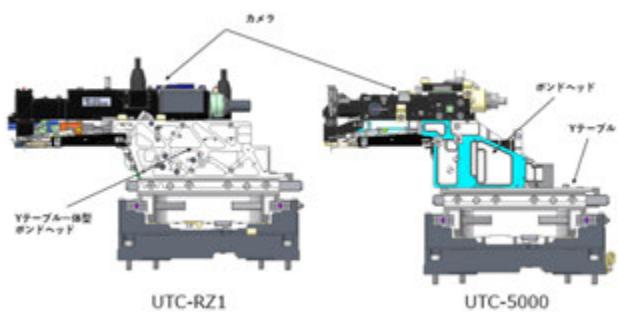


図8-1 ボンドヘッドとYテーブルの構造図

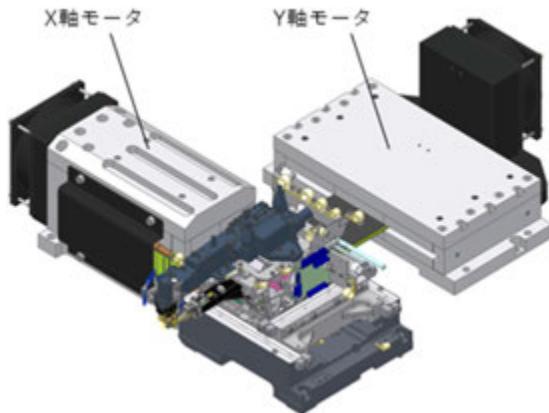


図8-2 X軸Y軸モータ構造図

### 3 おわりに

「UTC-RZ1」は、12年ぶりのフルモデルチェンジ機としてリリースされた。今回の開発では、過去の事例や蓄積されたデータをもとに基本性能を強化するとともに、次世代パッケージに対応する機能を搭載した。今後は、先端デバイスへの対応を見据え、派生機や新機能の追加によって製品力のさらなる強化を図る予定である。

製品名の「UTC-RZ1」は、ヤマハ発動機製の自動二輪車「RZ250/RZ350」をオマージュし、“Reborn(生まれ変わった)”と“Z(最高、究極)”の意味を込めた1号機と命名した。ヤマハブランドの顧客に感動を与える装置として、後継機につなげていく。

#### ■著者



富山 俊彦

Toshihiko Toyama

ヤマハロボティクス株式会社

アセンブリソリューション事業部

ボンディングソリューション技術部



石山 義人

Yoshito Ishiyama

ヤマハロボティクス株式会社

アセンブリソリューション事業部

ボンディングソリューション技術部

# ゴルフカー向け内製リチウムイオン電池パック開発

## In-house Development of Lithium-ion Battery Packs for Golf Cars

長澤 佑樹 宮崎 敏典 香田 昌伸 安藤 義崇汰 竹内 祐哉 宮田 明典 程 再冉  
三輪 千夏 笹本 侑紀 藤原 一樹 田中 洋充



### Abstract

In June 2025, Yamaha Motor launched the 5-seater electric golf cars G30Es and G31EPs, both equipped with newly developed in-house lithium-ion battery packs. To meet actual usage conditions and latent customer needs, two types of battery packs were developed in parallel: a 4 kWh and a 6 kWh model.

This lithium-ion battery pack is a fully in-house product, developed and manufactured entirely within the Yamaha Group. The development and design were carried out by Yamaha Motor Co., Ltd., while Yamaha Motor Power Products Co., Ltd. was responsible for manufacturing. This marks Yamaha's first in-house production of a traction battery pack for four-wheeled vehicles. As electrification accelerates across all forms of mobility, achieving cost competitiveness in traction battery packs has become essential. This product realized full in-house production not only through progress in development, manufacturing technology, quality assurance, and procurement strategies, but also through close collaboration with affiliated companies, leading to significant technological advancements.

As a result, compared with the previously installed battery, the 6 kWh pack achieved approximately 50% cost reduction. The following sections detail the specific technological approaches that made this possible.

### 1 はじめに

新たに開発したリチウムイオン電池パックを搭載した5人乗り電動ゴルフカー「G30Es」「G31EPs」を、2025年6月に国内市場へ投入した。ゴルフカーの使用実態および潜在的なニーズに対応すべく、4kWh および6kWh の2種類の電池パックを並行して開発した(表1)。

本リチウムイオン電池パックは、開発から製造までを一貫して社内で行った内製品である。開発設計は「ヤマハ発動機株式会社(以下当社)」、製造は「ヤマハモーター・パワープロダクツ株式会社」が担当し、四輪車の駆動用電池パックとしては当社初の内製事例となる。

今後、モビリティ全般の電動化が加速する中で、車両駆動用電池パックにおけるコスト競争力は必要不可欠である。本製品

は、開発のみならず製造技術、品質技術、調達推進、関連会社との協力を通じて内製化を実現し、技術の獲得に至った。その結果、従来搭載電池と6kWh電池を比較して、約50%のコスト削減を実現しており、その具体的な取り組み内容については、後述する各技術項目にて紹介する。

表1 仕様諸元表

項目	4kWh 電池	6kWh 電池
リチウムイオン電池種類	リン酸鉄系	
総エネルギー量 (kWh)	4.1kWh	6.1kWh
電圧 (V)	51.2V	
容量 (Ah)	80Ah	120Ah
外形サイズ	L889×W361×H228	
組電池、モジュール構成	4S2P × 4S	4S3P × 4S
重量	約63kg	約79kg

## 2 開発の狙い

本製品は、品質および信頼性を確保しつつ内製化を実現し、低コスト化を図ることを目的として、以下の5項目を重点に置いて開発を進めた。

- 1)高品質・低コストを両立する電池モジュールの実現
- 2)筐体部品および配索仕様の最適化設計
- 3)LFP(リン酸鉄リチウム)電池制御技術
- 4)電池制御システムの基板設計
- 5)電池検査およびパック組立ライン構築技術

次項より、上記5項目についての技術紹介を行う。

### 2-1. 高品質・低コストを両立する電池モジュールの実現

前述の開発の狙いを達成するため、ゴルフカー用途に適したLFP電池モジュールを開発した(図1-1)。LFP電池はレアメタルを使用せず、鉄やリンといった豊富に存在する資源を用いることで、低コストかつ安定した調達が可能である。さらに、性能・強度面および品質管理面においても、あらゆるゴルフコースやプレイスタイルに対応可能な高い信頼性を実現しており、これにより本電池モジュールをゴルフカー用途に最適な仕様へ仕上げた。

この高い信頼性を支えるために、性能・強度面に関する具体的な取り組みを行った。まず、ゴルフカーの走行データを分析した結果、コース路による頻繁な加減速や長時間連続走行が電池に与える影響が大きいことが判明した。これに対応するため、電池セルの選定においては高い電流供給と回生性能を重視しながら、容量保持性能を優先した設計方針を採用した。また、

走行中の振動や衝撃に耐えるため、モジュール構成部品の構造設計を行い、目標の耐久性を実現した。さらに、走行時および保管時の状態を正確に監視・制御するため、実測およびシミュレーションにより電池モジュールの温度分布を把握した(図1-2)。特定の電池セルにサーミスタを直接取り付けることで、温度ばらつきを抑えつつ、正確な温度計測機能を組み込んだ。これにより、性能面および強度面の最適な開発目標として設定し、電池モジュールの高性能と高信頼性設計を進めた。

性能・強度面だけでなく、安定した品質を確保するための体制整備にも注力した。調達部門・品質部門・電池セルメーカーと連携し、設計および製造段階においてFMEAおよびFTAを活用し、潜在的な故障モードの洗い出しと対策を実施した。さらに、特殊特性項目の策定や検査方法の協議を開発初期から行い、品質管理体制について協議を重ねた。

これらの取り組みにより、性能・強度面と品質管理面の双方において、ゴルフカーに最適な電池モジュールを実現した。今後は、本開発で得られた成果をもとに、品質とコストの両立を図りつつ、LFP電池モジュールを今後の電動化における中核技術の一つとしてさらに展開していく方針である。

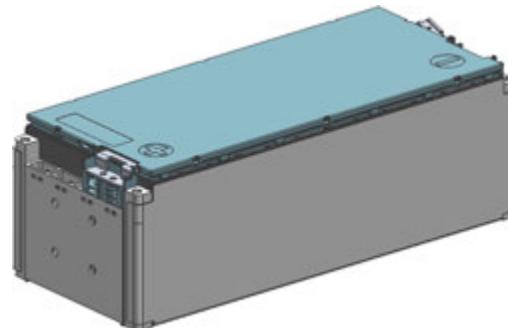


図1-1 電池モジュール

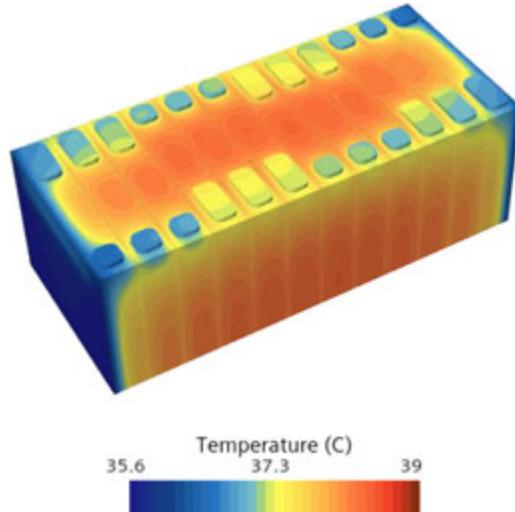


図1-2 電池モジュールの温度分布シミュレーション

## 2-2. 筐体部品と配索仕様の最適化設計

本開発では、筐体および配索仕様において電池を保護し、性能を最大限に発揮させるため、以下の3点を中心に最適化設計を行った。

### 2-2-1. 容量違い電池パック部品の共通化

車両ラインナップに応じて2仕様の電池パックを並行して開発した。電池モジュールやバスバーと呼ばれる導電部材の配置を工夫し、固定点の差異を最小限に抑えることで、電池パック全体の部品点数のうち95%以上の共通化を実現した(図2-1、図2-2)。

### 2-2-2. 板金筐体の採用

電池パック筐体は、プレス部品の多用と各締結点の最適化により、高張力鋼板を組み合わせることで軽量かつコンパクトな構造を実現した。剛性についてはシミュレーションにて強度が必要な箇所を明確化し、該当部位に集中して材料変更および剛性形状を織り込むことで、車両の振動や衝撃に耐え電池を保護できる強度を的確に確保する設計とした。

### 2-2-3. 大電流配線のバスバー化

大電流が流れる配線経路では発熱が危惧されるため、ケーブルに対して表面積が広く放熱性の高いバスバーを選定した。想定電流に応じた断面積の設定により発熱を抑制し、周辺部品の適切な耐熱グレードを設定することで、仕様の最適化を実現した。

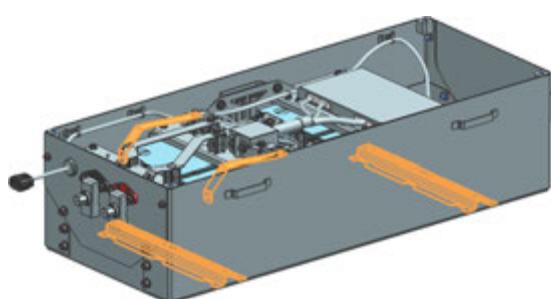


図2-1 4kWh 電池パック 共通部品(オレンジ以外共通)

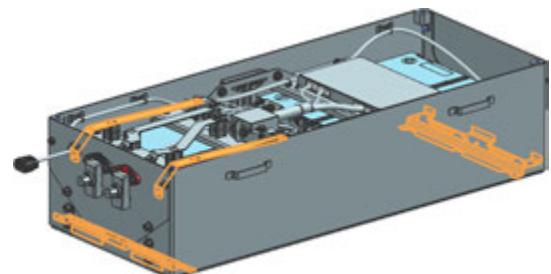


図2-2 6kWh 電池パック 共通部品(オレンジ以外共通)

## 2-3. LFP電池制御技術

現在、車載用途においてシェアを拡大しつつあるLFP電池に対応した電池制御ロジックおよび制御システムの技術開発は不可欠であった。そこで、電池制御システムの先行開発技術の応用と社内での電池評価技術の活用を踏まえ、限界利益の向上を主眼に内製開発を行った。特に商品性に寄与する充電率および劣化率の算出について、制御技術に加え、実電池による試験データの収集が重要な要素となるため、本項ではそれらの取り組みを紹介する。

リチウムイオン電池の充電率は、環境温度および放電速度(レート)に応じて変化する放電可能な電力量を基に算出される。様々な充放電試験結果から、LFP電池向けの温度とレートの相関係数を導出し、これに基づく充電率の推定制御を新たに組み込んだ(図3-1)。

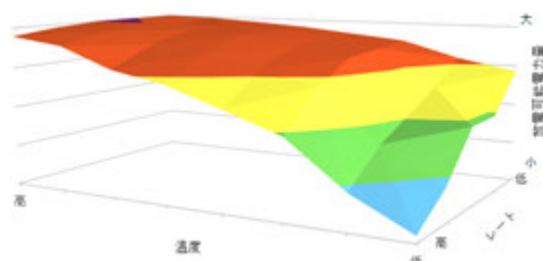


図3-1 LFP電池向け放電可能電力量

リチウムイオン電池の劣化は、充放電の繰り返しによる劣化と保管による劣化の二種類に大別され、それぞれが環境温度などの使用状況に応じて影響度が変化する。膨大なサイクル試験および保存試験のデータを解析し、LFP電池向けにサイクル数、環境温度、使用頻度、充電率を用いた多軸の劣化係数を導出する制御を開発した(図3-2、図3-3)。

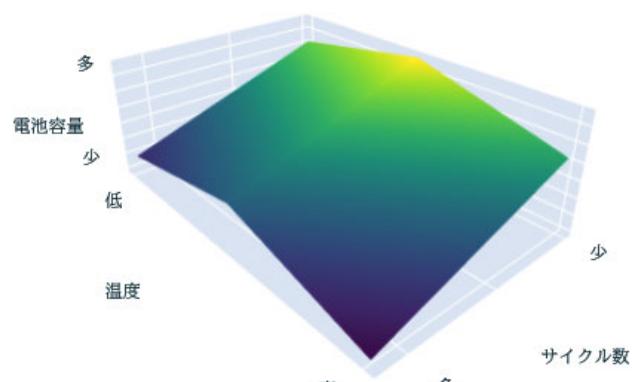


図3-2 LFP電池向け充放電の繰り返しによる劣化率

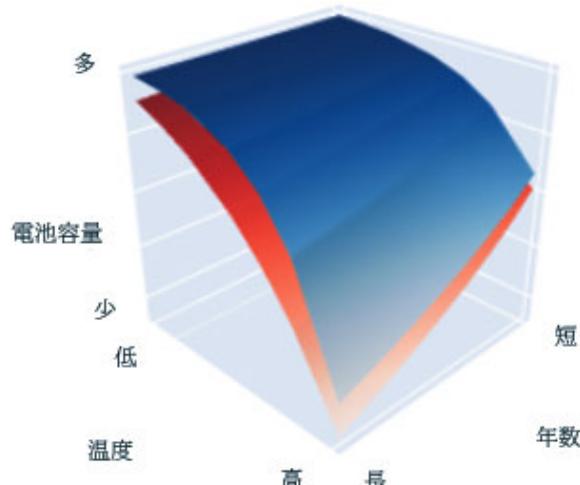


図3-3 LFP電池向け保管による劣化率  
(青:高充電率、赤:低充電率)

これらの制御により、幅広い気温条件下的ゴルフ場や多様なゴルフカーの使用状況に対応した、充電率および劣化率の推定が可能となった。

さらに、品質確保のため、基板製造およびパック組立の各工程に対応可能な共通の工程完成検査機能を組み込んだ。これにより、的確な機能検証を実施することで、開発工数の削減と品質向上の両立を実現した。

## 2-4. 電池制御システムの基板設計

本項では、前項の制御ロジックを実装するマイコンを搭載した基板設計の狙いと取り組みについて紹介する。

電池制御システム ASSY(図4-1)の回路基板は、低コスト化を目的として2層基板で設計した。2層化に伴う課題として、以下の点が挙げられる。

- 1) 基板サイズの大型化による機械強度の低下
- 2) グランドプレーンの省略によるノイズ耐性の低下

1)への対応として、基板の共振点解析技術を導入した。これにより、基板固定点の最適化および穴数の最小化を図り、生産組立時の工数削減を実現した。

2)への対応として、外来ノイズの影響を低減するため、基板の電源を電池パック自身からの内部供給とした(図4-2)。この構成により待機時の消費電流が増加する課題が生じたが、マイコンを活用した基板設計技術と漏れ電流を徹底的に排除する回路設計によって、課題の解決を行った。

また、本基板は限界利益の向上を目的として、既存の設備工程を活用する内製手段を選択した。「遠州森町工場(旧ヤマハ

モーターエレクトロニクス株式会社)」での生産にあたり、品質向上を図るべく、開発初期段階から森町製造部の各部門(生産技術、品質保証、調達部門等)と連携し、製造要件の事前設計織り込みを実施した。さらに、検査要件への円滑な対応を実現するとともに、電子部品の低コスト化や長期的な供給保証などの課題解決に取り組み、高品質な基板の安定供給を実現した。



図4-1 電池制御システム ASSY 写真

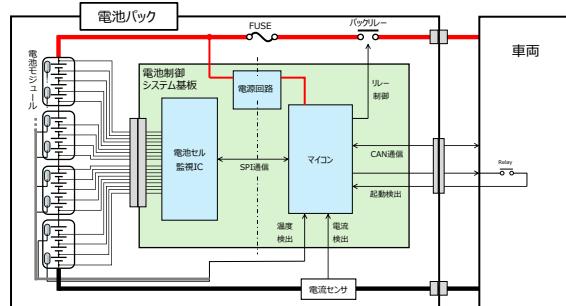


図4-2 電池システムブロック図

## 2-5. 電池検査およびパック組立ライン構築技術

全品質保証の重点目標として、電池パック組立ラインの構築を行った。

特に、量産タクトタイム内において電池モジュール内部の異常を全数検査するため、高速パルス診断法<sup>※</sup>という検査手法を導入した。導入に際して課題となっていた電圧への対応およびノイズ除去については、エンネット株式会社の協力のもと、本電池に適したソフトウェア開発を行い、解決を図った。この検査手法を組み込んだ自動検査装置の製作により、量産タクトタイム内での全数検査が可能となった(図5)。



図5 電池検査パック組立ライン

さらに、生産されたすべての電池パックに対して、電池モジュール製造メーカーの測定データおよび社内生産での測定データ、ゴルフカー製造シリアルとの紐づけを行い、トレーサビリティを確保した。これにより、市場の使用状況に応じた今後の車体開発や電池パック開発への活用が期待される。

※高速パルス診断法とは、電池内部の正極、負極、電解液といった部位ごとの異常を数分で測定できる手法である。

出典：エンネット株式会社 | 電池、デバイスおよび関係材料の開発 <https://energynet.co.jp/>

### 3 おわりに

当社の設計技術および製造技術を融合することで、内製化による高い性能、品質、信頼性とコストの両立を成す競争力の高い電動ゴルフカー向け駆動用リチウムイオン電池パックの商品化を実現した。この成果は、開発部門のみならず、社内外の製技が一体となって電動技術の具現化に挑んだ結果であり、部門の垣根を越えた協働の賜物である。

今後も電動技術を磨き続け、お客様の期待を超える新たな価値を提供し、カーボンニュートラルへの貢献を目指した技術開発を継続する。新技術・新製品・電動モビリティが描き出す当社の未来に、引き続きご期待いただきたい。

#### ■著者



長澤 佑樹  
Yuki Nagasawa  
パワートレイン開発本部  
電動開発統括部  
パワーサプライ開発部



宮崎 敏典  
Toshinori Miyazaki  
パワートレイン開発本部  
電動開発統括部  
パワーサプライ開発部



香田 昌伸  
Yoshinobu Koda  
パワートレイン開発本部  
電動開発統括部  
パワーサプライ開発部



安藤 羅崇汰  
Rasuta Ando  
パワートレイン開発本部  
電動開発統括部  
パワーサプライ開発部



竹内 祐哉  
Yuya Takeuchi  
パワートレイン開発本部  
電動開発統括部  
パワーサプライ開発部



宮田 明典  
Akinori Miyata  
品質保証本部  
ランドモビリティ品質保証統括部  
製品品質保証部



程 再冉  
Zairan Cheng  
パワートレイン開発本部  
電動開発統括部  
パワーサプライ開発部



三輪 千夏  
Chinatsu Miwa  
パワートレイン開発本部  
電動開発統括部  
パワーサプライ開発部



笹本 侑紀  
Yuki Sasamoto  
パワートレイン開発本部  
電動開発統括部  
パワーサプライ開発部



藤原 一樹  
Kazuki Fujiwara  
パワートレイン開発本部  
電動開発統括部  
パワーサプライ開発部



田中 洋充  
Hiromichi Tanaka  
ヤマハモーター・パワープロダクツ㈱  
工場技術部

# リスクアセスメントの質向上を狙ったVR活用 VR活用による設備導入前リスクアセスメント 精度向上

Use of VR to Improve the Quality of Risk Assessment  
Enhancing the Accuracy of Pre-Installation Risk Assessment through VR Utilization

上川 和輝 曽根 沙季

## Abstract

In recent years, the importance of risk management in occupational safety has increased significantly. Anticipating potential risks and implementing corresponding mitigation measures have become essential. In particular, when introducing new equipment, if there is a high risk of occupational accidents or health hazards, countermeasures such as modifications to the structure or specifications are required. Since discovering risks after equipment completion often leads to costly rework or retrofitting, it is crucial to identify potential risks in advance and incorporate preventive measures before installation.

Meanwhile, XR (Cross Reality) technology, which integrates the real and virtual worlds, has been attracting attention. XR is a collective term encompassing VR (Virtual Reality), AR (Augmented Reality), and MR (Mixed Reality), and in the manufacturing industry it is expected to serve as a means to reduce development time and costs. Details of these technologies will be discussed later in this document.

This paper describes a case study on conducting risk assessments using VR during the introduction of new equipment in the assembly process, along with the resulting effectiveness of this approach.

## 1 はじめに

近年、労働安全におけるリスクマネジメントの重要性が高まっており、リスクの予見とそれに基づく低減措置は不可欠である。特に設備導入に際しては、労働災害や健康被害のリスクが高い場合、構造や仕様の変更などの対策が求められる。設備完成後にリスクが判明すると、改造による手戻りが発生しやすいため、導入前にリスクを抽出し、対策を織り込むことが重要である。

一方、現実世界と仮想世界を融合するXR(クロスリアリティ)技術が注目されている。XRはVR、AR、MRなどを包括する総称であり、製造業においては開発期間やコストの削減手段として期待されている。これらの技術の詳細については後述する。

本稿では、組立工程における新設備導入時に、VRを活用したリスクアセスメントの実施とその効果の事例について述べる。

が望ましい。

一般的なリスクアセスメントの手順は以下の通りである<sup>[1]</sup>。

1. 危険性・有害性の特定
2. 労働災害の重篤性、頻度、可能性からリスクを見積り
3. 優先度を設定し、リスク低減措置を検討
4. リスク低減措置の実施と結果の記録

当社では、リスクレベルが5段階中3以上の場合、許容できないリスクとして低減措置が必須となる。

したがって、設備導入後のリスクアセスメントにおいてリスクレベル3以上の指摘があれば、設備改造が必要となり、手戻りが発生する。現状では、設備導入前に生産・保全部門とリスク評価を実施しているにもかかわらず、導入後に問題が発覚する事例が少なくない。

その原因として、現状の事前リスク評価が設備図面のみに基づいており、既存設備や製品などの情報を付加しにくい点が挙げられる。特に、サイズ感や作業性に関わる部分では実際の作業イメージを持ちにくく、図面と現物、担当者間で認識のギャップが生じ、リスク抽出漏れが発生しやすいという課題がある(図1)。

## 2 現状分析

リスクアセスメントとは、労働災害や健康被害のリスクを抽出し、評価および対策を施す一連の手順であり、事故災害の未然防止に有効である。事業者にはリスクアセスメントの実施と、その結果に基づく措置が努力義務として求められ、労働者も職場のリスクや安全上のルールを明確化できるため、積極的な参画

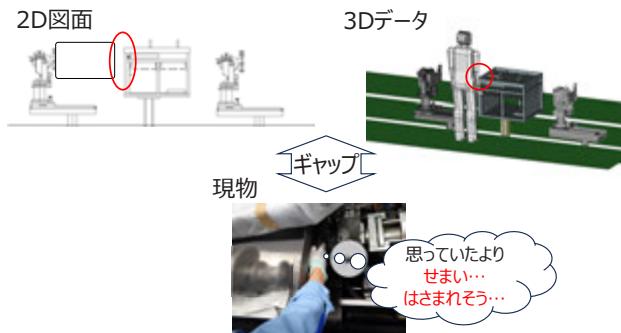


図1 認識のギャップのイメージ

トフォンなどのスマートデバイスを通して実際の風景を見ると、その中に映像や画像を映し出すことができる。

MR (Mixed Reality: 複合現実) は、現実世界と仮想世界を密接に融合させる技術である。VR と AR の中間的な技術であり、現実世界の情報を共有する点で、仮想世界に入り込む VR とは異なる。

事前評価における課題である、生産部門目線でのリスク抽出精度向上には、参加者全員が共通の作業イメージを共有することが重要である。また、新規ライン導入の検討段階であることから、現物が未完成の段階でも検証が可能であるという優位性を踏まえ、“VR”を活用することとした。

### 3 狙いと目標

本取り組みの狙いは、設備製作におけるリスク抽出精度の向上と、対策の織り込みである(図2)。特に、作業者の使用条件や作業内容に依存し、認識のギャップが生じやすい部分において、関係者が共通の作業イメージを持つことが、リスク抽出精度の向上およびリスクアセスメントの質向上につながると考えた。目標は“XRを活用した事前リスクアセスメント技法の確立”とし、具体的な目標を以下に設定した。

- 目標1. 事前安全確認時のリスク抽出と対策織り込み率100%  
目標2. 実機安全確認会時のリスクレベル3以上項目0件

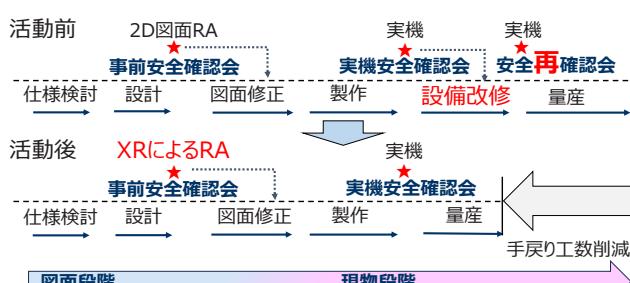


図2 XR活用による設備導入フローの変化イメージ

### 4 活用技術の選定

ここではXRに含まれる“VR”“AR”“MR”的特徴について説明する(表1)<sup>[2][3]</sup>。

VR(Virtual Reality: 仮想現実)は、CG技術を用いて仮想世界を構築する技術である。VR用HMD(ヘッドマウントディスプレイ)を装着することで、“現実にはない世界”や“体験しがたい状況”を、現実に近い立体的な環境で体験できる。

AR(Augmented Reality: 拡張現実)は、現実世界にデジタル情報を重ね合わせ、現実世界を拡張する技術である。スマ

表1 XRの特徴

	Real	MR	Virtual
使用デバイス	AR ・スマートフォン/タブレット ・ARスマートグラス	MR ・スマートグラス ・MR用HMD (ヘッドマウントディスプレイ)	VR ・VR用HMD (ヘッドマウントディスプレイ)
体験する空間	デバイス上の拡張された現実世界	現実世界+仮想世界	仮想世界
見え方の例			
特徴	現実に仮想を反映させ拡張現実の手前に仮想が配置される(前後関係を反映できない)	現実に仮想を反映させて現実世界に仮想世界を融合(前後関係が正確)	現物がベースで遠近感の評価が可能 現物を活用した評価向き 3Dデータのみで検証可能 現物がない検討初期や大規模改修時の検証向き
現実のもの	見える	見える	見えない
現物要否	必要	必要	不要
サイズ感	○	○	○
作業姿勢	×	○	○
新規ライン導入検討	×	△	○
既存ライン改修検討		○	△ (既存ラインに対するサイズ感のみ検証可)
工程作業性検証	×	○	○
検証の準備負荷(データ、実機準備)	○	△	○ (既存データ変換不要)

### 5 事前リスクアセスメント技法の確立

#### 5-1. VR空間における高さ調整

事前リスクアセスメントの目的を達成するには、参加者が実際の作業イメージを持つことが重要である。検証の結果、VR空間と実作業における高さ感覚に乖離があると、作業姿勢や視認性が異なり、正確なリスク評価が困難になることが判明した。したがって、リスクアセスメント用途でVRを活用する際には、VR空間と現実世界の目線高さおよび作業高さに注意を払う必要がある。今回、VRによるリスクアセスメントを実施するにあたり、高さ基準柱を設け、VR空間内の作業高さと現実世界の作業高さを一致させることで、高さ方向の位置を補正した(図3)。



図3-1 指摘事項の確認

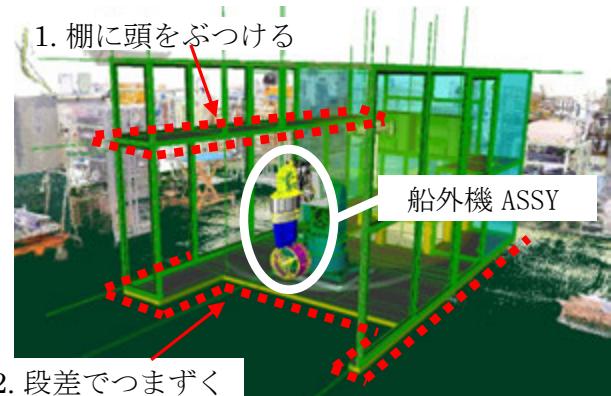


図4 指摘事項

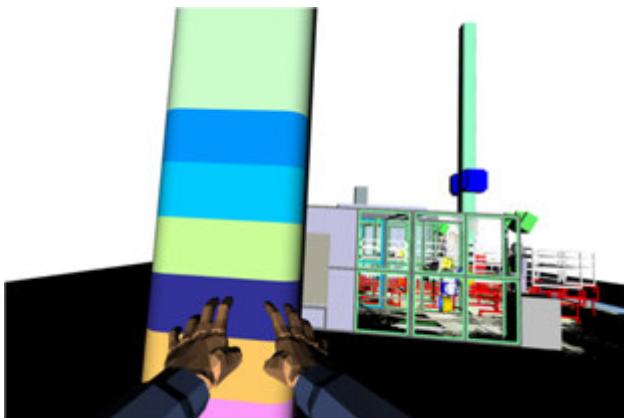


図3-2 VR空間内高さ調整 基準柱を設定

これにより、仮想空間内での作業高さを実作業と一致させることができるとなり、VR実施時に実作業と遜色ないレベルで作業イメージを体感できるようになった。また、VR空間の作業状況を周囲のメンバーが確認する際にも、仮想空間内の作業高さを一致させることで、作業イメージを正確に把握しやすくなる点で効果があった。

## 5-2. リスク指摘事項と対策事例

VRを活用したリスクアセスメントにより抽出されたリスク例として、船外機メインASSY組立機の事例を以下に示す(図4)。

### 1. 棚への頭部衝突リスク:リスクレベル3

指摘事項:作業中に落下部品を拾うなどでしゃがみ、立ち上がる際に設備内の棚に頭部をぶつけることで頭部裂傷を負う可能性がある

### 2. つまずきリスク:リスクレベル3

指摘事項:作業エリアにおいて設備床板と工場床面との段差があり、作業中につまずいて転倒し、捻挫や打撲を負う可能性がある

これらのリスクに対し、以下の対策を実施した(図5)。

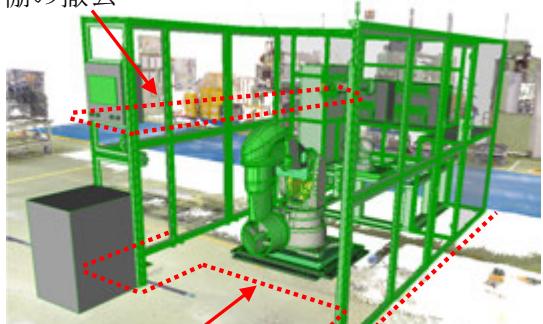
### 1. 棚への頭部衝突リスク:リスクレベル3→2

対策内容:頭部衝突の可能性が高い位置にある棚を撤去し、棚下部にクッション材を貼付。

### 2. つまずきリスク:リスクレベル3→2

対策内容:設備床板を撤去し、工場床面を基準として作業エリアをフラット化。

### 1. 棚の撤去



### 2. 設備床板を撤去し、作業エリアのフラット化

図5 指摘事項に対する対策

## 5-3. 實施結果と効果

設定した目標に対する結果は以下の通りである。

まず、目標としていた事前安全確認時のリスク抽出と対策織り込み率100%を目指した取り組みは、設計・図面化された領域について100%達成することができた。

具体的な成果として、事前安全確認会の中で8件のリスクを抽出することができた。そして、そのうち2件はリスクレベル3以上であった。それらの安全上の設備指摘事項を設備仕様にあらかじめ織り込みすることにより、設備完成後の想定修正費用を40%削減し、改造工期も1カ月短縮することができた。

図面によるリスクアセスメントでは、ロボット動作中の挟まれ

リスクなど、設備の安全装置や仕様検討に関する指摘が中心であり、人の作業内容に依存するリスク抽出が不足していた。今回VRを活用することで、“作業中”を想定したリスクが抽出することができたと考える。従来の図面ベースのリスクアセスメントと比較して、関係者間で共通の作業イメージを共有できた点も有効であった。

次に、リスクレベル3以上項目0件を目標とした実機安全確認会においては、リスクレベル3以上の要修正項目が2件発生し、目標は未達となった。なお、指摘事項は通常の設備設計時には詳細に製図を行わない設備配線に起因するつまずきリスクであり、今回のVRの適用範囲外であった。事前のVR活用により、設備発注前に8件の対策を織り込み、安全対応設備改造を防止できることから、一定の効果があったと評価できる。

## 6 今後の展開

今回、袋井南工場の新規ライン構築において、“VR”を活用した設備導入前のリスクアセスメントを実践した。本取り組みにより、VRは実物がない早期段階において、関係者間で作業イメージを共有する手段として有効であることが示された。

なお、今回のVR検証は静止状態の設備を対象として実施したが、アニメーションを付加することで自動運転時の設備の部品組付動作まで検証可能であることが確認されている。設備導入前のリスクアセスメントに動きの要素を加えることで、設備稼働時の挟まれリスクや、大型・重量物の搬送工程における作業者の心理的圧迫感を抽出・改善する施策の織り込みが期待できる。これにより、リスクアセスメントをはじめとした設備導入検討のさらなる質向上につながるものと考えている。

## 7 おわりに

今回、リスクアセスメントの質向上を目的として、VRを活用し、工場内外の関係者に広く体験の機会を設けた。特に、日常業務で3D技術に触れる機会の少ない製造部門から寄せられた前向きな反応や、技術進化への期待の声が印象的であった。リスクアセスメントの質という観点では、実際に作業する人の声は不可欠であり、今回の取り組みはそのきっかけとして有意義であったと考える。一方、VRによる空間認識効果は大きいものの、重量物を取り扱う際の力学的負荷など、作業者への身体的な影響は評価できない。将来的な展望として、VR空間における部品の取り置き作業とアシストスーツの連携による作業負荷の体感が期待されている。また、今回のVRシステムでは、操作者自身の身体モデルを取り込むことができないため、評価は主

に目線ベースで行われている。現時点の技術水準においても、身体モデルの取り込みは実現可能と考えられるが、これに組み合わせて作業姿勢毎の人間工学的負荷を数値的に評価する技術が確立されることで、作業現場における負荷低減を通じた労働災害のさらなる低減に貢献できると期待される。今後も新技術を活用し、工場一丸となって安全な職場構築を牽引していきたい。

### ■参考文献

- [1] 厚生労働省「職場のあんぜんサイト」: <https://anzeninfo.mhlw.go.jp>
- [2] 日本政策投資銀行「AR/VRを巡るプラットフォーム競争における日本企業の挑戦」: [https://www.dbj.jp/topics/investigate/2021/html/20211129\\_203602.html](https://www.dbj.jp/topics/investigate/2021/html/20211129_203602.html)
- [3] キヤノン ITソリューションズ「XR(クロスリアリティ)とは? VR・AR・MRの違い」: <https://www.canon-its.co.jp/solution/mr/vr-ar-mr/>
- [4] ラティス・テクノロジー「XVL Studio VR オプション(XVL VR)」: <https://www.lattice.co.jp/products/lineup/xvl-studio/xvl-vr>

### ■著者



上川 和輝  
Kazuki Kamikawa  
マリン事業本部  
製造統括部



曾根 沙季  
Saki Sone  
マリン事業本部  
製造統括部

## Abstract

On production floors that use surface-mounting machines to mount electronic components and manufacture electronic circuit boards (hereafter referred to as “SMT floors,” where SMT stands for Surface Mount Technology), several challenges have emerged — including labor shortages caused by declining birth rates and a shift away from manufacturing careers, rising labor costs due to global increases in wages and prices, and variations in work quality caused by differences in worker skill levels. These challenges pose significant risks to customers not only in expanding their businesses but even in maintaining and continuing them. To address these issues, Yamaha Motor Co., Ltd. (hereafter, “the Company”) has been developing automation solution technologies aimed at achieving unmanned SMT floors. This report introduces the envisioned future image of SMT floors that the Company aims to realize through these technologies.

## 1 はじめに

電子部品を実装し、電子基板を生産する表面実装機を用いた生産フロア(以下 SMT フロア (SMT: Surface Mount Technology(表面実装技術)))においては、少子化や製造業離れによる労働力不足、世界的な賃金・物価高騰による人件費の増加、作業熟練度の差による作業品質のばらつきなどが課題として挙げられる。

これらの課題は、お客さまにとって事業の拡大のみならず、維持・継続にも大きなリスクとなる。ヤマハ発動機株式会社(以下 当社)では、これらの課題解決を目的として、SMT フロア無人化に向けた自動化ソリューション技術の開発を進めている。本報告では当技術により実現を目指す、未来の SMT フロア像を紹介する。

## 2 SMT フロアの目指す姿

## 2-1. 現状の課題

多くの人員を必要とする SMT フロアの現状の課題として、電子部品などの資機材が点在すること(図1)、各電子部品補給作業が突発的かつ同時に要求されること、さらに生産装置の高速化による補給やゴミ回収などの作業量の増加が挙げられる。

これらの現状の課題に対応するため、人員は SMT フロア内を不定期に各所へ移動・待機を繰り返し、時間毎に不定量な作業が求められている。こうした同時多発的かつ不定量な作業に迅速に対応するには、訓練や学習を経て一定の熟練度を持つ人員を複数配置する必要がある。さらに、作業量の増加により、追加の人員配置が必要となる。



図1 SMT フロアに点在する作業と資機材、人員

## 2-2. 課題を解決した SMT フロアの姿

前述の課題を解決し、人員が不定期に各所へ移動・待機することなく、時間毎に定量の作業を行い、配員後すぐに一定の熟練度で作業を継続できる、すなわち、“定点”“定期”“定量”および“作業品質の安定化”を実現することが重要であると考え、そのための手段を開発することとした。

SMT フロアには、表面実装機による生産に必要な印刷機や検査機など、各種装置に対する作業が存在する。の中でも、特に多くの人員配置が必要となる表面実装機への電子部品補給作業の課題を優先して解決することとし、本稿ではこの補給作業に焦点を当てて記載する。

## 3 目指す姿の実現に向けて

## 3-1. 定点の実現

表面実装機への電子部品補給作業は、補給用の電子部品が梱包されたテープリールを部品供給器(図2)に組付ける作業、補給を必要とする表面実装機まで部品供給器を運ぶ作業、電子部品を使い終わった部品供給器を表面実装機から抜き取り、

運んできた部品供給器を挿し込む、補給作業などを行う。

これらの作業を“中間作業エリア”(図3)と呼称する場所に集約し、また、部品供給器を運ぶ作業と部品供給器を表面実装機から抜き取り及び挿し込む作業を自動で行う“チェンジャー”(図4)と呼称するロボットを配備する。

電子部品補給に関わる作業を集約した場所、すなわち“定点”で実施することにより、人員は移動や待機に費やしていた時間を作業に充てることができ、効率化が可能となる。さらに、人員と作業の集約により、管理の簡易化など、追加の効率化要素も見出すことができる。



図4 チェンジャー外観

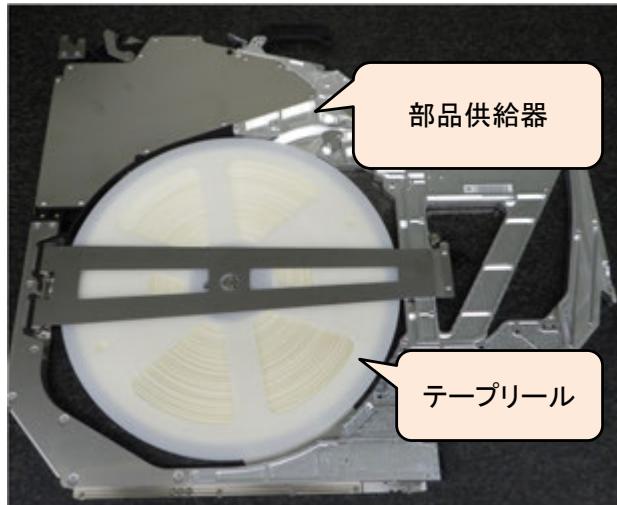


図2 テープリールを部品供給器に組付けた状態

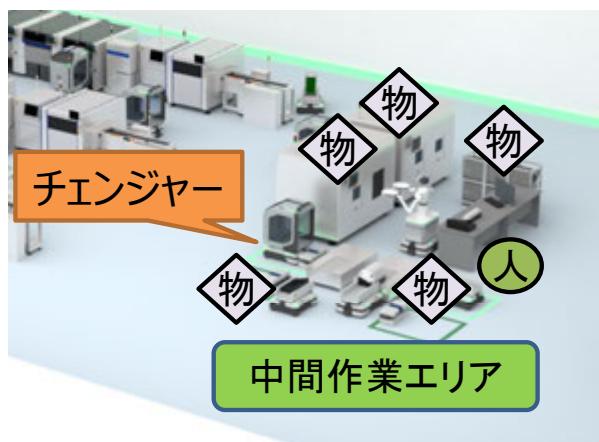


図3 中間作業エリア

### 3-2. 定期および定量の実現

表面実装機には、形状や機能の異なる複数の電子部品が搭載されており、部品ごとに使い切られるタイミングが異なる。これが、突発的かつ同時に補給作業が発生する要因となっている(図5)。



図5 フロア各所での部品切れ発生

この補給タイミングを“定期”とする手段として、“生産計画を基にした補給シナリオを作成するシステム”および“事前に補給可能な機器構成”を実現する。これにより、特定の時間帯における補給作業を“定期”とすることが可能となる。

さらに、補給用の部品供給器を複数格納し、表面実装機が要求するものを自動選別してチェンジャーに受け渡す機能を有する“デリバリーステーション”(図6)と呼称する装置を中間作業エリアに配置する。これにより、当作業を“定量”とすることが可能となる。

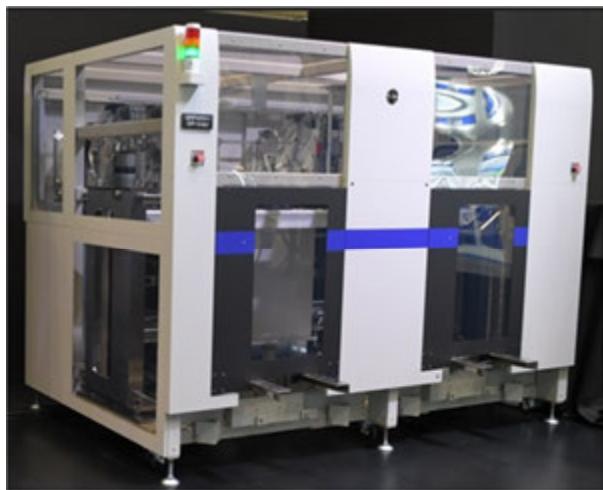


図6 デリバリーステーション外観

また、デリバリーステーションに一定数の部品供給器を格納できるスペースを設けることで、格納された部品供給器が尽きるまで、無人で生産を継続する時間を確保することも可能となる。

このように実現された“定期”・“定量”作業は、人員の作業計画立案を容易にし、作業効率の向上および人員削減に大きく寄与すると考える。

### 3-3. 作業品質の安定化

前述の“定点”“定期”“定量”作業の実現により、人員は“どこへ行けばよいのか”“何が、どれだけ、いつまでに必要なのか”を探したり、判断したりする必要が従来よりも少なくなり、高い熟練度を求められる場面を減らすことが可能となる。また、エンジニアはロボットであり、熟練した人員のように状況に応じた柔軟かつ迅速な思考や作業はできないが、導入初日から特定の作業を一定品質で実行でき、増員時に育成期間を要さない。

これにより、複数の生産ラインにおける作業品質の確保が可能となる。

### 3-4. 目指す姿の実現

最小限の人員で、図1に示すようなSMTフロアを運営できること。これが目指す姿である(図7)。

SMTフロアには、先に紹介したテープに梱包された電子部品のほかに、対応すべき項目として、トレイに梱包された電子部品や電子基板の補給や、基板回路に合わせてハンダを塗布する印刷機など、資材補給を要する他の装置があげられる。

これらに対しても、同様の構想により作業人員の削減を図る。さらに、残存する人員についてもロボットや装置、システムへの置き換えによる無人化を視野に入れ、段階的な開発を進めている。

また、段階的な開発であることを活かし、お客様のニーズに応じた自動化提案を部分的に導入可能であり、それぞれの異なる要望に対して最適な提案を行うことができる。



図7 少人数で運営されるSMTフロア

## 4 おわりに

多くの人員や育成期間を確保せずとも、生産のさらなる効率化、維持、事業拡大を実現することが可能である。その手段として、無人化に向けた自動化ソリューション技術の提供に向け、開発を進めている。

本ソリューション技術により、多くの人が生産における価値の高い作業、新製品の開発、さらなる事業創造に向けて、より多くの時間を確保できれば、それは新たな夢や感動を生み出すと信じている。そのような未来を実現することが、感動創造企業としての使命である。

ロボティクスを活用し、人々の可能性を拓げ、より良い生活と社会の実現を目指す。当社らしいソリューションとして、絶えず挑戦を続けながら開発に取り組んでいく。

### ■著者



天内 真

Shin Amanai

ロボティクス事業部

技術統括部

先行開発部

# クラウドで実現する迅速簡便なシステム検証環境の構築

Rapid and simple setup of system evaluation environment by the Cloud

落合 寛彰 吉成 雅通 人見 瑛一 浜田 善夫

## Abstract

At Yamaha Motor Co., Ltd. (hereinafter referred to as "the Company"), the IT Division is advancing strategic initiatives in IT, digital technology, and data utilization under the long-term vision "Art for Human Possibilities" toward 2030. Under the slogan "Yamaha Motor to the Next Stage," the Company promotes three key digital transformation (DX) programs: Y-DX1 (Reforming the Management Foundation), Y-DX2 (Strengthening the Present), and Y-DX3 (Creating the Future)<sup>[1]</sup>.

Through new Yamaha experiences, services, and products that blend the real and digital worlds, the Company aims to strengthen customer engagement, enhance brand value, and foster lifelong Yamaha fans. To realize this goal, the IT Division's role is to provide rapid and effective solutions that meet business needs promptly, while building a cloud-native environment as the foundation of DX activities.

This paper introduces one such initiative — the development of a rapid and easy-to-set-up system verification environment. This environment has already been adopted in several projects and is contributing as a core foundation for Yamaha's DX initiatives.

## 1 はじめに

ヤマハ発動機株式会社(以下当社)IT本部では、2030年までの長期ビジョン "Art for Human Possibilities" に向けて、IT・デジタル技術・データの戦略的活用を図るため、"Yamaha Motor to the Next Stage" を掲げ、Y-DX1(経営基盤改革)、Y-DX2(今を強くする)、Y-DX3(未来を創る)から成る3つのデジタルトランスフォーメーション(DX)活動を推進している<sup>[1]</sup>。リアルとデジタルで創出するヤマハの新たな体験・サービス・製品を通じてお客様とつながり、ブランド価値の向上を図り、生涯にわたるヤマハファンの創造を目指している。その実現に向け、ビジネス部門からの要求に迅速に応えるソリューションを提供することがIT部門の役割であり、DX活動の基盤としてクラウドネイティブな環境の実現に取り組んでいる。

本稿では、その取り組みの一つとして、迅速かつ簡便に構築可能なシステム検証環境を紹介する。この環境は既に複数のプロジェクトで活用されており、DX活動の基盤として貢献している。

定を個別に行っていたため、多くの工数とリードタイムが発生していた。そこで、クラウド技術を活用し、迅速かつ簡便で安全に技術検証を行えるシステム検証環境を構築した。

## 3 解決すべき課題

当社が新しいシステムの検証に際して抱えていた課題として、従来の構成では次の3点が挙げられる(図1)。

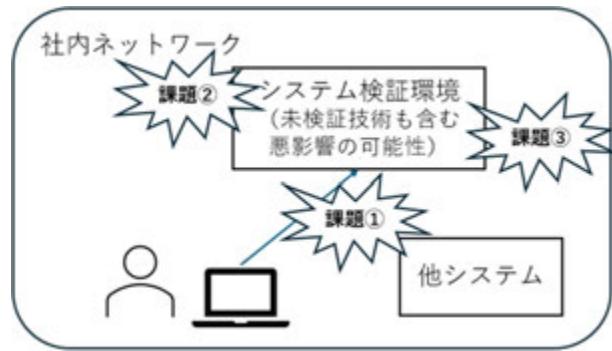


図1 システム検証環境構築の課題

### 課題①: 安全性の確保

新しいシステムを稼働させる際、他の社内システムやネットワークに悪影響を及ぼさない安全性の確保が必要である。特に、セキュアな通信経路の確保や、検証環境が他システムと隔離された構成が求められる。

## 2 背景

企業は市場のニーズに迅速に対応し、競争力を維持することが求められており、その実現手段としてクラウド技術は重要な要素である。クラウド技術は急速に進化しているが、当社ではこれまでクラウド技術を活用したシステム検証環境を提供しておらず、プロジェクトごとにインフラやアプリケーションの構築・設

#### 課題②：迅速かつ簡便な環境構築

オンプレミス(自社内に設置・運用するITインフラ)で検証環境を準備する場合、最短でも2週間を要し、その後の環境設定やアプリケーション構築には1ヶ月以上を要する。これでは、冒頭で述べた市場のニーズに迅速に対応したい業務とシステム開発との間に乖離が生じる。したがって、システム稼働に必要なインフラやアプリケーションを迅速かつ簡便に構築することが求められる。

#### 課題③：環境の再現性および変更管理の容易さ

本番環境構築時には、検証環境と同等の構成を用意する必要がある。一部設定に差異が生じる場合もあるが、安定した再現性の確保は、規模が大きくなるほど困難となる。また、構築済みの環境に対しても随時変更が発生するため、変更管理の容易さも重要な要件である。

## 4 課題解決の利用技術

3章で述べた課題を解決するために、以下に AWS(Amazon Web Services<sup>※</sup>)を中心とした6つのクラウド技術を説明する。

※Amazon Web Services、“Powered by Amazon Web Services”ロゴ、[およびかかる資料で使用されるその他のAWS商標]は、米国その他の諸国における、Amazon.com, Inc. またはその関連会社の商標です。

#### 利用技術①：クラウド(AWS)

インターネット経由でクラウド環境のサーバ、データベース、ネットワークなどの機能を利用できるサービス群である。社内ネットワークと隔離された環境であるため、社内に悪影響を及ぼすことなく、新しいシステムを安全に検証できる。

#### 利用技術②：AWS Session Manager

操作端末からサーバへ高いセキュリティで接続可能な、AWSが提供する通信技術サービスである。従来の接続方式では、クラウドサーバにインターネットからの接続口を開ける必要があり、攻撃者による脆弱性の悪用が懸念されていた。

しかし、Session Managerでは、接続口を設けることなく操作端末から接続できる。通信には暗号化されたhttpsプロトコルを用いるため、通信の盗聴や改ざんに対して高い安全性を有する。さらに、接続認証にはAWSが標準提供するIAM(Identity and Access Management)を使用し、高い信頼性と堅牢性を確保している。

#### 利用技術③：IaC(Infrastructure as Code)

クラウド上のサーバ、データベース、ネットワークなどの機能をプログラムコードとして定義・実行することで、インフラ環境を自動構築できる技術である。代表的なIaCツールには

TerraformやAWS CDKがある。IaCコードを実行するだけで環境構築が可能となり、構築時間は約1～15分程度に短縮される。

また、類似のインフラ環境が必要な場合には、過去のIaCコードを再利用することで、最小限の修正で同一環境を構築できる。

#### 利用技術④：コンテナ

システムと、その実行に必要なプログラムを独立したパッケージとしてまとめる技術である(図2)。

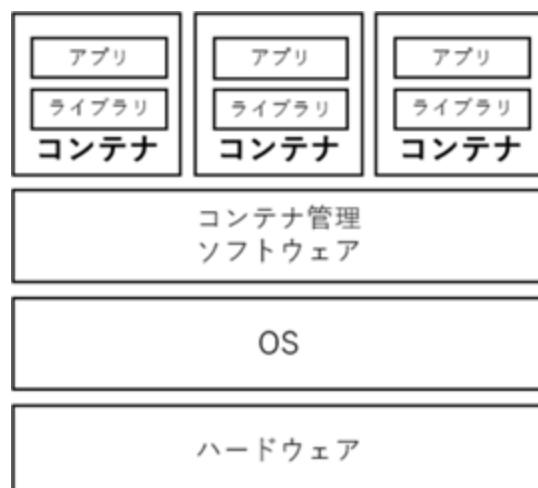


図2 コンテナの概念

システム間の独立性により、稼働の安定性とセキュリティの確保を両立し、起動の速さやメモリなどのリソース消費の抑制にも寄与する。Dockerが最も広く利用されているコンテナ技術であり、“コンテナ”という用語はDockerコンテナを指す場合が多い。

コンテナは設計図(テキストファイル)から作成され、起動は簡単なコマンドで行う。実行時間は数秒から数分と比較的短時間である。

また、コンテナを複製しても環境は同一であり、同様に動作することが保証されるため、複数のメンバーがそれぞれの環境でシステム技術を評価する際にも、環境の差異に起因する問題が生じにくい。

#### 利用技術⑤：DevContainers

Microsoftが提供するエディタ“VSCode”内で呼び出される機能であり、システム動作環境全体をDockerコンテナ内にカプセル化する技術である。複数のメンバーでシステム検証を行う場合、設定ファイルを共有することで、全員が同一の環境に同一の方法で接続して使用できるため、環境差異による問題を抑止しつつ、開発スピードの向上が可能となる。

### 利用技術⑥: GitHub

ソフトウェア開発プロジェクトにおいて、主にバージョン管理とソースコード管理を提供する SaaS 型サービスである。コンテナの設計図(テキストファイル)、DevContainers の設定、IaC のプログラムコードの変更管理やバージョン履歴の管理に使用される。バージョン履歴を活用することで、過去の任意の時点のバージョンにいつでも戻すことが可能である。

これらのクラウド技術を組み合わせたシステム検証環境(図3)により、図1に示した課題①～③を解決することができた。

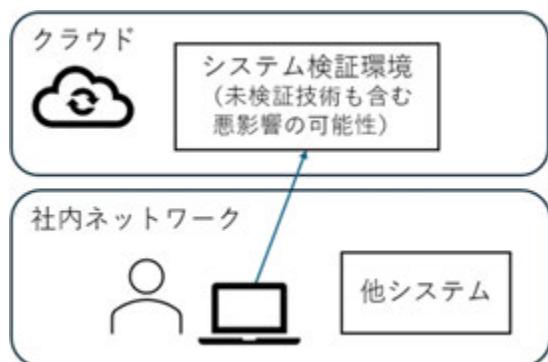


図3 システム検証環境の解決策

## 5 活用事例

この活用事例として、システム開発支援 AI エージェントシステムの検証を目的とした環境構築を紹介する(図4)。

検証対象のシステム(オープンソースソフトウェア)は、自然言語による指示でプログラムコードの作成、デバッグ、起動および動作確認を行うものであり、エディタ(VSCode)内でツールを呼び出して使用する。

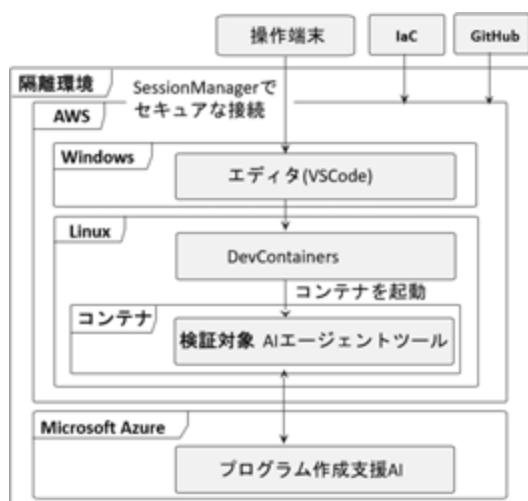


図4 活用事例の環境構成

### 対応①: 安全性の確保

#### ・ AWS、AWS Session Manager

検証対象システムに有害なバグやセキュリティリスクが含まれる可能性を考慮し、操作端末は画面情報のみをやり取りするリモート操作端末として使用し、実際の処理はすべてクラウドの隔離環境で検証を行うことができた。

### 対応②: 迅速かつ簡便な環境構築

#### ・ コンテナ、DevContainers、IaC

検証対象システムを動作させるコンテナを作成し、VSCodeから接続した。手作業の場合、手順書に基づき毎回作成から接続後の設定までを行う必要があるが、自動化により手順のミスなく容易に環境構築が可能となった。従来では1カ月以上を要していた環境構築を、今回は約2日間で完了し、93%以上の納期短縮を達成した。これにより、多くの工数とリードタイムのストレスから解放され、迅速に検証を開始できた。

### 対応③: 環境の再現性および変更管理の容易さ

#### ・ GitHub、IaC

AWS 上の Windows や Linux、ネットワーク構成の設計図は、250行のテキストで記述されている。今回の活用事例では、試行錯誤を繰り返しながら環境変更を行ったが、混乱することなく一貫性のある変更管理が可能であった。

また、検証環境を一度削除し、別のメンバが再度検証を行う際には、手作業であれば半日要するところ、数分で環境構築を完了できた。

## 6 おわりに

本稿ではクラウド技術を活用し、迅速かつ簡便に構築可能なシステム検証環境を紹介した。これにより、従来のオンプレミス環境に比べて、環境構築のリードタイムを93%以上短縮し、安全で効率的なシステム検証が可能となった。

冒頭に述べたとおり、クラウド技術は日々進化しており、ITエンジニアはこれに継続して追従する必要がある。当社 IT 本部およびヤマハモーターソリューション株式会社では、クラウド活用を推進する組織体として CCoE(Cloud Center of Excellence)を両社合同で立ち上げ、AWS を中心にクラウド技術の相談・支援および人材教育をミッションとして活動している。この CCoE 活動を通じて、本稿で紹介した技術が社内において標準技術として浸透することが、クラウドネイティブな環境の実現につながり、DX 活動に寄与すると考える。

また、DX 推進に向けた基盤機能の拡張として、現在は生成 AI を活用したプログラムコードの自動生成、テストの自動化、グローバルでのコード共有などに挑戦している。これらの活動

は、システム開発におけるリードタイムを劇的に短縮する可能  
性を有しており、研究開発を継続することで、お客様へのサー  
ビス即時提供を目指している。

### ■参考文献

[1] ヤマハ発動機株式会社:「中期経営計画 2025-2027」,  
<https://global.yamaha-motor.com/jp/profile/mtp/pdf/2025/2025medium-plan.pdf>

### ■著者



落合 寛彰  
Hiroaki Ochiai  
IT 本部  
プロセス・IT 部



吉成 雅通  
Masamichi Yoshinari  
IT 本部  
プロセス・IT 部



人見 瑛一  
Eiichi Hitomi  
IT 本部  
プロセス・IT 部



浜田 善夫  
Yoshio Hamada  
IT 本部  
プロセス・IT 部

# 社会共創による新価値創造「Town eMotion」Vol. 2 —まちなか R&D における GSM 活用と循環車 共創の実践と展望—

Creating New Value through Social Co-Creation: "Town eMotion" Vol. 2  
—Practical Applications and Future Outlook of GSM Utilization and Collaborative  
Development of Circular Vehicles in Urban R&D—

住友 剛 植原 瑞穂 杉浦 知志

## Abstract

Yamaha Motor's "Town eMotion" project, launched in 2020, promotes open innovation aimed at enhancing the well-being of people and society through field-based urban R&D. This paper, following the introductory report, presents the second phase of the initiative, focusing on the implementation and co-creation processes related to green slow mobility (GSM) and the development of a novel "Junkan-sha" model — a mobile co-creation hub that serves as a center for circulating resources and human connections, generating moments of serendipity through unexpected encounters and discoveries enabled by mobility — within local communities. From 2021 to 2025, eight PoC tests have been conducted in four regions, including Setagaya, Kamakura, Iwata, and Ueno, to test the hypothesis that mobility can foster community engagement and help solve local challenges. Building on this, the Kamakura Creative Field initiated a co-creation process in collaboration with local stakeholders and the COI-NEXT program, resulting in a new prototype designed not only as a means of transport but also as a mobile hub for dialogue, circular economy practices, and social engagement.

Through agile development from versions V1 to V4, and a transition to public-road-capable models, the circulation vehicle has evolved to support various community functions. The project also established partnerships with local DAO initiatives to co-develop materials and services, demonstrating the vehicle's potential as a scalable platform for regional co-creation and social innovation.

This study positions the "Junkan-sha" as a tool for redefining mobility—not merely as a means of transport, but as an enabler of new social interactions, resource cycles, and collective creativity. The findings suggest a scalable model for localized innovation aligning with the project's vision of "Art for Human Possibilities" while supporting long-term social value creation.

## 1 はじめに

ヤマハ発動機株式会社クリエイティブ本部では、2020年より、人と社会の "Well-being" 向上を目指し、オープンイノベーション型の研究開発プロジェクト「Town eMotion(タウンエモーション)」を推進してきた<sup>[1]</sup>。その基盤として定義された「クリエイティブフィールド(CF)」における社会価値創造の実践は、地域課題へのアプローチや社会的インパクトの創出に資する重要な試みである。本稿では、その第2段階として、CFでのまちなか R&D 活動におけるグリーンスローモビリティ(GSM)<sup>[2]</sup>を用いた検証、および "循環車" という新たな地域連携モビリティの共創開発プロセスとその可能性について報告する。

## 2 GSM を起点とした価値検証と共創プロセス

### 2-1. GSM 活用による PoC 検証(2021–2025年)

本プロジェクトでは、「モビリティが地域の多様な場づくりと移動ニーズ両方に対応することで普及促進となり社会的価値を創出する」との仮説のもと、当社 GSM 事業化部門と協力し、以下8件の PoC(概念実証)を実施した。

- ・ 東京都世田谷区(三宿エリア):3件
- ・ 神奈川県鎌倉市:3件
- ・ 静岡県磐田市:1件
- ・ 東京都台東区上野公園:1件

これらの実証では、GSM 車両を基盤に、休憩・交流・モビリティ発着拠点など、地域ニーズに基づく「場づくり」の可能性が検証されると同時に、運用面・機能面での課題を抽出改良を行った(図1)。

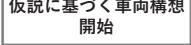
2021	2022	2023	2024
仮説「モビリティが地域の多様な場づくりと移動ニーズ両方に対応することで普及促進となり社会的価値を創出する」 2021-2024まで多様な場づくりのニーズを検証&改良			
2021年 10月 青空図書館 ：世田谷PoC  地域ニーズ 「場づくり」とし ての初トライ  	2022年 3月 モバイルパークレット ：世田谷PoC  室内外の自由度を高め る必要性を確認 デザイナーによる場 づくり仕様を起案   	2023年 5月/8月 モバイルパークレット ：世田谷PoC  ワークショッピング内容を 反映した仮説をプリテ スト車で検証開始    	2023年 9月/11月 ：上野・御厨PoC  プリテスト車によ る検証と改良    
仮説に基づく車両構想 開始			

図1 車両のニーズに合わせた改良(2021-2024)

## 2-2. 鎌倉 CF における循環車構想創出と共創活動

2023年以降、鎌倉 CF では、リサイクリエーション慶應鎌倉ラボが主導する「COI-NEXT プロジェクト(共創の場形成支援プログラム)」<sup>[3]</sup>(以下、慶應 COI-NEXT)に参画し、その活動の中で地域資源循環を軸とした「循環車」構想を創出した。本構想では、GSM ベースのモビリティを、地域住民による自主運用を視野に入れた新たな社会地域活動の HUB(集いや活動の起点)とするこどをを目指している。「※循環者になるまち ～社会でまわす、地球にかえす、未来へのこす～PJ」<sup>[1]</sup>では、建築・都市・素材リユースの観点から、環境意識を醸成しつつ地域資源を活用する共創モデルとしての循環車の需要と有効性が確認されつつある。

循環車とは、「人のご縁と資源の循環の HUB となる」を中心コンセプトに、市民が地域資源の循環を含めた地域課題や、活性化等 Wellbeing 向上につながる様々な活動に向け、車両の仕様や使い方をクリエイターとして活用できるモビリティである。それにより、「循環者になるまち ～社会でまわす、地球にかえす、未来へのこす～PJ」を含めた様々な形で地域の役に立つことを目標としている。車両の形態は、今回は GSM を起点としているが、活動が自律分散型で普及することも見据え、多様なモビリティに発展することを想定している(図2)。

(※循環者:資源循環活動を自分で実行する人)

## 2-3. 鎌倉の共創活動を通じて実現した多様な検証

循環車をテーマとした共創活動を通じた車両開発・場づくり検証を行った(図3)。

共創パートナーは以下の通り

- ・慶應 COI-NEXT
- ・鎌倉ウェルビーイングラボ<sup>[2]</sup>
- ・ゴミフェス532<sup>[3]</sup>
- ・鎌倉竹部<sup>[4]</sup>
- ・ZAMPU PROJECT<sup>[5]</sup>

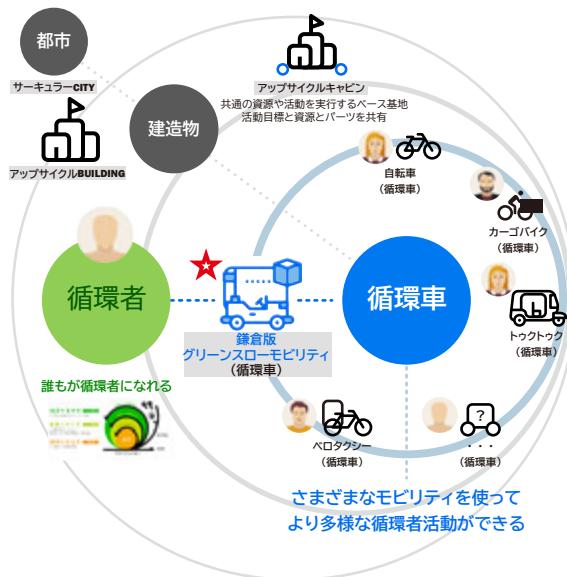


図2 循環車の概念図



図3 共創パートナーとの活動風景

左上:鎌倉ワーケーションウィーク<sup>[7]</sup>における慶應 COI-NEXT 関係者対話  
右上:鎌倉ワーケーションウィークにおける鎌倉竹部ワークショップ  
下:鎌倉竹部× ZAMPU PROJECT 共創による  
くつろげるアート空間トライアル

PoC ではアンケート調査を実施し、移動型の場づくりというコンセプトに対し多くの賛同を得ることができている。

最新の循環車仕様のプリテスト車を使った、鎌倉市役所での市民参加イベントでのアンケートから、鎌倉のまちなかでモビリティでの場づくりを使うことについては、8-9割の回答者が「非常にそう(良いと)思う」「そう思う」と回答した。

活用方法としては、「休憩スペース」という回答が、最も多い支持を集めた。参加者では特に「アートに触れる場の票も多い。活用場所については、海辺や公園など、開放的な自然の要素がある環境に票が集まった。

2025年5月24日鎌倉市役所での市民主体イベントにて実施したアンケート結果を抜粋して示す(図4)。

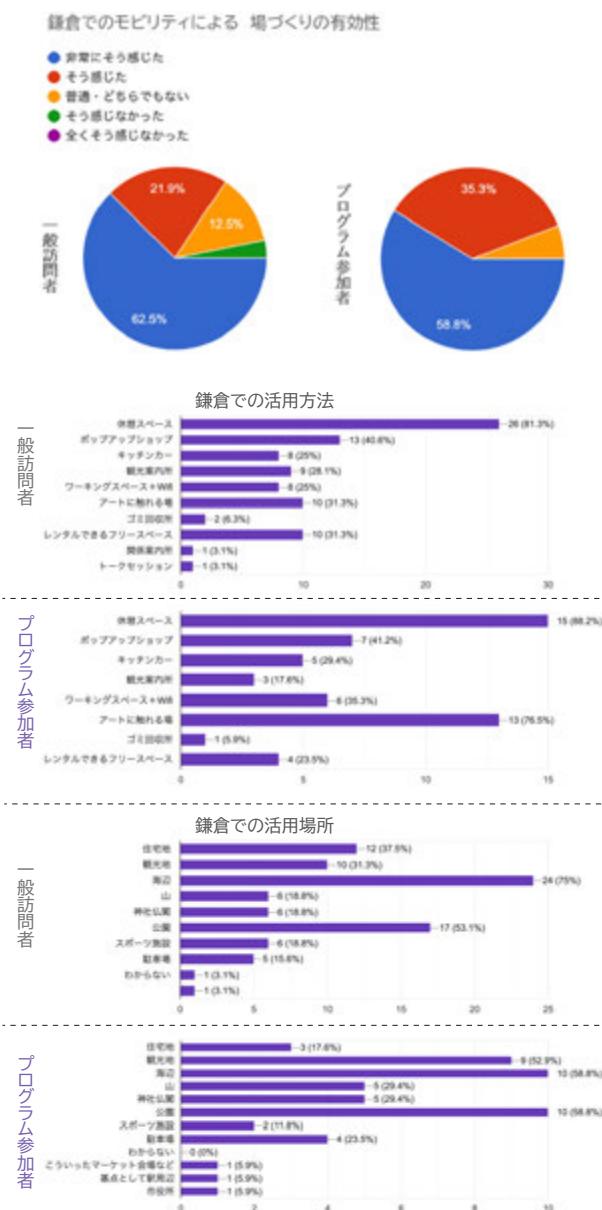


図4 2025年5月24日鎌倉市役所での市民主体イベントにて実施した主なアンケート結果(一部抜粋)

共創パートナーそれぞれの目的に応じた活動との組み合せにより、様々な仕様とトライアル、それに合わせた改良開発～検証が実現した(図5)。



図5 鎌倉ワーケーションウィークで検証された多様な目的での場の使い方と空間アレンジとその分析

### 3 循環車の開発

#### 3-1. プリテスト車のアジャイル開発による進化(V1~V4)

GSM をベースとしたプリテスト車を用い、PoC 結果から得られた知見を反映し、段階的に改良を実施してきた。

V1~V3: 地域の多様なニーズを把握し、それに合わせたアジャイルな検証開発を実施(図6・図7・図8・図9)。場づくり仕様にて走行なしで開発検証を実行。

V4: 慶應 COI-NEXT での循環型まちづくりを目指した大成建設の移動型建築<sup>[4]</sup>と共通モジュールをエクステリアパーツとして採用し、循環車検証プロトタイプとした(図9)(部品の共用交換による活動一異業界連携が狙い)。



図6 V1:中古カートを改造し、取り外し式シート・テーブル可変ルーフ等を装備



図7 V2:ルーフ・テーブル軽量化・簡素化、設置性向上に対応した改良

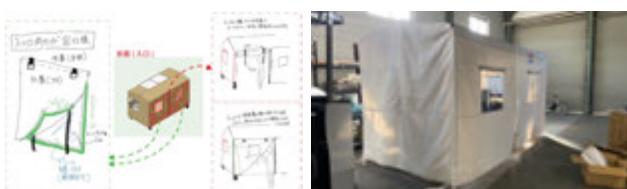


図8 V3:建築視点を取り入れ、天候や防犯、多様な空間アレンジ可能なフルカバー型テントを開発



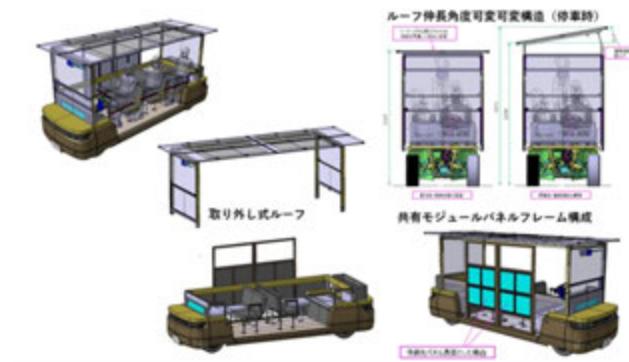
図9 V4:地域資源などの循環素材を適用できる建築との共通モジュールを搭載可能なマウント構造を採用

#### 3-2. 地域での運行を伴う検証を見据えた開発

##### 3-2-1. 公道走行型プロトタイプ開発(2023年末~)

V1~V4で得た知見を基に、当社製 GSM AR-07<sup>[2]</sup>をベースとして活用し、公道走行可能な循環車プロトタイプ一次車の開発に着手。

プリテスト車で有効性が確認された基本構造(取り外し式ルーフ・ピラー、シート、ステアリング構造)を採用。慶應 COI-NEXT の共通モジュールパネル対応構造と、地域ニーズに応じた仕様変更が可能となっている(図10)。



走行時運転操作系→停車時 ステアリング取り外し & インパネがシートに可変

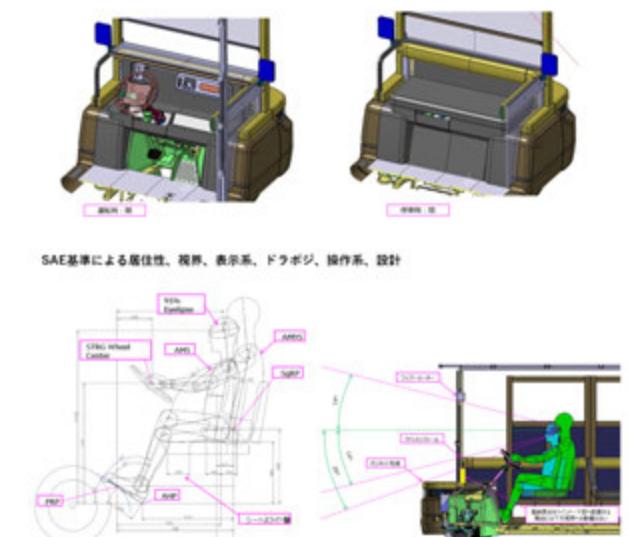


図10 公道走行型プロトタイプの機上検討例

##### 3-2-2. 共創目的に沿った車両運行検証

AR-07を使い、ライセンス講習と市内のトライアル運行についても検証を実施し、“場づくり”と“旅客、輸送”の組み合わせの有効性を共創パートナーが実際に体感する機会となった(図11)。



図11 左:ライセンス講習会の様子  
右:共創パートナーによるGSM運行  
トライアル竹林ツアー検証



## 4

## 循環車共創のモビリティが担える社会的可能

### 4-1. 地域 DAO との共創手法の有効性

循環車プロジェクトでは、地域 DAO<sup>6)</sup>との連携によって、資源循環と社会的包摶の実装を進めている。

各共創パートナーに対しては対話を通じて相互理解を深め、デザイン思考やサービスデザインの手法でパートナーの持つ課題やビジョンを明確にした。こうした活動を通じてパートナーが共創メリットを感じており、PJに対する期待とエンゲージメントも高まっている(図12、13)。

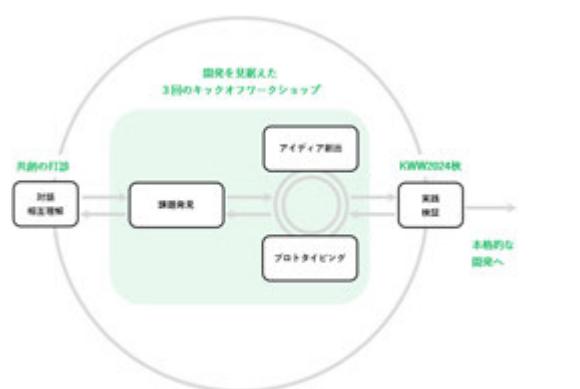


図12 循環車共創パートナーに対する共創アプローチの事例



図13 循環車共創活動鎌倉竹部との共創 WS の様子

### 4-2. 循環車によるモビリティの新たな社会的価値役割の発見

循環車は移動手段の枠を超えて、地域資源の集積・再配分のHUB、多様な活動の拠点、対話を誘発する“場”として機能し得る可能性が示されつつある。

特に、鎌倉ワーケーションパーク<sup>7)</sup>との連携により、偶発的な出会いや創造性の誘発=セレンディピティ<sup>8)</sup>を生み出す手段として有効性を示している(図14)。また本活動は、慶應 COI-NEXT の掲げる「産官学民共創」において、特に「民-民」「民-産」といった接続領域の関係構築・強化における HUB として機能し始めている(図15)。

- 1 地域パートナーとの循環車共創の実践
- 2 モビリティによる縦づくりの可能性探索
- 3 COI-NEXT関係者への活動発表 + 循環車へのフィードバック取扱

資源循環活動を行う地域パートナーとの共創の内容を実践する場として、KWWのプログラム提供の場を活用します。

KWW2024春、七里ヶ浜駅周辺において、循環車での縦づくり実験。沿岸先の検索を行ったことに続き、今回は駅周辺を鶴岡ヶ崎公園に変え、「縦寄せ」というテーマを追加し、同様の検証を行います。

COI-NEXT関係者に循環車のコンセプトを説明し、実際の車両・場づくりのプロトタイプを体験してもらうことで、COI-NEXT実践の場での循環車の活用可能性を探ります。

※KWW：鎌倉ワーケーションパーク

図14 循環車検証を含む複数活動の統合の「場」としての  
鎌倉ワーケーションパーク

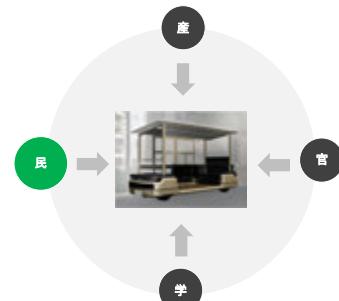


図15 慶應 COI-NEXT コンソーシアム参加団体の接続領域  
関係構築・強化 HUB のイメージ

## 5

## 考察:まちなか R&D による価値創造の新展開

循環車の開発は社会的価値の創出(アウトカム)として、当社の掲げる2030長期ビジョン「Art for Human Possibilities」<sup>[5]</sup>および、中期経営計画2025<sup>[6]</sup>でも定義されている「心躍る移動体験」「社会課題の解決」「生活の質の向上」「脱炭素・循環社会への貢献」を体現している活動である。以下に考察を示す。

### 5-1. “場の創出”による課題解決、新価値創造

従来の「移動手段としてのモビリティ」に加え、「地域資源の循環」「対話の創出」「関係性の再構築」など、社会的課題や新

価値創造活動に対して「場の創出」、セレンディピティ創出、マルチステークホルダーの共創という、包括的アプローチに本プロジェクトの独自性がある。

### 5-2. モビリティを“共創装置”として再定義

循環車は、「人とモノを運ぶ手段」にとどまらず、「多様な価値が交差・蓄積・再分配される共創のためのプラットフォーム」として再定義できる。

### 5-3. スケーラブルな地域共創モデルとしての展開

鎌倉で得られた知見（循環車構想とそれを通じて実現できる価値－共創開発の実行手法）を他地域へ展開することで、特定の車両形式に依存せず、地域課題と接続したスケーラブルな共創モデルとして、循環車は持続可能性と社会実装性の高いプラットフォームとなり得る。

循環車プロジェクト一環として活動内容を型化してワーキングショップパッケージを開発しており、それを使い次のステップで国内外への普及活動を実施予定（次回報告予定）。

## 6 おわりに

本稿では、Town eMotion プロジェクトにおける GSM の検証を出発点とし、鎌倉 CF における循環車共創のプロセスと意義を報告した。まちなか R&D を通じて地域住民や活動団体との協働による“移動する場”的開発は活動を通じたセレンディピティにより、多くの共創パートナーと共に随時新たな発見と、人と社会の Well-being の向上につながる想定以上の新価値創造の可能性を示しつつある。

今後は、活動家との場つくりと車両運行の同時検証にて、検証と開発の深化を狙う。さらに共創プロセスの汎用ガイド作成、非財務的価値の可視化を通じて、多地域展開に向けた実装モデルの構築を推進していく。

本活動は、循環車というモビリティの新しい在り方とその開発活動を通じて生まれる社会的意味、共創で作り上げることにより生まれる複数の価値が実現できる可能性が活動の中から見出されたことが非常に意義深い ART for Human Possibilities の実践の好例となるようにこれからも信念をもって続けていきたい。

### ■参考文献

[1] ヤマハ発動機株式会社. “社会共創による新価値創造「Town eMotion」Vol. 1 まちなかR&D クリエイティブフィールドの可能性”. ヤマハ発動機株式会社HP. <https://global.yamaha-motor.com/jp/profile/mtp/> (参照2025-06-25)

[yamaha-motor.com/jp/design\\_technology/technical/presentation/pdf/browse/59gs04.pdf](https://yamaha-motor.com/jp/design_technology/technical/presentation/pdf/browse/59gs04.pdf) (参照2025-06-25)

[2] ヤマハ発動機株式会社. “グリーンスローモビリティ（電動カート公道仕様）”. ヤマハ発動機株式会社HP. <https://www.yamaha-motor.co.jp/gsm/> (参照2025-06-25)

[3] 慶應義塾大学. SFC研究所 “デジタル駆動超資源循環参加型社会共創コンソーシアム”. 慶應義塾大学HP. <https://coinext.sfc.keio.ac.jp/> (参照2025-06-25)

[4] 大成建設株式会社. “アップサイクル建築プロトタイプ「アップサイクルキャビン」”. 大成建設株式会社HP. [https://www.taisei-techsolu.jp/news/2025/250303\\_10303.html](https://www.taisei-techsolu.jp/news/2025/250303_10303.html) (参照2025-06-25)

[5] ヤマハ発動機株式会社. “長期ビジョン「ART for Human Possibilities」”. ヤマハ発動機株式会社HP. <https://global.yamaha-motor.com/jp/profile/long-term-vision/> (参照2025-06-25)

[6] ヤマハ発動機株式会社. “中期経営計画2025年～2027年”. ヤマハ発動機株式会社HP. <https://global.yamaha-motor.com/jp/profile/mtp/> (参照2025-06-25)

### 用語解説

1) 慶應義塾大学と鎌倉市が推進する、COI-NEXT「循環者になるまち～社会でまわす、地球にかえす、未来へのこす～」プロジェクト：文部科学省／科学技術振興機構（JST）、「共創の場形成支援プログラム（COI-NEXT）」2023-2032 リスペクトでつながる「共生アップサイクル社会」共創拠点。慶應義塾大学と鎌倉市による資源循環型社会の担い手育成プロジェクト。プラスチック問題を背景に、使い捨てから循環への意識と社会構造の転換を目指す取り組み。慶應義塾大学HP. <https://junkansha.jp/> (参照2025-06-25)

2) “鎌倉ウェルビーイングラボ：人のウェルビーイングを中心に、生活・空間・環境デザインの観点から社会課題の探求と実践を行う研究・実践組織。鎌倉ワーケーションウィークを主催”。鎌倉ウェルビーイングラボHP. <https://zampu-project.studio.site/> (参照2025-06-25)

3) “ゴミフェス532（ゴミフェスゴミニティ）：日本一ゴミ問題が深刻な鎌倉での地域活動／「コミュニティの力でごみ問題を楽しく解決」をテーマに、市民・団体・企業・行政が連携した活動”. ゴミフェス532HP. <https://gomifes532.com/> (参照2025-06-25)

4) “鎌倉竹部：鎌倉市を拠点に、放置竹林問題への対処と竹資源の有効活用を進める地域団体”. 鎌倉竹部HP. <https://www.kamakuratakebu.com/> (参照2025-06-25)

- 5) "ZAMPU PROJECT(ザンプ・プロジェクト):衣料廃棄物の再資源化と、障害者福祉施設との連携による創造的なアップサイクル活動を行う共創型プロジェクト". ZAMPU PROJECT HP. <https://zampu-project.studio.site/> (参照2025-06-25)
- 6) "DAO(Decentralized Autonomous Organization):自律分散型組織"
- 7) "鎌倉ワーケーションウィーク:生き方働きかた先進地域鎌倉において、人と社会のWellbeing向上を目指した地域活動". 鎌倉ワーケーションウィークHP. <https://www.kamakura-workation.com/> (参照2025-06-25)
- 8) "セレンディピティ(Serendipity):偶然の発見や出会いによって新たな価値や創造性が生まれる現象"

### ■著者



住友 剛  
Takeshi Sumitomo  
クリエイティブ本部  
プランニングデザイン部



榎原 瑞穂  
Mizuho Sakakibara  
クリエイティブ本部  
プランニングデザイン部



杉浦 知志  
Satoshi Sugiura  
クリエイティブ本部  
プランニングデザイン部

# 高出力応答が求められる小型モビリティへの適用を 目指した FCV の電力マネジメントシステムの開発

Development of a Power Management System for Fuel Cell Vehicles (FCVs)  
Aimed at Application in Small Mobility Vehicles Requiring High Power Response

鈴木 聖也 中田 暢大

## Abstract

In recent years, the introduction of small electric mobility vehicles equipped with fuel cells has been proposed as part of efforts to realize a carbon-neutral society<sup>[1][2]</sup>. Compared with BEVs (Battery Electric Vehicles), fuel cell vehicles offer advantages such as longer driving range. On the other hand, fuel cells degrade under load fluctuations, leading to reduced output performance.

Therefore, Yamaha Motor Engineering Co., Ltd. (hereafter “the company”) has developed a hybrid system combining a fuel cell and a LiC (Lithium-ion Capacitor) for application in small electric mobility vehicles requiring high power response. This system suppresses load fluctuations on the fuel cell while achieving approximately 53% reduction in size, 50% reduction in weight, and 23% improvement in acceleration performance compared to systems using LiB (Lithium-ion Battery) as the secondary battery.

Although LiC is compact, lightweight, and provides high output, its discharge capacity is small. Accordingly, we first built a prototype hybrid system combining a fuel cell and a LiB to verify the load-fluctuation suppression function. Next, we developed a system in which the LiB was replaced by LiC, obtaining favorable results in both the suppression of fuel-cell load fluctuations and the improvement of driving performance.

This paper introduces these technical development efforts and presents the evaluation results.

## 1 はじめに

近年、カーボンニュートラル社会の実現に向けて、燃料電池を搭載した小型電動モビリティの導入が提案されている<sup>[1][2]</sup>。燃料電池車は、BEV(Battery Electric Vehicle)と比較して、航続距離などの面で優位性を有する。一方、燃料電池は負荷変動によって劣化し、出力が低下するという課題がある。そこで、ヤマハモーターエンジニアリング株式会社(以下当社)では、高出力応答が求められる小型電動モビリティへの適用を目的として、燃料電池と LiC (Lithium-ion Capacitor)によるハイブリッドシステムを開発した。本システムにより、燃料電池の負荷変動を抑制しつつ、二次電池に LiB (Lithium-ion Battery) を採用した場合と比較して、LiC にて約53%の小型化、約50%の軽量化、約23%の加速性能向上を実現した。LiC は小型軽量かつ高出力であるが、放電容量が小さい。そのため、まずは燃料電池と LiB によるハイブリッドシステムを試作し、負荷変動抑制機能の検証を行った。次に、LiB を LiC に置き換えたシステムを開発し、燃料電池の負荷変動抑制効果および走行性能の向上について良好な結果を得た。本稿では、これらの技術開発の取り組みと評価結果について紹介する。

## 2

## LiB を用いた燃料電池の負荷変動抑制システム

燃料電池をモビリティに搭載するには、加減速時の負荷変動や高負荷走行時の出力要求に対応する必要がある。このため、自動車をはじめとする多くのモビリティでは、燃料電池と二次電池を組み合わせたハイブリッドシステムが採用されている<sup>[3]</sup>。

二次電池を併用することで、燃料電池が停止している間の電力供給が可能となるほか、負荷変動を吸収することで燃料電池の劣化を抑制できる。また、燃料電池単独では利用できない回生エネルギーを二次電池に回収することが可能であり、システム全体のエネルギー効率向上にも寄与する。これらを踏まえ、本開発においても、燃料電池と二次電池を組み合わせたハイブリッドシステムを試作し、燃料電池の負荷変動抑制効果を検証した。ここで、ダイナミックな出力特性を持つ小型電動モビリティとして、レーシングカートを選定した。

### 2-1. システム構成

試作したシステムの構成を図1に示す。燃料電池システムの諸元は表1に示す。燃料電池、二次電池、モータは市販品を使用し、電流制御回路は試作した。燃料電池およびモータは、エンジン搭載型レーシングカートを参考に、走行時の楽しさを確

保できる出力特性と、市販の空冷式燃料電池の仕様を踏まえて選定した。二次電池は、入出力電流特性を考慮して選定した。

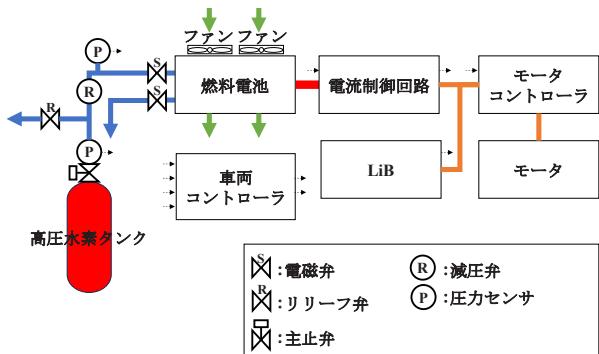


図1 試作した燃料電池システム

表1 システム諸元

燃料電池 - 型番	H-5000 @HORIZON FUEL CELL
-タイプ	PEFC
-定格出力	5000W (72V at 70A)
LiB -セル型番	20Ah-HP @TOSHIBA
-公称電圧	48V (22直列)
-最大放電電流※	120A
-最大充電電流※	120A
電流制御回路	試作
-入力電圧	70-120V
-出力電圧	35-60V
モータ -型番	MA01B @MITSUBA
-最大トルク	20Nm @210Arms
-出力	6kW @48V

※独自に設定

本システムでは、スロットル操作などのユーザ入力を受けた車両コントローラが、燃料電池、LiB、モータなどの状態を考慮し、モータコントローラにトルク指令値を、電流制御回路に燃料電池の発電電流指令値を出力する。電流制御回路は、入力された発電電流指令値を目標電流値とし、実測電流値との差分に基づいてPWM(Pulse Width Modulation)のデューティ比を決定する。

## 2-2. 電力マネジメント

本開発では、燃料電池の負荷変動抑制を目的とした電力マネジメント制御を検討した。電力マネジメントの全体構成を図2に示す。

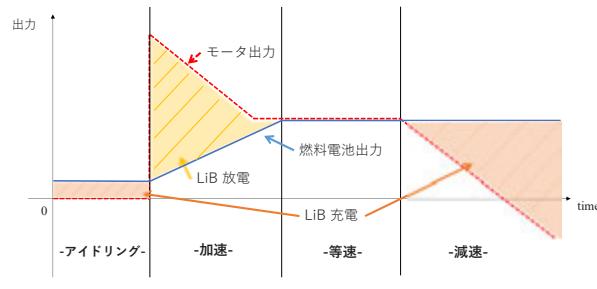


図2 電力マネジメント制御構成

アイドリング時には、燃料電池で発電した電力をLiBへ充電する。等速走行時や減速時などの低負荷走行時には、燃料電池の発電電力でモータを駆動するとともに、余剰電力をLiBに充電する。一方、高負荷走行時や急激な負荷変動が発生した場合には、燃料電池の発電電力に加え、不足分の電力をLiBから供給することで、燃料電池の負荷変動を抑制しつつ、モータの負荷要求に対応する。このような制御により、燃料電池の発電量を任意の値で維持しながら、モータの電力需要に柔軟に対応することが可能となる。

電力マネジメント制御の成立性検証および制御パラメータの適合を目的として、車両電力収支シミュレータを構築した。本シミュレータは、車両運動モデルと、モータ、LiB、燃料電池からなる電気モデルを組み合わせたものである。シミュレーション環境にはMATLAB R2022bを使用した。各モデルの粒度はシステムレベルとし、ツールボックスに含まれるライブラリブロックを活用した。モデルパラメータおよび走行パターンは、実測データに基づいて設定した。シミュレーション結果の一例を図3に示す。

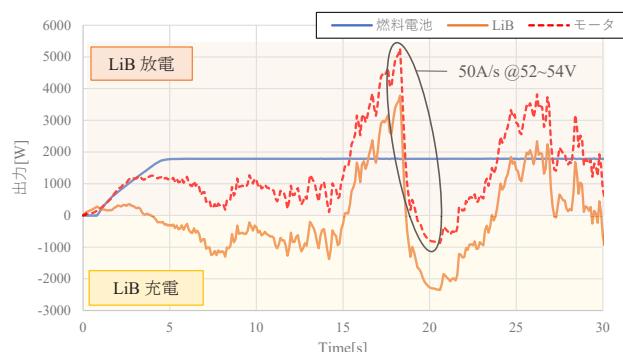


図3 各コンポーネントの電流比較

図3より、モータ出力が燃料電池の発電電力を上回る場合には、不足分をLiBから放電して補っており、逆にモータ出力が燃料電池発電電力を下回る場合には、LiBへの充電が確認できる。また、モータ出力の変動に対して燃料電池発電電力は一

定であり、負荷変動が抑制されている。さらに、本シミュレータを用いて、発電電流指令値などの制御パラメータの適合を行った。具体的には、走行パターンにおいてLiBのSOC(State Of Charge)に対する発電電流指令値を変更し、水素消費量、LiBの消費電力量、航続距離を網羅的にシミュレーションした。その結果、電欠回避のために、燃料電池の発電電力量がLiBの消費電力量を上回り、かつ航続距離が最大となる発電電流指令値を採用した。その他の制御パラメータも同様の手法で適合した。このように、シミュレーションを活用することで、電力マネジメント制御の成立性検証および制御パラメータの適合期間を、実機主体で検証／適合した場合と比較して約60%程短縮できた。

### 2-3. 実走評価

実走評価のため燃料電池システムを試作し、レーシングカートのフレームに搭載した。試作車両の外観を図4に示す。

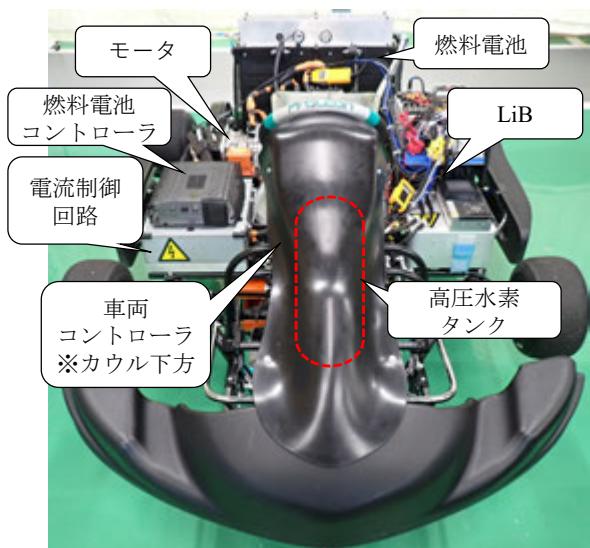


図4 試作車両

モータは、一般的なエンジンカートにおけるエンジン搭載位置と同様の場所に配置した。燃料電池は車両後方に、LiB、電流制御回路、車両コントローラ、水素タンクおよび補機類は、車両の左右バランスと積載スペースを考慮して配置した。

本試作車両を用いて、実走行時における負荷変動抑制効果を検証した。計測結果の一部を図5に示す。

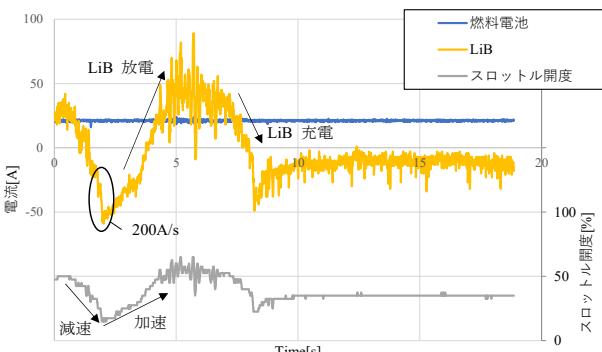


図5 実走行時の電流、スロットル開度

図5は、特に大きな電流変動が発生した場面のデータである。LiBの電流は最大90A、最小-60Aとなり、電流変化率は最大200A/sに達している。一方、燃料電池の発電電流は走行中を通じて一定値で維持されており、LiB電流やスロットル入力の急変動の影響を受けていない。以上の結果より、本システムが燃料電池の負荷変動抑制に有効であることが確認された。

### 3 二次電池の換装によるシステムの小型化および動力性能の向上

本章では、二次電池をLiBからLiCへ換装することによる、システムの小型化・軽量化および動力性能の向上に関する取り組みを紹介する。一般に、走行に必要な電力を燃料電池で貯うことが可能なシステムでは、二次電池は主に負荷変動の吸収に用いられる。この場合、二次電池を負荷変動の吸収に加え走行に必要な電力を貯う目的でも使用した際と比較し、必要な電力量が小さくなるため、放電容量を抑えることが可能である。ここで、燃料電池の発電量を制御しない場合、発電量は二次電池の電圧に依存し、電圧が高い場合には発電量が小さく電圧が低くなると発電量が多くなる。モータ負荷が大きい場合、燃料電池の発電が間に合わず二次電池の電力が消費されるため、コース走行中に電欠しないだけの十分な放電容量を確保する必要がある。そこで、二次電池の電圧に依存せず、任意の発電電流を設定可能な電流制御機能を備えた燃料電池システムでは、モータ負荷に応じて発電電流を最適化できる。すなわち、モータ負荷が大きい場合には発電電流を増加させ、小さい場合には低減させる制御が可能となる。試作したシステムは、燃料電池により走行に必要な電力を直接供給でき、かつ電流制御により発電電流を柔軟に設定可能である。このため、二次電池の放電容量をさらに小型化できると考え、LiBに比べて放電容量は小さいものの、小型・軽量かつ高出力であるLiCを採用した。採用したLiBおよびLiCの仕様比較を表2に示す。

表2 LiBとLiCの仕様比較

	LiB	LiC
セル型番	20Ah-HP	CPQ3300SD
公称電圧	48V(22 Series)	48V(16 Series)
放電容量	20Ah	1.47Ah
最大放電電流*	120A	500A
最大充電電流*	120A	500A
寸法(L×W×H)	460×173×204mm	335×178×127mm
重量	18kg	9kg

\*独自に設定

LiCへの換装により、二次電池の重量を約50%、体積を約53%削減でき、システムの小型化・軽量化が可能となる。また、LiCは充放電能力が約4倍高く、動力性能の向上も期待される。なお、今回採用したLiCは、LiBと同一の電圧範囲を有しており、二次電池以外のシステム構成部品は変更せずに換装した。一方で、LiCの放電容量はLiBの約8%に相当するため、過放電を防止する電力マネジメント制御が重要となる。

### 3-1. LiC換装後の電力マネジメント

レーシングカートとしての走行性能を確保するためには、走行中に過放電による駆動トルクおよび過充電による回生トルクの制限が生じないことが望ましい。これらの制限は、意図した加減速を阻害し、走行性能の低下を招くおそれがある。そこで、本開発では、走行パターンにおいてLiCが過放電および過充電に至ることなく走行可能な電力マネジメント制御を、シミュレータを用いて検討した。

シミュレータは、LiBのパラメータをLiC仕様に置き換えて構築した。ここで、LiCはSOC低下に伴い電圧が低下する特性を有し、SOC60%時に約48Vとなる。この特性を踏まえ、LiCのSOCを60から70%の範囲で維持することを目標に、網羅的なシミュレーションを実施し、燃料電池の発電電流を決定した。また、SOCが設定範囲を逸脱する場合には、駆動トルクおよび回生トルクに制限を加える制御ロジックを適用し、同様に網羅的なシミュレーションを行って、制限値を決定した。シミュレーション結果の一例を図6に示す。

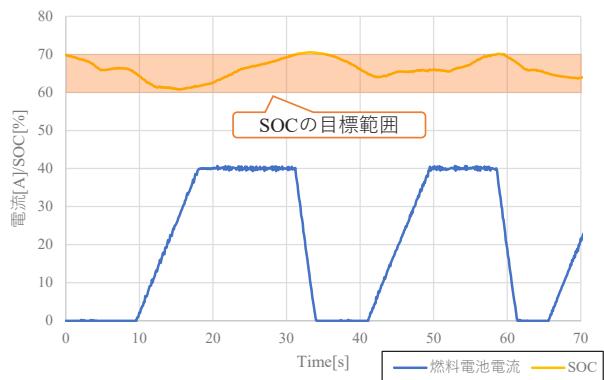


図6 コース1周走行時の発電電流に対するSOC

図6より、コース1周走行した時、SOCは目標範囲内で維持されており、電力マネジメント制御の成立性が確認された。

### 3-2. LiC換装後の実走評価

実走評価により、システム成立性の検証および、LiB搭載時との走行性能比較を実施した。評価は、広く平坦なアスファルト敷地に設定した特設コースにて行った。電力収支の結果を図7に示す。

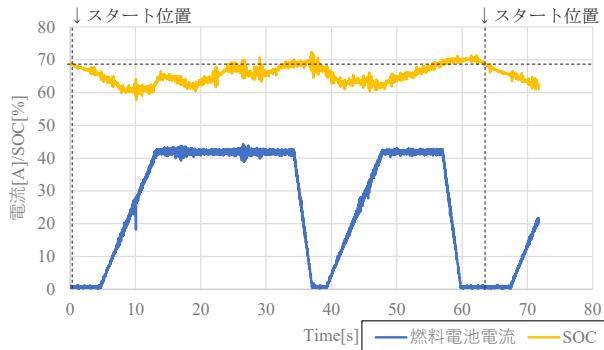


図7 コース走行時の電力収支

図7は、特設コースにて高負荷走行をした時の任意の周回を抽出したものである。各周回におけるスタートライン通過時のLiCのSOCを比較した結果、いずれも同一であり、走行中を通じて一定のSOCを維持できていることが確認された。この結果より、燃料電池の負荷変動抑制機能を維持したまま、二次電池の小型化・軽量化が成立することが示された。加速性能比較結果を図8に示す。

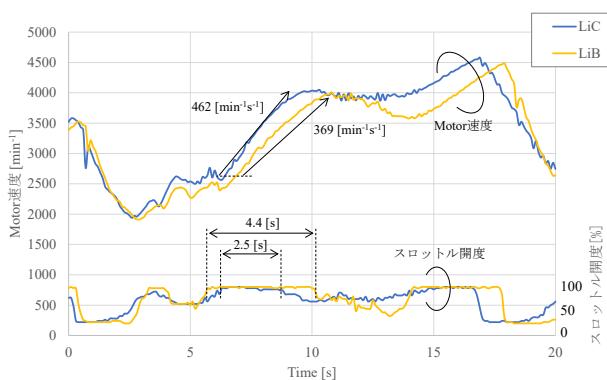


図8 LiB および LiC 搭載時の加速性能比較

図8より、LiC 搭載時には加速性能が約25 %向上することが確認された。さらに、LiC の充電能力の高さから回生トルク上限値の引き上げも実現でき、加減速の応答性に優れた車両挙動が得られた。

## ■著者



鈴木 聖也  
Masaya Suzuki  
ヤマハモーターエンジニアリング㈱  
先行技術開発部



中田 輝大  
Nobuhiro Nakata  
ヤマハモーターエンジニアリング㈱  
先行技術開発部

## 4 おわりに

本開発では、小型電動モビリティへの燃料電池システム適用における課題である負荷変動の抑制、システムの小型化・軽量化、ならびに走行性能の向上を目的として、レーシングカートを用いたシステム開発および検証を実施した。本システムにより、燃料電池の負荷変動を抑制しつつ、二次電池に LiB を採用した場合と比較して、LiC にて約53%の小型化、約50%の軽量化、約23%の加速性能向上を実現することができた。また、開発に際してシミュレータを活用することで、開発期間の短縮を実現した。

今後は、開発したシステムと電力収支シミュレータにより、ヤマハ発動機グループ内外への技術的貢献を目指す。

## ■参考文献

- [1] Yasuyuki Muramatsu, Masatsugu Ohishi, Shuhei Adachi: "Fuel Cell System for Two-Wheeled Vehicles", Proceedings The 11th. Small Engine Technology Conference No. 2005-32-0077, (2005).
- [2] Kengo Ikeya, Kazuyuki Hirota, Yohei Takada, Toru Eguchi, Kazuhiko Mizutani, Tohru Ohta: "Development and Evaluation of Air-Cooled Fuel Cell Scooter", Proceedings The 17th. Small Engine Technology Conference No. 2011-32-0644, (2011).
- [3] 陣内 亮典 (編集): "燃料電池の原理と応用 (自動車工学シリーズ 5)" , (2023).

# 板鍛造工法による新「NMAX」「YECVT」部品の 一体成形開発

Integrated forming development of new "NMAX" and "YECVT" components using the plate forging method

田丸 翔吾 北郷 博成

## Abstract

In the pursuit of carbon neutrality, it is essential to reduce not only CO<sub>2</sub> emissions generated during product use but also those produced in the manufacturing processes of component parts. In the field of plastic forming, continuous technological advancements are required to achieve lightweight, high-strength, high-precision, and low-cost manufacturing, while maintaining a balance with CO<sub>2</sub> reduction.

The plate-forging method enables a reduction in material usage, near-net shaping, and process simplification through the integration of multiple components into a single part, making it one of the most suitable manufacturing methods for achieving carbon neutrality.

Yamaha Motor Engineering Co., Ltd. provides plastic-forming technologies based on fuel-tank press-forming expertise and engages in the development of various plastic-forming methods, including forging, as well as the design and manufacture of the dies and equipment that make these processes possible.

In this project, the core component of the "YECVT unit" — the continuously variable transmission mechanism installed in Yamaha Motor Co., Ltd.'s small premium scooter "NMAX" — was redesigned. Instead of the conventional method of welding together individual parts, the new approach forms the entire component integrally from a single thick plate using a cold-forming process. This enables the creation of a lightweight, compact, high-strength, and high-precision part. By fully leveraging material properties and combining advanced analysis and die technologies, a new hybrid plastic-forming method called the "YEC Plate Forging Process" was developed. This process creates new added value through the integration of multiple forming techniques and has been successfully implemented in mass production. This paper outlines the development and key technical features of this forming process.

## 1 はじめに

カーボンニュートラルの実現に向けて、商品使用時のCO<sub>2</sub>排出のみならず、構成部品の製造過程におけるCO<sub>2</sub>排出の削減も急務である。塑性加工分野では、軽量・高強度化、高精度化、低コスト化の技術進化が継続的に求められており、CO<sub>2</sub>削減と両立するバランスの取れた工法が必要とされている。板鍛造工法は、投入材料の削減、ネットシェイプ化、複数部品の一体化による工程削減が可能であり、カーボンニュートラルに適した工法の一つである。

ヤマハモーターエンジニアリング株式会社では、燃料タンクのプレス成形技術をベースとした塑性加工技術を提供しており、鍛造を含む各種塑性加工工法およびそれを具現化する金型装置の開発を行っている。

今回、ヤマハ発動機株式会社の小型プレミアムスクーター「NMAX」に搭載された変速機構「YECVTユニット」の基幹部品において、従来の単品部品同士を溶接する仕様ではなく、一枚の厚板から冷間工法により一体成形することで、軽量・コンパク

トかつ高強度・高精度な部品として成立させた。材料の特性を活かした複合的な塑性加工工法を、解析技術と金型技術を駆使して取り入れることで、新たな付加価値を創出する複合塑性加工工法「YEC 板鍛造工法」を開発し、量産化を実現した。本稿では、当該工法開発技術の概要について述べる。

## 2 YEC 板鍛造工法を適用した YECVT 駆動部品

本稿のテーマである新型「NMAX」に搭載された YECVT の駆動系スクリュ部品(図1)について概要を述べる。当該部品は、小型スクーターパワーユニット内に収めるため、軽量・高強度かつコンパクトな形状が求められるとともに、駆動系部品として高精度な機能を有する必要がある。特に、モーターで駆動する外周のギア部、シーブを駆動するボス外周のスクリュ部、相手部品を固定するピン部において、強度および精度を満たしつつ、安定した一体成形が求められる。基本形状は、スターターギア部品に類似しているが、駆動スクリュ部を担うため、ボス

部がより長く厚くなっているため、成形難易度が高くなっている(図2)<sup>[1]</sup>。



図1 「NMAX」YECVT 板鍛造スクリュウ部品

本部品の特徴として、フランジ部とボス部の板厚比は約1.7であり、ボス部の長さもスターターギアと比較して約1.3倍程度と長くなっていることが挙げられる。このため、ボス部を座屈させずに材料を塑性流動させることができることが成形上の重要な課題である。また、フランジ部に設けられた小径ボスはハーフピアス成形により一体成形されており、従来の別体ピン圧入方式からピニット化構造へと変更することで、コスト低減を図っている。次章では、YEC 板鍛造工法による初の量産部品として「NMAX」に搭載されたYECVT 基幹部品である板鍛造スクリュウ部品の開発ポイントについて述べる<sup>[2]</sup>。

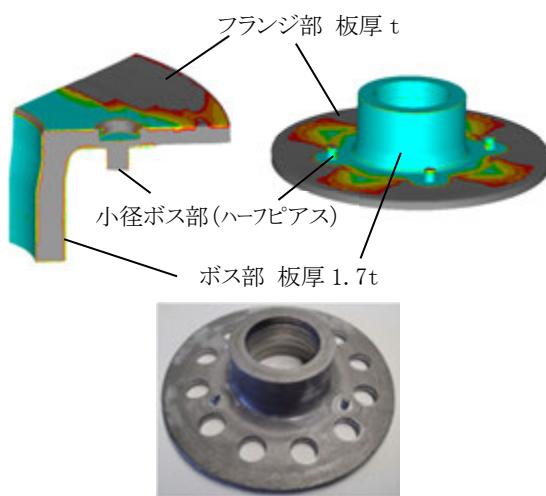


図2 YECVT スクリュウ板鍛造

### 3 工法開発のポイント

#### 3-1. 絞り 引張り成形による塑性流動

板素材からの成形は、通常の鍛造成形とは異なり、材料が圧延により薄く延伸された状態にあるため、必要な部位へ少しづつ塑性流動させる必要がある。そのため、本工法では引張りを主体とした絞り成形を複数回繰り返す再絞り成形を採用している。板素材は圧延方向により材料の流動特性が異なる“異方性”を有しており(図3)、絞り成形時にはこの影響を最小化する技術が求められる。

さらに、板幅方向においても材料特性に差があり、コイル端部と中央部では引張特性が大きく異なる(図4)。

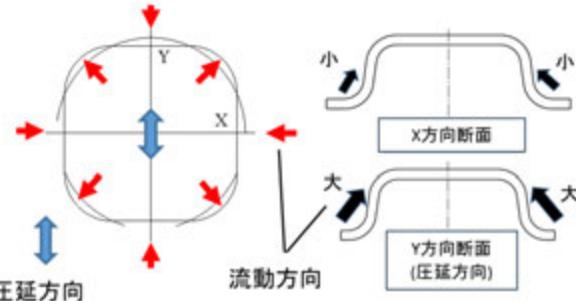


図3 絞り成形ポイント

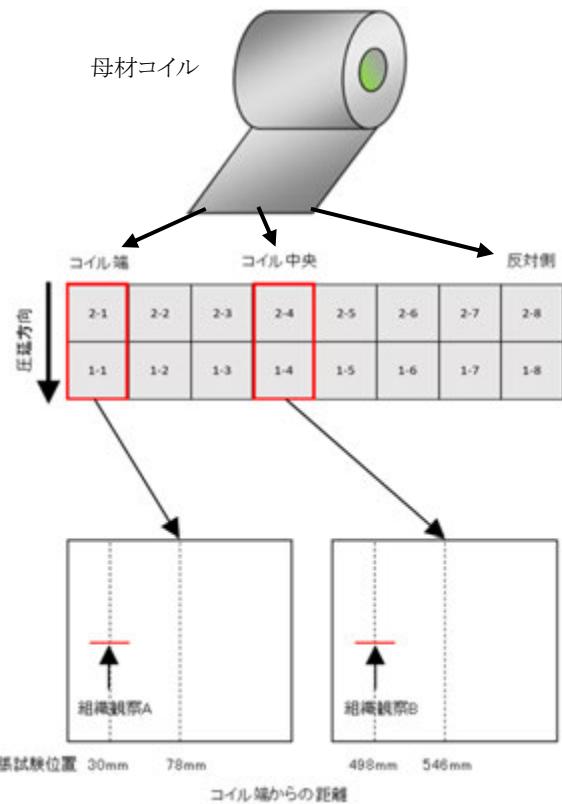


図4 材料レイアウト

実際に端部および中央部にて引張試験と組織調査を実施した結果、コイル端に近いほど伸びが小さく、材料組織が粗大化していることが確認された(図5、6、7)。また、面内異方性はJIS Z 2254に規定される式 $\Delta r = (r_0 + r_{90})/2 - r_{45}$ により評価され、 $\Delta r$ が0に近いほど真円に近いワーク形状が得られる。特にコイル端から30mm位置では $\Delta r$ が極端に小さく、絞りフランジ外径ヒケ(図8)の要因となっている。このような材料特性を踏まえ、材料歩留まりを最適化した工程設計を行うことで、YECVTスクリュウ板鍛造部品のコスト低減および量産化を実現した。

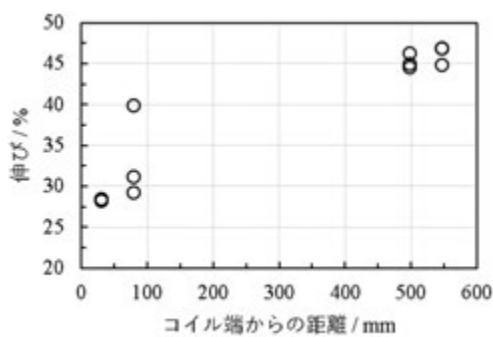


図5 引張試験 伸び比較 (SPHC)

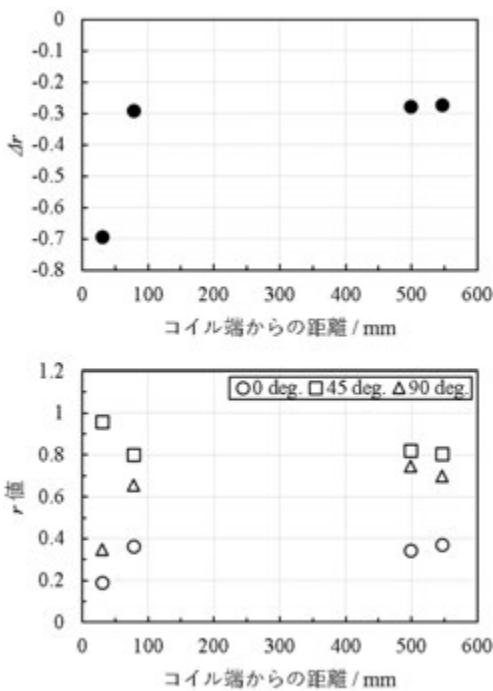


図6 引張試験  $\Delta r$  比較 (SPHC)

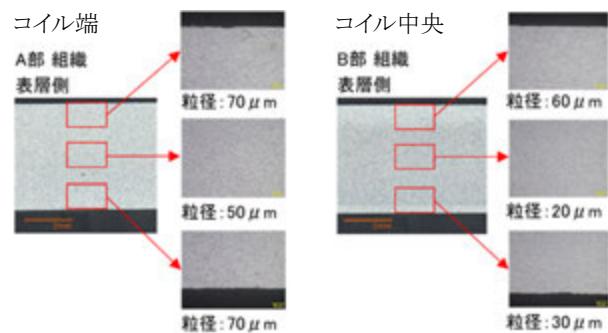


図7 材料組織比較 (SPHC)



図8 絞りフランジ外径ヒケ

本開発を通して得られた材料異方性パラメーター(ランクフォード値 =  $r$  値)を解析へフィードバックすることで、1部品内で異なる材料特性を再現したマルチマテリアル解析技術を確立することができた(図9)。このように、工法開発領域における解析技術の活用範囲は拡大しており、製造アプローチに加えて材料アプローチを含めた工法開発が求められている。

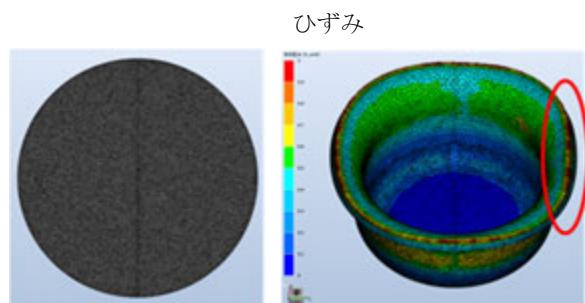


図9 マルチマテリアル解析

### 3-2. 鍛造 圧縮成形による増肉成形

スクリュウ部を構成するボス部に材料の増肉を施す方法として、鍛造による圧縮成形加工を適用し、絞り成形では得られない大きな塑性流動を実現している。YECVTスクリュウ板鍛造工程図を図10に示す。成形の要点は、ボス部を座屈させることなく材料を圧縮成形することである。技術的には、鍛造前の素材

形状の最適化と材料の拘束方法の検討が必要である。最適化に際しては、図11に示す予備成形形状から重要な機能パラメータを抽出し、鍛造成形解析を実施しながら導出している。



図10 YECVT 板鍛造工程図

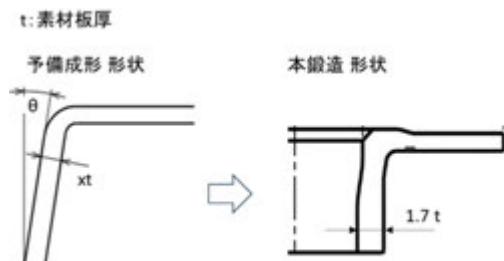


図11 予備成形 形状検討

本工法の最大の要点は、予備成形形状における板厚  $xt$  とテーパ角度  $\theta$  という重要パラメーターに対して条件振り解析を行い、座屈の有無を確認したことである。図12、図13にその結果を示す。予備成形から本鍛造工程まで両立可能なテーパ角度  $\theta$  を選定し、板厚比を適切に配分することで、YEC 板鍛造工法による初の量産部品である YECVT 駆動部品では、従来にない高い増肉率(素材板厚  $t$  に対するボス板厚1.7倍)を達成している。材料の座屈を防止するには、金型による拘束条件の設定が不可欠である。圧縮成形の初期段階では非拘束状態であるが、成形が進行するにつれて外径部および内径部を拘束することで座屈を防止している(図14)。

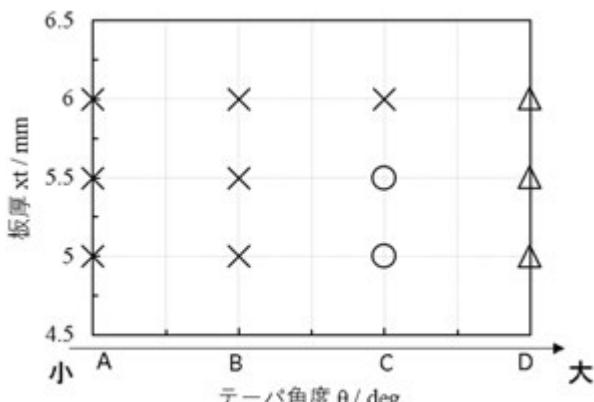


図12  $\theta / xt$  条件振り解析結果

- : 予備成形、本鍛造ともに座屈なし
- ×: 予備成形にて座屈あり
- △: 予備成形座屈なしだが本鍛造にて座屈あり

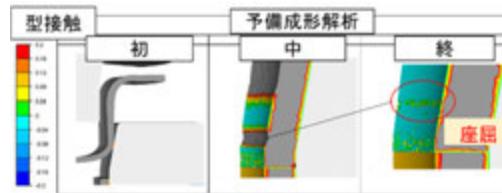


図13 鍛造座屈 NG 例

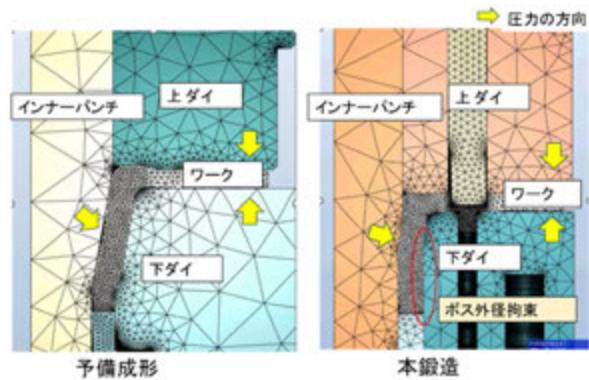


図14 鍛造増肉 成形過程

### 3-3. カーボンニュートラルおよび工程削減への取組み

本開発で採用した YEC 板鍛造工法は、鍛造完成品までの全工程を冷間成形で校正しており、従来の熱間鍛造品と比較してカーボンニュートラル効果が高い。鍛造工程全体での  $\text{CO}_2$  排出量は約56%削減される見込みである(図15)。また、標準的な絞り成形では多工程となるため、型数およびショットの削減が生産性向上および  $\text{CO}_2$  削減の鍵となる。量産化を実現した YECVT 板鍛造部品では、絞り工程における絞り率配分および荷重バランスの最適化により、ショット数を従来の5ショットから2ショットへ削減し、生産性向上と環境性能の両立を達成している。今後も、従来の成形手法にとらわれず、材料特性を活かした自在な成形制御により、さらなる工程削減を目指した多様な成型方案を検討していく。

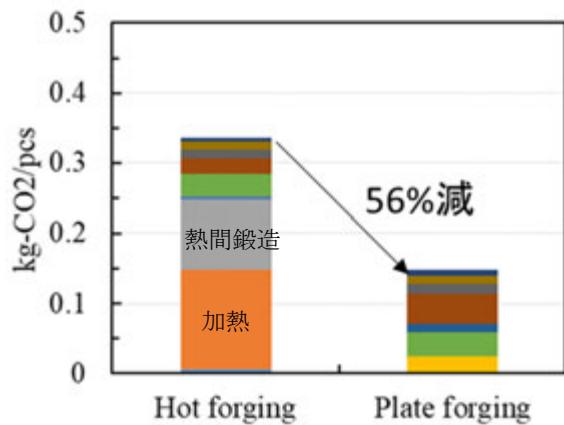


図15 鍛造ワーク1個あたり  $\text{CO}_2$  排出量

## 4 今後の展開

インドネシアにて量産を開始した新「NMAX」のYECVTスクリュウ部品を皮切りに、今後も量産効果の高い部品への技術展開を図り、ASEAN各拠点への導入を進めることで、現地量産機種のコスト低減および商品価値の向上に寄与する方針である。さらに、従来のモーターサイクル部品に加え、EV製品などへの適用範囲を拡大し、複合塑性工法(板鍛造工法)による新たな付加価値を低コストで提供していく考えである(図16)。

### ■著者



田丸 翔吾  
 Shogo Tamari  
 ヤマハモーター  
 エンジニアリング株式会社  
 生産技術部



北郷 博成  
 Hironari Hongo  
 ヤマハモーター  
 エンジニアリング株式会社  
 生産技術部

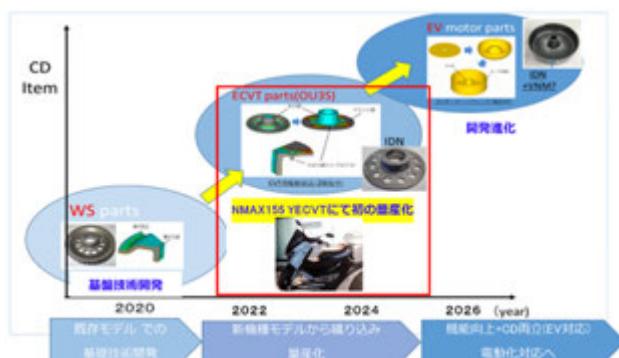


図16 YEC 板鍛造技術展開

## 5 おわりに

ヤマハ発動機グループにおけるYEC板鍛造工法は、鍛造成形領域とプレス板金成形領域の中間に位置し、技術開発の観点から今後さらなる進化が期待される領域である。また、カーボンニュートラルにおいても有用な技術であり、今後は材料メーカーとの連携を図り、リサイクル材料を含めた素材段階からのCO<sub>2</sub>排出削減とコスト低減の両立を目指した新たな取り組みを推進していく。

今後もヤマハ発動機グループの新たな商品価値の向上を目指し、これまでにない新工法をタイミングで提案することで、商品のコスト低減、機能向上およびCO<sub>2</sub>削減を両立させる“物創り技術”的提供を継続していく所存である。

### ■参考文献

- [1]「板鍛造工法によるパワートレイン部品のコストダウン開発」北郷博成ほか:ヤマハ技報, 54, (2018) 15-18 [https://global.yamaha-motor.com/jp/design\\_technology/technical/publish/pdf/browse/54gs03.pdf](https://global.yamaha-motor.com/jp/design_technology/technical/publish/pdf/browse/54gs03.pdf)
- [2]「走りの幅を広げる 新型「NMAX」用「YECVT」の開発」勝山祐紀ほか:ヤマハ技報, 59(2024) 56-59 [https://global.yamaha-motor.com/jp/design\\_technology/technical/presentation/pdf/browse/59gs01.pdf](https://global.yamaha-motor.com/jp/design_technology/technical/presentation/pdf/browse/59gs01.pdf)

# マリン用水素直噴エンジンと水素ボートシステムの開発

Development of a Hydrogen Direct-Injection Engine and Hydrogen Boat System for Marine Applications

鬼形 朋良 平本 成実 弘田 雅士 能勢 幸憲 Hien H. Nguyen

## 要旨

プレジャーボートの動力に必要なエネルギーは自動車の約10倍と見積もられており、電動化でのカーボンニュートラル達成の実現にはバッテリの大容量化など多くの課題がある。そこで本研究ではカーボンニュートラルの実現を目指し、水素直噴エンジンを適用した水素ボートシステムの開発を行った。エンジン開発では異常燃焼を抑制し、信頼性を確保しつつ、自然吸気水素直噴エンジンとして航走に適した性能を実現した。また水素ボートシステムは従来のプレジャーボートのパッケージングを維持し、水素タンクをはじめとした水素供給システムを搭載し、水素を安全に取り扱うため、水素に特化したフェールセーフシステムを開発した。

## Abstract

The energy required to power a pleasure boat is estimated to be approximately ten times that of an automobile, and achieving carbon neutrality through electrification presents numerous challenges, including the need for significantly larger battery capacity.

In this study, we developed a hydrogen direct-injection engine system for pleasure boats with the aim of realizing carbon neutrality.

In the course of engine development, we successfully suppressed abnormal combustion, ensured reliability, and achieved performance suitable for cruising with a naturally aspirated hydrogen direct-injection engine.

Furthermore, the hydrogen boat system maintains the packaging of a conventional pleasure boat while incorporating a hydrogen supply system, including hydrogen tanks, and features a hydrogen-specific fail-safe mechanism designed to ensure safe operation.

## 1 はじめに

地球環境と人々の豊かな活動を持続させるには、マリン製品におけるカーボンニュートラル対応が不可欠である。プレジャーボートの動力に必要なエネルギーは自動車の約10倍とされており、特に大型船外機ではバッテリおよびモータへの置き換えが困難と予測される。そこで、当社は環境性能と商品性の両立を目指し、水素船外機ボートの可能性を探求した。

水素エンジンの開発においては、エンジン単体の開発に留まらず、水素供給系やフェールセーフ機構なども含めたシステム全体の実現性を検討することが重要である。加えて、安全な運用方法の確立やマリン製品への適用性の評価も、開発初期からの重要な検討項目とした。開発は社内の各部門をまたいだ連携に加え、社外の専門パートナーとの協力により(図1)、それぞれの強みを活かしながら技術を蓄積し、着実に前進している。

本稿では、水素を安全に使用しつつ、エンジンの持つ力強さを維持しながら、使用時に二酸化炭素を排出しないボートの試作開発の状況について報告する。

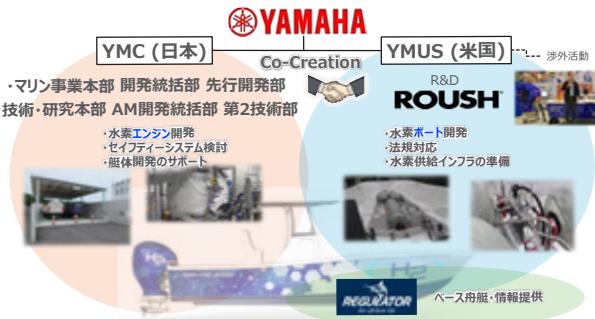


図1 開発パートナー

## 2 エンジン開発

### 2-1. 概要

本開発では、“ヤマハ船外機”のフラッグシップとして市場から高い評価を受けているガソリン船外機「F450A」(V8 5.6L)をベースに、水素燃料対応に向けた仕様変更を行った(図2)。「F450A」を選定した理由は、大出力帯への挑戦により水素エンジンの可能性をより広く探求できることに加え、直噴レイアウ

トが採用されており、水素用直噴インジェクタの搭載に大幅な設計変更を要しないためである。開発においては、自然吸気水素エンジンとして航走に適した性能を目標とし、水素燃焼に特有の異常燃焼への対応や、燃焼特性の違いに起因する部品への対策を講じることで信頼性の確保を図った。



図2 水素エンジン船外機

表1 参考\_「F450」(ガソリン仕様)のスペック

Maximum power	450 HP
Bore × stroke	96×96mm
Compression ratio	12.3
Number of cylinders	8
Displacement	5560cm <sup>3</sup>
Fuel injection method	Direct injection

水素エンジンとしての主な変更点は、水素供給配管および専用インジェクタの装着である。さらに、水素燃焼に合わせて、ピストンとスパークプラグの仕様を変更し、インジェクタ先端には筒内に噴射する水素噴流を最適化するアダプタを装着した。これらの改良により、当初の出力性能と比較して約40%の向上を確認した。現時点でのエンジン性能曲線を図3に示す。同図にて水素エンジンでは高回転域においてガソリンエンジンに對して出力が低下する傾向にあるが、これは点火前に水素が着火する異常燃焼(以下ブレイク)の回避および耐久性への配慮から、最大筒内圧を意図的に抑制した結果である。有効とみられる各構成部品の選定に関しては、「F450A」を単気筒化したエンジン(図4)による試験結果をフィードバックすることで、点火プラグ仕様をはじめとした各部品の純粋な性能を効率的に把握することができた。なお、更なる出力向上に向けて研究開発は現在も継続中である(図5)。

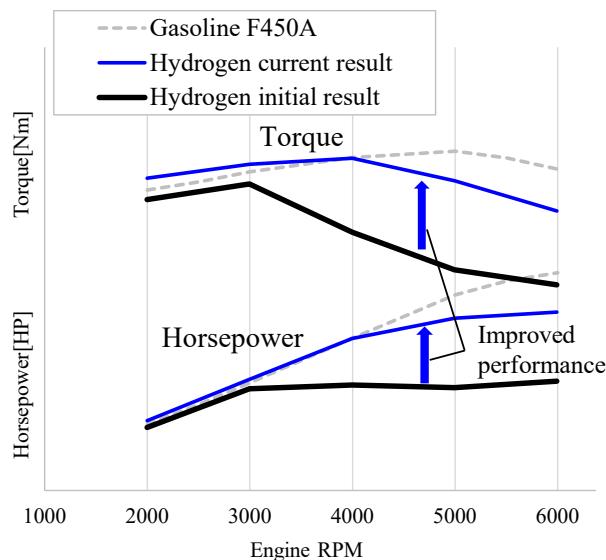


図3 エンジン性能曲線

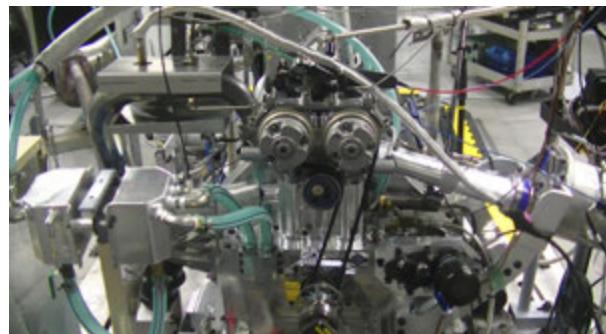


図4 試作単気筒水素エンジン台上試験



図5 水素船外機台上試験

## 2-2. 異常燃焼への対応

水素は燃焼室内へ充填される際の流動状態において空気と混合しにくく、また、ガソリンと比較して着火しやすい特性を有する。このため、燃焼室内に残留した熱源や混合気の不均質性を原因として、図6に示すようなブレイクが発生しやすいことが知られている。したがって、ブレイクの抑制は性能向上および信頼性確保における重要な課題である。

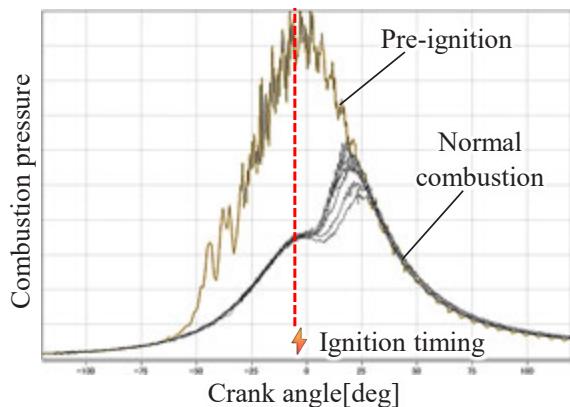


図6 プレイグ発生時の燃焼圧イメージ

本エンジンにおいても、他の多くの水素エンジンと同様に高回転域で燃焼安定性が低下し、プレイグが発生しやすい傾向が確認された。筒内流動のシミュレーション結果からは、水素における筒内混合気の均質度が低く、可燃範囲が局所的に偏在することが明らかとなり、これがプレイグの要因であると考えられた。この課題に対し、筒内混合気の均質度向上を目的として、吸気ポートを高流量ポートおよび高タンブルポートへ変更し、筒内流動場が混合気形成に与える影響をシミュレーションにて評価した。また、インジェクタ先端にアダプタを装着し、アダプタ先端の噴流孔形状を変更することで、噴流設計による改善の可能性についても検討を行った。これらのシミュレーション結果から以下の知見を得た。

- 筒内に直接噴射される水素噴流は、噴射直後から噴射場にとても強い乱流エネルギーを生成する。
- 筒内均質度の向上にはタンブル流の強化が有効である。当エンジンでは、インジェクタを燃焼室ボア中心にマウント(センタマウント)することで、水素噴流によるタンブル強化(噴流によるタンブルアシスト)が期待できる。
- センタマウントインジェクタのベース噴孔(円形単孔)による噴射では、噴射直後のタンブルは強まるが、その効果は持続せず、筒内の均質度も低く、局所的なリッチ領域が残存しやすい。

以上の結果を踏まえ、図7に示す先端形状を試作した。当該アダプタは図8の噴流コンセプトに示す通り、インジェクタから噴射された水素がガイド構造に沿って扇状(Fan-shaped)に拡散して噴射することで筒内の均質度を高める効果を狙った。また扇状の噴流を排気側に傾けることにより水素の噴流によって強いタンブル流を生成し(図9)、圧縮上死点前まで高い乱流強度を維持できる(図10)ことが分かった。結果として、図11および図12に示すように、円形単孔のBase水素噴射に比べて筒内の均質度を向上させることができた。図12は筒内のプレイグが発生し易い比較的リッチなLambda領域の存在分布を水素

噴流改善前後で比較したものである。

※図9から12は同噴射量、同噴射時期での比較結果

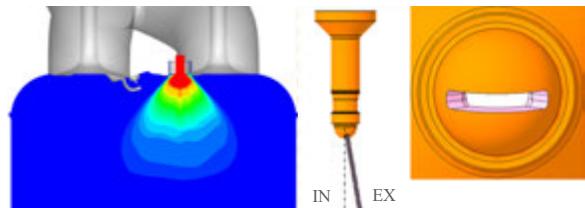


図7 噴流改善アダプタ

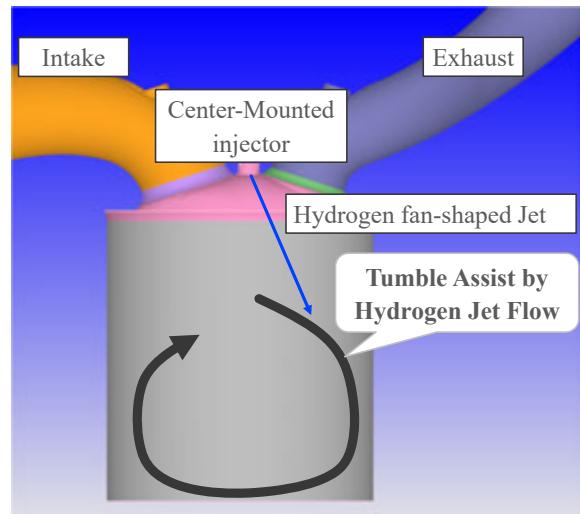


図8 噴流コンセプト

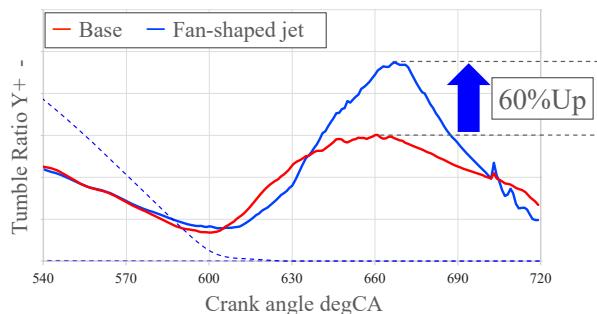


図9 噴流違いによるタンブル流の差

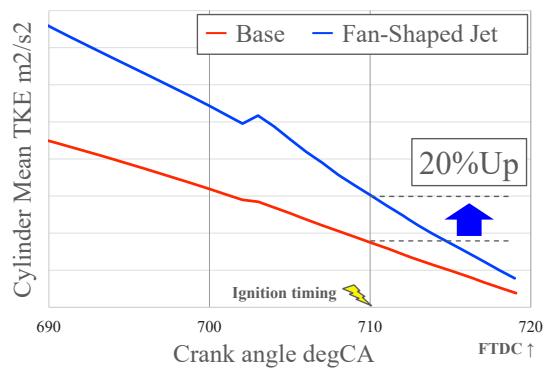


図10 圧縮上死点付近の乱流強度の差

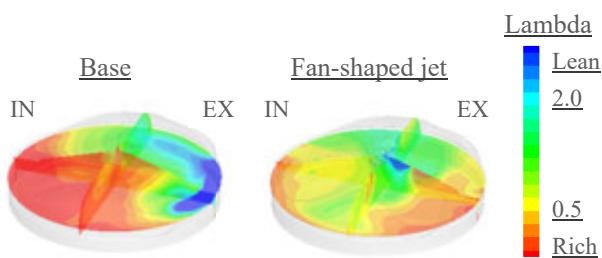


図11 筒内均質度シミュレーション結果 @20BTDC

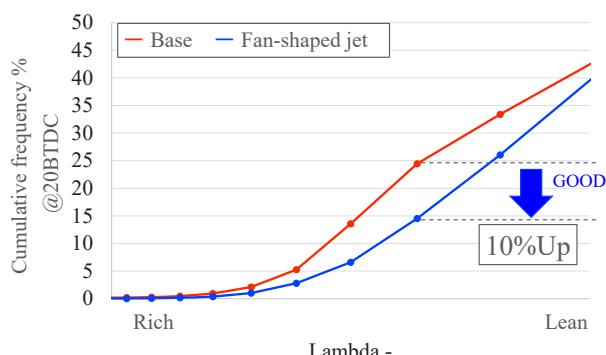


図12 噴流デザイン改善によるリッチ領域の低減効果

本アダプタを実機に適用することで、燃料投入量を増量し、Lambdaをリッチ側に設定してもプレイグの発生を抑制することが可能となり(図13)、出力性能および信頼性の向上を実現した。

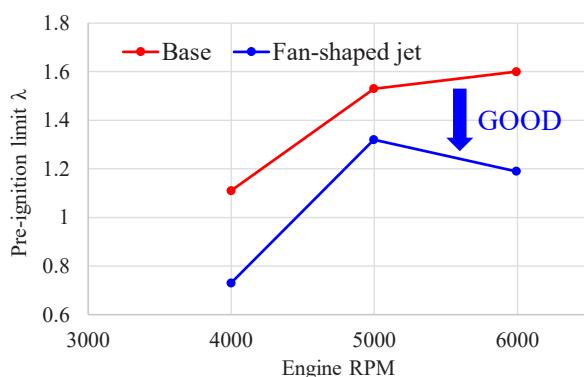


図13 実機での筒内均質度改善によるプレイグ発生λの差

さらに、混合気の圧縮による自着火を抑制する目的で、圧縮比をガソリン仕様よりも低下させた。加えて、残留熱による自着火を防止するため、ガスケットの少ないスパークプラグへ変更するなど、複合的な対策を講じた。

以上のエンジン単体開発を踏まえ、当エンジンをボートへ搭載するにあたっては、万が一プレイグが発生した場合に備え、フェールセーフシステムによりエンジンの破損防止を狙っている。ボートに搭載した水素制御システムは、エンジンの燃焼圧

を常時監視しており、プレイグにより規定値を超える異常燃焼圧を検知した際には、水素インジェクタの噴射を停止することで、エンジンを速やかに停止させる。

### 2-3. 信頼性の確保

水素燃焼はガソリン燃焼と比べて高温となり、かつ燃料に気化潜熱がないため、燃焼室温度が上昇する傾向にある。この燃焼温度上昇が原因とみられるバルブシート摩耗を試験中に確認したことから、摩耗の発生原理(図14)をもとにバルブシート材質を見直し、対策を講じた(図15)。また、水素燃焼により発生する水分に起因する錆への対策として、バルブおよびバルブシートの材質の変更を実施した。発生する水によるオイルの希釈については、定期的な交換により大きな問題は生じていないが、長期使用に伴う劣化に対して調査と検討を継続している。

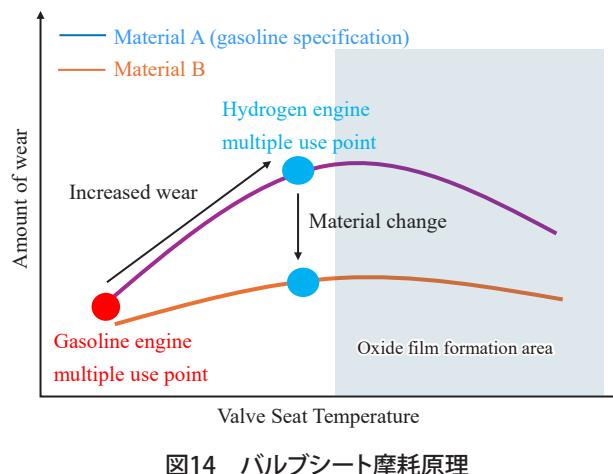


図14 バルブシート摩耗原理

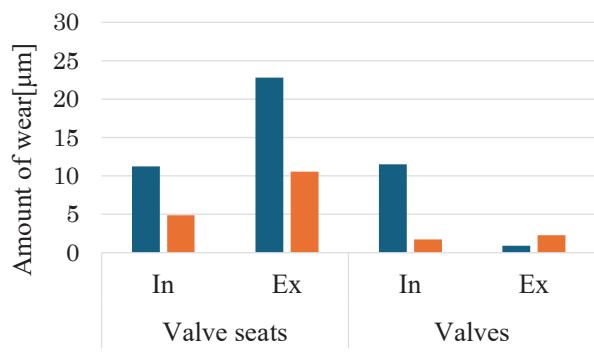


図15 バルブシート摩耗量

## 3 ボート開発

ボート開発にあたっては、大型船外機の世界最大市場である米国において、“ヤマハモーターコーポレーション,U.S.A.(以下YMUS)”が中心となり、ボートビルダーの“Regulator社”、水素

タンクシステム設計を担当する“Roush社”、および“ヤマハ発動機株式会社(以下当社)”を含む4社による協力体制を構築し、共同で開発に取り組んだ。ボートの作成から航走テストは米国にて実施している。

### 3-1. ベースボート選定

船外機は、さまざまな形状や大きさのボートに取り付け可能である必要があり、高い汎用性が求められる。本開発では、“Regulator社”的ペイボートタイプ「26XO」(図16)を選定した。「26XO」は入り江や近海に加え、河川や湖などの多様な水域に対応可能であり、優れた操縦性(ドライバビリティ)を備えている。これにより、異なる環境下での船外機の性能を幅広く評価することが可能となる。また、船体サイズや構造が一般的なプレジャーボートに近いため、新たに開発した水素エンジンの汎用性や実用性を検証する上でも適していると判断した。



図16 Regulator 26XO

Regulator「26XO」全長:8.2m、幅:2.83m、乾燥重量:2,871kg

### 3-2. 水素ボートのタンクレイアウト

水素ボートの製作にあたり、水素タンクの大きさ、搭載数、配置方法について“Regulator社”、“Roush社”とともに協議を重ねた結果、デッキ下への搭載するレイアウトを採用した。これにより低重心化と居住性の両立を図った。デッキ下には本来、ローワーハル(船底)を補強する構造体(レインフォースメント)が存在するが、本設計では水素タンクを固定するブラケット等が船体の構造部材を兼ねることで、必要な強度を確保している。

採用したタンクは気体タンク3本で構成され、1本あたりの容積は244L、重量188kg、充填最大圧力は70MPaである。

船体は滑走状態に移行する前にバウ(船首)が上がり、姿勢が大きく変化した後に滑走状態へ移る特性があり、重心位置が後方に偏ると航走性能を損なう可能性がある。このため、タンク3本のレイアウトはバウ側に2本、スター(船尾)側に1本とし、重心バランスに配慮した(図17)。

また、タンク搭載時には、ネック部に緩衝材を設け、航走時の衝撃に耐えられる構造とした(図18)。

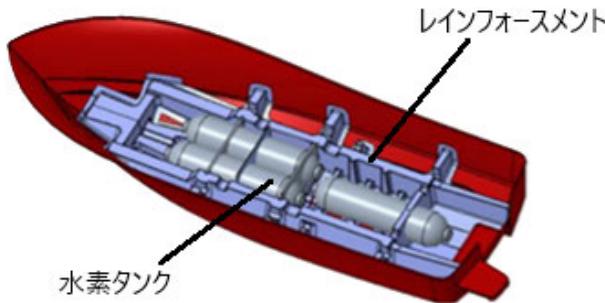


図17 水素タンク搭載レイアウト

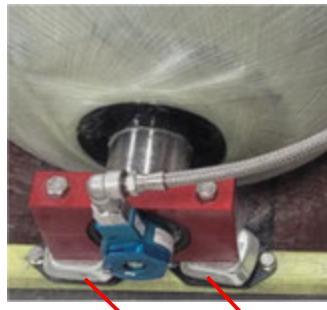


図18 水素タンクネックマウント

### 3-3. 水素供給システム

ボートに搭載する水素供給配管には、水素脆化対策として全てステンレス鋼(SUS材)を採用した。配管継ぎ手には、水素燃料モビリティ分野で実績のある専用のフレアナット方式を用い、メンテナンス性の向上および組み付け時のばらつき吸収を可能とした。

3本の高圧水素タンクから供給される水素は、まずハイプレッシャマニホールドで合流し、各タンク間の圧力を均一化する。その後、減圧機構としてレギュレータを介し、ロープレッシャマニホールドを経由してエンジンへ供給される。供給圧力は10MPaに設定している。

各マニホールドには圧力センサおよび温度センサを装着し、船内に設置された水素システム監視モニタにて常時状態を確認できるようにしている(図19)。水素漏洩や圧力異常などの異常を検知した際には、水素配管各部に設置された電磁遮断弁が作動し、水素の供給を停止するとともに、配管内の水素をバージするシステムが作動する。

水素は着火しやすく可燃域が広いため、バージ時にデッキ内部へ滞留すると火災に繋がる恐れがある。そこで、水素が非常に軽い気体である特性を踏まえ、確実に大気中へ開放することを目的として、バージラインを船のルーフ部に設ける構造とした(図20)。

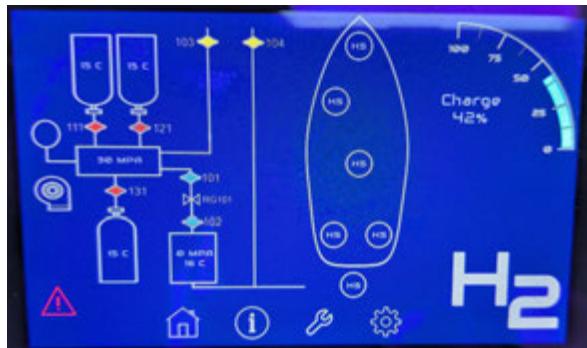


図19 水素システム監視モニタ

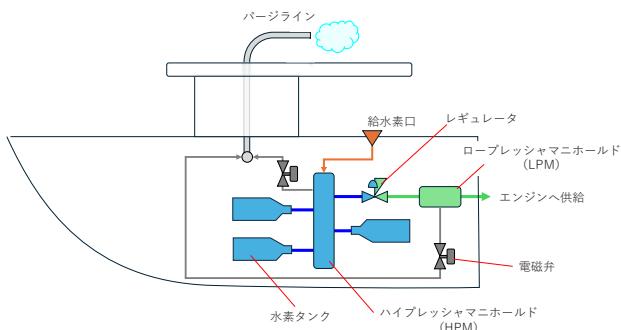


図20 水素供給システム

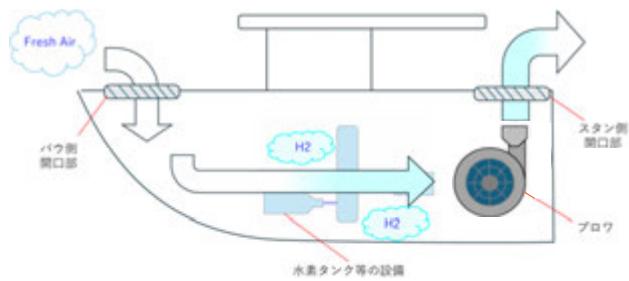


図21 艇体内部の水素掃気

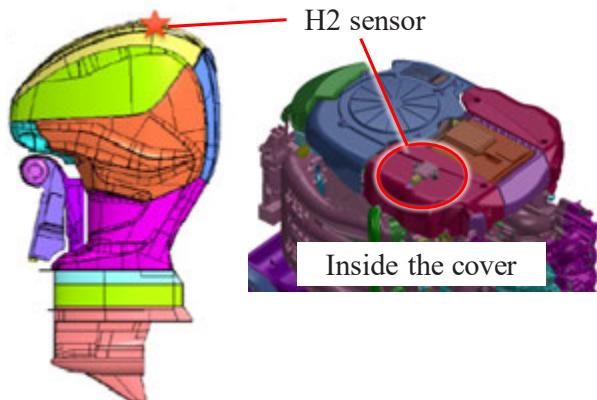


図22 船外機の水素センサ装着位置

### 3-4. 水素漏洩への備え

水素は分子サイズが小さく、漏れやすい性質を有する。このため、万が一の水素漏洩に備え、艇体内部5カ所に水素濃度センサを配置した。

また、漏洩時には艇体内部に水素ガスが滞留することが想定されるため、デッキ部に掃気構造を設け、バウ側に吸気開口部、スタン側に排気開口部を配置し、プロワによる強制掃気を行う構成とした(図21)。エンジン始動前には必ず水素ガス掃気システムを作動させ、プロワで艇体内を掃気し、水素ガス濃度があらかじめ設定した閾値以下にならない限り、エンジンを始動できない制御としている。

一方、エンジン側で水素が漏洩した場合、船外機ではカウル内の上方に水素が滞留することが想定される。このため、カウル内のエンジン最上部に水素濃度センサを設置し、早急な漏洩検知を可能とした(図22)。なお、船外機はチルト操作によりエンジンの傾きが変化する特有の使用方法があるため、いかなるチルト角度においても検出可能となるよう、センサは適切な位置に配置している。

### 3-5. 転舵時の水素配管の対応

船外機は、船の転舵、トリム操作、チルトに応じて、船体に対して上下左右に傾ける必要がある。このため、エンジンと船体を接続する水素配管にも、船外機の可動範囲に追従する柔軟性が求められる。本システムでは、金属性のフレキシブル配管を採用し、ボート側の水素配管に接続している。しかし、高圧の水素供給時には、フレキシブル配管が硬直し、可動性を阻害する懸念が生じる。この課題に対しては、フレキシブル配管の取り回しにゆとりを持たせたレイアウトとし(図23)、曲げ角度を低減する工夫を施した。さらに、船外機側に水素配管用の堅牢なブラケットを新設することで(図24)、従来の船外機と同等の転舵角、トリム角、チルト操作を可能とした。



図23 水素フレキシブル配管の取り回し

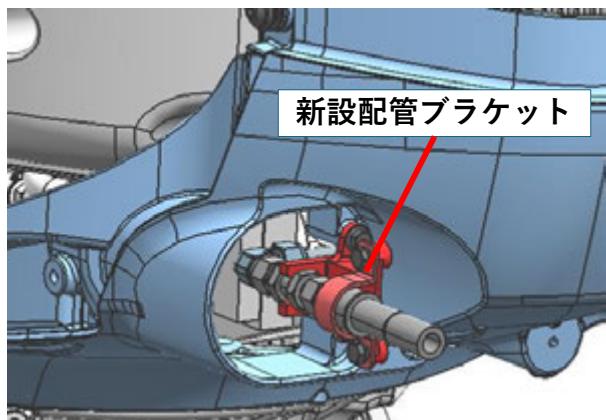


図24 水素配管プラケット

## 4

## 水素充填

## 4-1. 水素船外機ボートへの水素供給システム

水素供給システムは、海洋環境における安全性、性能、環境要件を満たすよう設計されている。本システムでは、15~30分という迅速な燃料補給時間を実現し、最大70MPaに対応した先進的なタンク技術を採用した。燃料インターフェースには、“SAE J2600”準拠のコネクタと、耐腐食性に優れた材料を使用することで、海洋環境における高い耐久性を確保している。急速充填時の高圧および熱負荷に対応するため、船内タンクには圧力開放装置と温度センサを内蔵し、安全性を確保した。さらに、水素漏れ検知機能、自動供給停止機構、適切な接地、防爆仕様の構成部品、効果的な換気など、包括的な安全対策を講じることで、高い信頼性と安全性を備えた給水素システムを実現している(図25)。



図25 陸上で水素供給作業

## 4-2. 移動式水素供給装置 (Air Products HF-150)

水素供給装置として採用した“Air Products HF-150”は、貯蔵と供給を一体化したコンパクトな移動式水素燃料供給装置であり、1台のトレーラーにより構成される(図26)。本装置は、一回の積載で最大80kgの35MPa水素を供給可能であり、設置寸法は長さ7.9m×幅2.6m×高さ2.1mで、一時的な設置にも柔軟に対応できる。本プロジェクトにおいて70MPaではなく35MPaを採用した理由は、操作の簡便性、および、コストの観点によるものである。圧力が低いため複雑な圧縮システムが不要となり、メンテナンス頻度の低減および運用の信頼性向上が期待できる。さらに、35MPaシステムは材料への負荷が少なく、熱安定性にも優れているため、初期段階における海洋用途において、安全性およびコスト効率の両面で優れている。安全な運用には、訓練を受けた作業者による取り扱い、定期的なリークチェック、静電気防止のための接地、供給元であるAir products社の安全プロトコルの遵守が必要となる。本システムは操作性にも優れており、マリーナドッグでの海洋用途に適している。テスト期間を通じて“HF-150”は高い信頼性を発揮し、水素船外機の初期開発フェーズにおいて、実用的かつ有効なソリューションとして機能した。なお、装置の立ち上げや運用に際しては、“Air Products社”より現地対応を含む技術サポートが提供され、円滑なシステム構築に大きく貢献いただいた。



図26 移動式水素供給装置 (Air Products HF-150)

## 5 まとめ

本研究では、水素船外機の実用性と課題を多角的に検証した。

### 5-1. エンジン開発

出力性能と信頼性確保のため、水素燃焼に合わせた変更や工夫が必要であり、これらは船外機に限らず他の水素エンジンモビリティにも共通する課題である。本研究では、水素噴射による吸気タンブル流の強化と噴流形状の工夫により筒内均質度を高め、課題の1つである異常燃焼を抑制した。

また、水素の広い可燃範囲は、セッティングの柔軟性を高め、短時間でボート特性への適合を可能にした。

### 5-2. ボート開発

大型水素タンクの搭載による航走性能や艇体剛性への影響は、搭載位置やマウント方法の工夫により、最小限に抑制した。高圧水素配管の張力による船外機の可動制限も、フレキシブル配管の選定により解消し、ボートのパッケージングとして十分に成立することも分かった。

航走時には、排出ガス特有の臭気がなく快適性が向上し、V8エンジンの高回転域での加速感および心地よいエンジンサウンドは維持されており、水素燃料と大型船外機の高い親和性を体感することができた(図27)。



図27 水素船外機プレジャーボートの航走状態

### 5-3. 水素供給

高圧タンクと安全機構を備えたボートの水素供給システムにより、比較的迅速かつ信頼性の高い充填が可能となった。テストを実施した米国では、ボートへの水素供給は移動式の水素供給装置により実現することができた。マリーナドックにおいて水素ステーションを建設することなく、比較的手軽な水素供給の可能性を探ることができた。

## 6 おわりに

水素船外機ボートは、2024年のマイアミインターナショナルボートショーにて初公開され、その後も国際展示会や講演を通じて技術発信を継続している。これらの活動を通じて、マリン製品のカーボンニュートラル対応および新たなエンジン技術に対する高い期待と関心が多くの関係者から寄せられた。2025年1月には、本システムを用いた実水域での航走試験を実施し、水素船外機プレジャーボートとしての基本性能を確認した。現時点では、性能向上および運用実証に向けた基準点を確認できた段階にあると認識している。今後、製品化を見据えた解決すべき課題は多岐にわたる。まず、今回のテスト環境とは異なる幅広いマリン用途の条件に適合した水素供給インフラの整備が必要である。エンジン技術においては、燃費改善、出力向上、排気中の NOx 低減の両立といったさらなる性能向上が求められる。加えて、ユーザーが安心して利用できるよう、ハードウェアおよびシステム両面における安全性とメンテナンス性の強化も不可欠である。今後も技術の蓄積と改良を継続することで、単なるカーボンニュートラル対応という環境価値にとどまらず、商品性および市場性の両面からも選ばれる価値ある製品の実現を目指し、挑戦を続けていく所存である。

### ■著者



鬼形 朋良  
Tomoyoshi Onikata  
技術・研究本部  
AM 開発統括部  
第2技術部



平本 成実  
Narumi Hiramatsu  
技術・研究本部  
AM 開発統括部  
第2技術部



弘田 雅士  
Masashi Hirota  
マリン事業本部  
開発統括部  
先行開発部



能勢 幸憲  
Yukinori Nose  
マリン事業本部  
開発統括部  
先行開発部



ヒエン H. グエン  
Hien H. Nguyen  
Yamaha Motor Corporation, U.S.A.  
Marine Innovation Engineering Division

■映像ライブラリー

水素船外機ボート 開発の軌跡 紹介動画

[https://global.yamaha-motor.com/jp/design\\_technology/  
technical/mov/YMC-TR60\\_TechPaper\\_01.html](https://global.yamaha-motor.com/jp/design_technology/technical/mov/YMC-TR60_TechPaper_01.html)



# 船外機のクランク打音予測のための ジャーナル荷重予測手法の開発

Developing a Journal Load Prediction Method for Crank Knock in Outboard Motors

井上 祐人 高比良 拓人 藤田 英之 小林 光司 近藤 拓 高橋 徹

## 要旨

船外機の商品性向上に向け、低騒音化が求められている。騒音の主要因の一つとして、クランク軸が主軸受を叩くことに起因する“クランク打音”が挙げられる。この騒音を予測するには、主軸受に作用する動的荷重の予測が必要であるが、計測の難易度が高いため、これまで十分な検証が行われていなかった。

本報では、エンジン実稼働時における主軸受荷重の計測手法とその結果、さらに計測結果を検証データとして用いた予測モデルの構築結果について報告する。荷重計測には、ひずみ測定値から荷重を推定する手法に着目し、実稼働時の荷重を推定した。予測モデルの構築に際しては、クランク軸およびエンジンケースを弾性体とした動解析モデルを構築した。実測結果と解析結果を比較し、時間領域および周波数領域における解析結果の精度を検証した結果、主軸受の動的荷重を高精度に予測可能であることを確認した。

## Abstract

To enhance the market competitiveness of outboard motors, reducing noise has become an important requirement. One of the primary noise sources is “crank knock,” which occurs when the crankshaft strikes the main bearings. Accurate prediction of this noise requires estimating the dynamic loads acting on the main bearings. However, due to the high difficulty of direct measurement, sufficient verification has not been achieved to date. In this study, we report on a measurement method for main bearing loads during actual engine operation, the obtained results, and the development of a predictive model validated using these measurements. For load measurement, we focused on a method that estimates bearing loads from strain measurements, thereby enabling estimation of loads during actual engine operation. In constructing the predictive model, we developed a dynamic analysis model that treats the crankshaft and engine case as elastic bodies. By comparing measured results with analytical results in both the time and frequency domains, we confirmed that dynamic loads on the main bearings can be predicted with high accuracy.

## 1 はじめに

船外機エンジンから発生する騒音は、船上の乗員に不快感を与えることがあるため、その低減が求められている。エンジン騒音は、燃焼騒音と機構系起因の騒音に大別され<sup>[1][2]</sup>、その発生メカニズムは多岐にわたる<sup>[3]</sup>。ここでは、特に機構系騒音の主要因であるクランク軸挙動に起因する騒音(以下、クランク打音)に注目する。クランク打音は、図1に示すように、①クランクピニオンにて力を受けたクランク軸が、②変形を伴って回転運動した際、③その運動を支える主軸受(すべり軸受)が荷重を受け、④エンジンケースにジャーナル荷重が伝達され、ケースの振動となって音として放射されるというメカニズムで発生する。

本開発では、クランク打音予測の第1ステップとして、クランク軸を支える主軸受のジャーナル荷重予測手法を開発した。供試には、当社製品「F425」に搭載されたV8エンジンを用い、ジャーナル荷重を予測する解析モデルの構築と、解析結果との比較を目的としたジャーナル荷重の実測を実施した。荷重の実

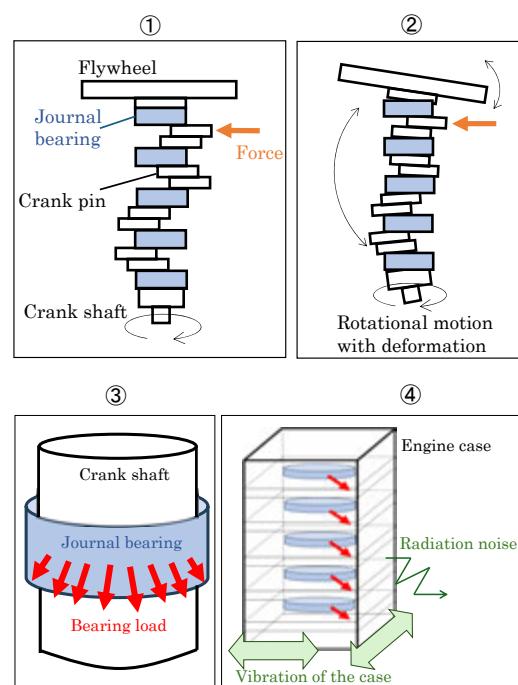


図1 クランク打音発生のメカニズム

測では、ひずみ測定値から荷重や圧力分布を推定する手法に着目し<sup>[4][5]</sup>、主軸受孔周辺のひずみ測定値から、実稼働中のジャーナル荷重を推定した。推定に際しては逆解析を用いず、事前に治具を用いて静的荷重を主軸受部に印加した際のひずみデータを測定することで構築した校正式を利用した。また、クランク軸およびエンジンケースを弾性体とした動解析モデルを構築し、実測結果と解析結果を比較することで解析モデルの精度検証を行った。

## 2 ジャーナル荷重の実測

### 2-1. 概要

主軸受孔周辺の圧縮ひずみの測定値を荷重に変換することで、ジャーナル荷重の実測値を得た。図2に示すフローチャートの通り、事前に実施した校正試験により校正係数を取得し、エンジン実稼働時のひずみ測定試験で得られた圧縮ひずみデータを荷重へと変換した。荷重変換に際しては、以下の仮定を置いた。

- ・主軸受孔自体の局所変形モードは、計測対象の周波数帯において無視できる。
- ・実稼働時の圧縮ひずみから荷重への変換は、校正係数を用いた線形関数で表される。

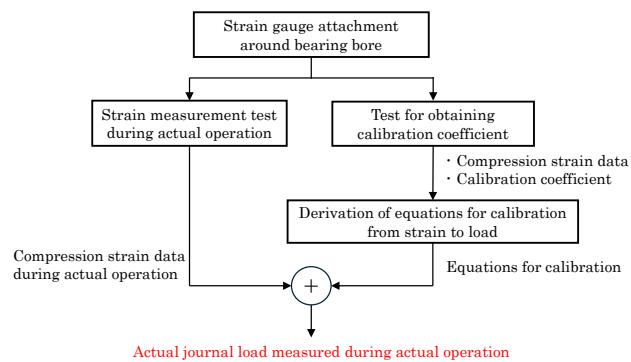
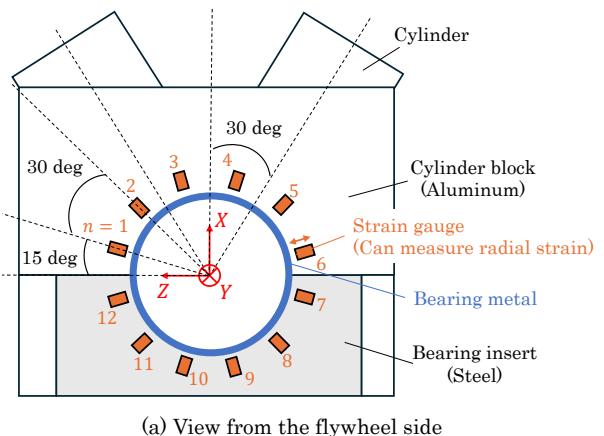


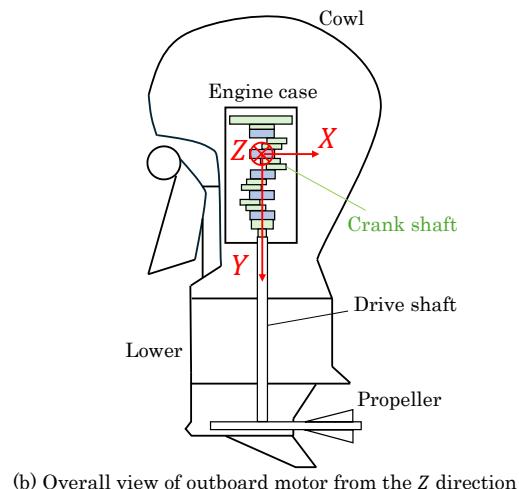
図2 実測ジャーナル荷重取得のフローチャート

ここで図3には、軸受孔周辺に貼付したひずみゲージの位置概要と座標系の定義を示し、図4および図5には実際に貼付したひずみゲージの様子を示す。以後、座標系は図3に示す定義に従う。原点は各軸受孔の中心とし、シリンダバンクの中間方向をX方向、クランク軸に沿った鉛直方向をY方向、X方向に垂直な水平方向をZ方向とする。エンジンケースの軸受孔は、シリンダブロック部(アルミニウム)とベアリングインサート部(鉄)に分かれている。図3のように、軸受孔のフライホイール側(a側)およびドライブ軸側(b側)それぞれの面において、孔近傍に同一円周上で30deg間隔に12枚のひずみゲージを貼付

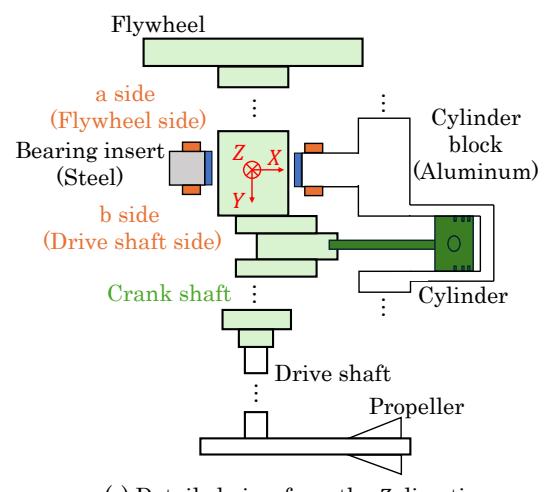
し、軸受径方向の圧縮ひずみを測定できるようにした。各ひずみゲージの位置は、図3(a)に示す角度番号  $n=1, 2, \dots, 12$  により定義する。各ひずみゲージ(アクティブゲージ)の近傍にはダミーゲージを貼付し、アクティブ・ダミー法により温度補償を行った<sup>[6][7]</sup>。



(a) View from the flywheel side



(b) Overall view of outboard motor from the Z direction



(c) Detailed view from the Z direction

図3 座標系の定義とひずみゲージ貼付の概要図

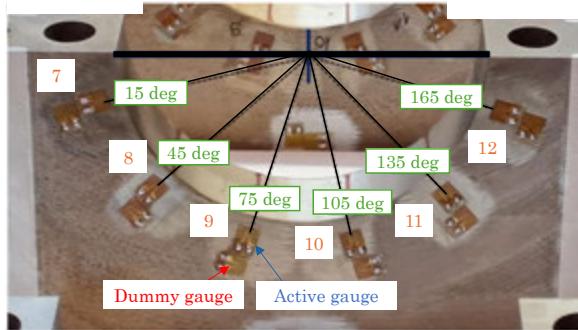


図4 ベアリングインサートに貼付したひずみゲージ(b側)

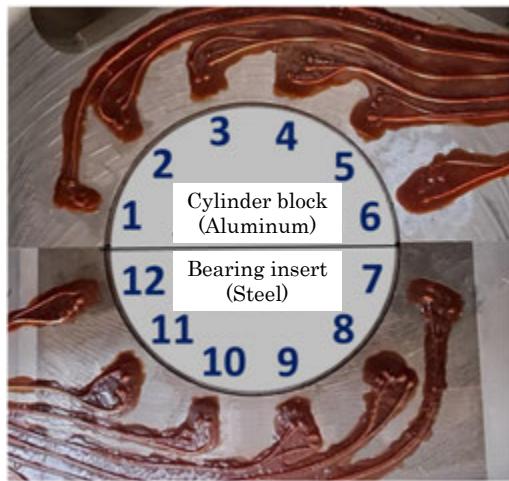


図5 a側に貼付したひずみゲージ(配線保護後)

## 2-2. 校正試験

図6に示すように、1対の半円筒金属をボルトで連結した治具を主軸受に押し当て、既知の荷重  $F_c$  を印加した際の圧縮ひずみ値を記録した。治具の半円筒金属は軸受内径よりもわずかに小さい径を有しており、ボルト軸力によって金属が押し当てられることで、荷重の印加が可能となる。ボルト軸力は、治具に貼付したひずみゲージにより計測し、印加荷重の大きさを算出した。角度方向  $n=1, 2, \dots, 12$  に対してそれぞれ治具を押し当て、各方向に荷重を印加した際の圧縮ひずみ値を記録した。図7に示すように、ある角度方向  $n$  に荷重を印加した条件について、印加荷重  $F_{c_n}$  を、角度  $n-1, n, n+1$  位置の a 側および b 側、計6カ所の圧縮ひずみ値  $\epsilon_{a_{n-1}}^n, \epsilon_{b_{n-1}}^n, \epsilon_{a_n}^n, \epsilon_{b_n}^n, \epsilon_{a_{n+1}}^n, \epsilon_{b_{n+1}}^n$  を用いてフィッティングした。ここで添え字  $a, b$  はそれぞれ a 側 b 側を示し、 $\epsilon_m^k$  は角度  $k$  方向の荷重が印加された際の角度  $m$  位置における圧縮ひずみを表す。

$$F_{c_n} = K_n \left\{ C_{n-1} \left( \epsilon_{a_{n-1}}^n + \epsilon_{b_{n-1}}^n \right) + C_n \left( \epsilon_{a_n}^n + \epsilon_{b_n}^n \right) + C_{n+1} \left( \epsilon_{a_{n+1}}^n + \epsilon_{b_{n+1}}^n \right) \right\} \quad (1)$$

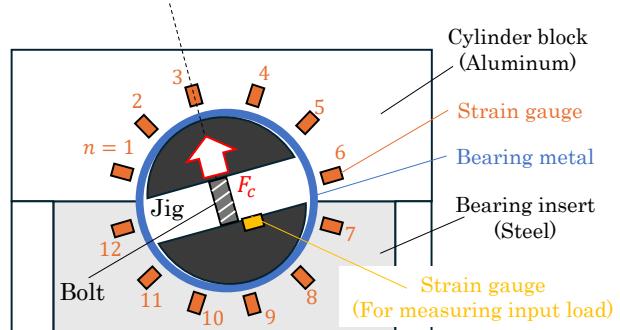


図6 校正試験の概要図

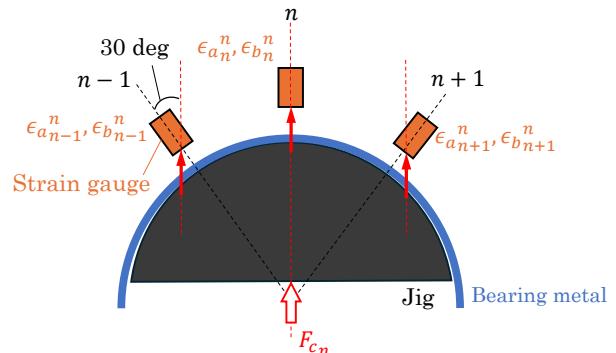


図7 校正試験時のフィッティングで考慮するひずみ

ここで、 $K_n$  は校正係数、 $C_n$  は印加荷重方向に対するひずみゲージの傾きを補正する係数である。荷重  $F_{c_n}$  をフィッティングする際には、 $C_{n-1} = C_{n+1} = 1/\cos 30^\circ$ 、 $C_n = 1$  とした。また、角度  $n-1, n, n+1$  位置において、シリンドラブロックとベアリングインサートを跨ぐ場合での荷重 ( $F_{c_1}, F_{c_6}, F_{c_7}, F_{c_{12}}$ ) をフィッティングする際は、シリンドラブロック(アルミニウム)とベアリングインサート(鉄)のヤング率の比率を考え、 $C_{n-1}$  や  $C_{n+1}$  を補正した。例えば  $F_{c_6}$  をフィッティングする際、角度  $n=5, 6$  位置では材質がアルミニウム、角度  $n=7$  位置では材質が鉄であるため、 $C_5 = 1/\cos 30^\circ$ 、 $C_6 = 1$ 、 $C_7 = (1/\cos 30^\circ) * 210 [GPa] / 70 [GPa]$  として、角度  $n=7$  位置の圧縮ひずみをアルミニウム相当の圧縮ひずみに換算してフィッティングを行った。なお、ヤング率は鉄を  $210 [GPa]$ 、アルミニウムを  $70 [GPa]$  とした。

校正試験時の荷重を式(1)に基づいてフィッティングした結果の一例を図8に示す。校正試験では、時間の経過に伴い印加荷重を段階的に増加させた。最大の印加荷重に対してフィッティング誤差が最小となるように、角度方向  $n=1, 2, \dots, 12$  に対して校正係数  $K_n$  をそれぞれ算出した。図8の結果から、実際の印加荷重を精度よく算出可能な校正係数  $K_n$  を得ることができた。

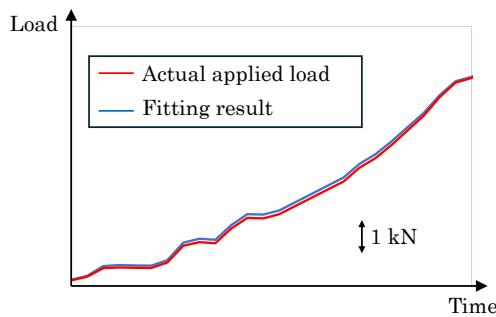


図8 フィッティング結果の例( $F_{c_3}$ )

式(1)に基づく6カ所の圧縮ひずみを用いたフィッティング(6フィッティング)に加え、角度  $n$  位置の a 側・b 側の2カ所のみを用いたフィッティング(2フィッティング)、および角度  $n-2 \sim n+2$  の範囲における a 側・b 側計10カ所の圧縮ひずみ値を用いたフィッティング(10フィッティング)も実施した。比較の結果、印加荷重に対するフィッティング誤差が最も小さかったのは、式(1)に基づく6フィッティングであった。したがって、実稼働時の圧縮ひずみを荷重に変換する際には、6フィッティング手法を採用することとした。

### 2-3. 実稼働ひずみ測定と荷重変換

定常回転条件および回転スイープ条件(スロットル開度全開)において、各ひずみゲージの圧縮ひずみ値を取得した。測定対象外の周波数帯のノイズを除去するため、取得した測定値にはローパスフィルタ処理を施した。

- 実稼働中の圧縮ひずみ値を、実稼働中の荷重に変換するため、式(1)を拡張した。拡張に際しては、以下の2点を考慮した。
- 測定されるひずみ量は、測定部位の構造剛性の影響を受けるため、最大圧縮ひずみを取る測定部位方向が必ずしも荷重印加方向に一致するとは限らない。また、荷重がひずみゲージ間の角度方向に作用する場合も考慮する必要がある。そこで、全角度方向  $n=1, 2, \dots, 12$  における荷重  $F_n$  の総和により、X, Z 方向の実稼働荷重を推定する。
  - 実稼働時には、主軸受が油膜によりクランク軸を支持するが、校正試験では軸受内径よりも小径の金属治具を押し当てた固体接触状態となる。このため、図9に示すように、同一荷重を印加した場合でも圧縮ひずみの発生面積が異なることが想定される。よって、圧縮ひずみ発生面積の違いによる影響を補正する係数を校正式に組み込むこととした。

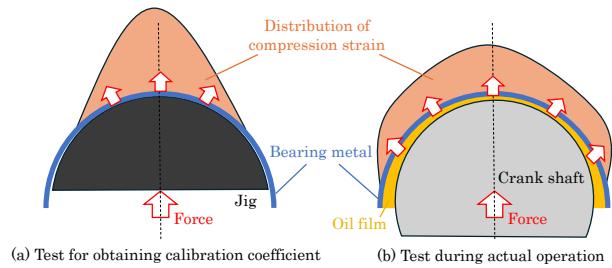


図9 校正時と実稼働時の圧縮ひずみ分布の差(想定)

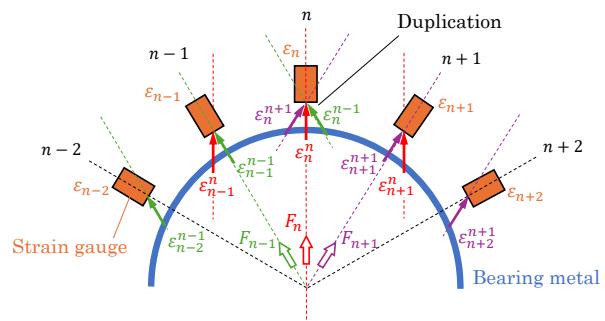


図10 足し合わせの際の重複

(i)について、ある時刻  $t$  における X 方向の実稼働荷重  $F_X(t)$ 、Z 方向の実稼働荷重  $F_Z(t)$  は、各角度方向  $n=1, 2, \dots, 12$  における実稼働荷重  $F_n(t)$  の加算により、以下の式で表される。

$$F_X(t) = \sum_{n=1}^{12} F_n(t) \sin(30^\circ n - 15^\circ) \quad (2)$$

$$F_Z(t) = \sum_{n=1}^{12} F_n(t) \cos(30^\circ n - 15^\circ) \quad (3)$$

式(2)、(3)において  $F_n(t)$  を加算する際には、図10に示すような重複の影響を考慮する必要がある。校正試験では、式(1)に基づき、荷重作用方向の圧縮ひずみ(中心ひずみ)とその両隣の圧縮ひずみ(周辺ひずみ)から荷重を算出している。したがって、実稼働中の角度  $n-1, n, n+1$  位置の圧縮ひずみ  $\varepsilon_{n-1}, \varepsilon_n, \varepsilon_{n+1}$  (a 側・b 側)をそのまま式(1)に代入して  $F_n(t)$  を算出し、これを式(2)、(3)に代入した場合、 $\varepsilon_n$  が中心ひずみ1回、周辺ひずみ2回の計3回計算に用いられることとなり、重複が生じる。そこで、式(4)、(5)、(6)を用いて、各角度位置における実稼働時の圧縮ひずみ  $\varepsilon_n$  を、中心ひずみ成分  $\varepsilon_n^n$  および両隣の周辺ひずみ成分  $\varepsilon_n^{n-1}, \varepsilon_n^{n+1}$  に分解した。ここで、 $\varepsilon$  は実稼働時の圧縮ひずみ、 $\varepsilon$  は校正試験時の圧縮ひずみを表す。

$$\varepsilon_n^{n-1}(t) = \left( \frac{\varepsilon_n^{n-1}}{\varepsilon_n^{n-1} + \varepsilon_n^n + \varepsilon_n^{n+1}} \right) \varepsilon_n(t) \quad (4)$$

$$\varepsilon_n^n(t) = \left( \frac{\varepsilon_n^n}{\varepsilon_n^{n-1} + \varepsilon_n^n + \varepsilon_n^{n+1}} \right) \varepsilon_n(t) \quad (5)$$

$$\varepsilon_n^{n+1}(t) = \left( \frac{\varepsilon_n^{n+1}}{\varepsilon_n^{n-1} + \varepsilon_n^n + \varepsilon_n^{n+1}} \right) \varepsilon_n^n(t) \quad (6)$$

式(4)、(5)、(6)により、角度  $n$  位置における実稼働圧縮ひずみ  $\varepsilon_n(t)$  は、校正試験時に角度  $n-1, n, n+1$  方向へ荷重を印加した際に、角度  $n$  位置で発生する圧縮ひずみ値  $\varepsilon_n^{n-1}, \varepsilon_n^n, \varepsilon_n^{n+1}$  の比率を用いて、以下の3成分に分解した。

- $F_n(t)$ のみが印加された相当の圧縮ひずみ  $\varepsilon_n^n(t)$
- $F_{n-1}(t)$ のみが印加された相当の圧縮ひずみ  $\varepsilon_n^{n-1}(t)$
- $F_{n+1}(t)$ のみが印加された相当の圧縮ひずみ  $\varepsilon_n^{n+1}(t)$

この分解により、ある時刻  $t$  における各角度方向  $n=1, 2, \dots, 12$  の実稼働荷重  $F_n(t)$  を算出するための校正式は次式で表される。

$$F_n(t) = AK_n \left\{ C_{n-1} (\varepsilon_{a_{n-1}}^n(t) + \varepsilon_{b_{n-1}}^n(t)) + C_n (\varepsilon_{a_n}^n(t) + \varepsilon_{b_n}^n(t)) + C_{n+1} (\varepsilon_{a_{n+1}}^n(t) + \varepsilon_{b_{n+1}}^n(t)) \right\} \quad (7)$$

ここで係数  $A$  は、(ii)で述べた校正試験時と実稼働時における圧縮ひずみ発生部の面積の違いを補正するための係数である。係数  $A$  は、図9(a)に示す校正試験時の圧縮ひずみ値から求めた平均圧縮ひずみ  $\bar{\varepsilon}_c$  (単位荷重印加相当)と、図9(b)に示す油膜を介して軸が支持されている実稼働状態で、単位荷重が印加された場合を想定して得られる平均圧縮ひずみ  $\bar{\varepsilon}_o$  の比率により定義される。

$$A = \bar{\varepsilon}_c / \bar{\varepsilon}_o \quad (8)$$

2-2節において、6カ所の圧縮ひずみを用いた6フィッティングが、2フィッティングおよび10フィッティングと比較して高い精度で荷重を再現できたことを踏まえ、校正試験時に角度方向  $n$  に荷重を印加した際の圧縮ひずみ分布は、角度  $n-1 \sim n+1$  の範囲に存在すると仮定する。この仮定に基づき、図9(a)に示す校正試験時に角度方向  $n$  に荷重を印加した際の平均圧縮ひずみ  $\bar{\varepsilon}_c$  は、角度  $n-1, n, n+1$  位置の圧縮ひずみの平均値として定義する。さらに、図9(b)に示す油膜支持状態において単位荷重が印加されている場合の平均圧縮ひずみ  $\bar{\varepsilon}_o$  は、単位荷重を軸受投影面積で除した軸受平均面圧(目安として用いられる圧力)<sup>[8][9]</sup> から求められるひずみとし、以下の式により算出する。ここで、 $W$  は単位荷重、 $L$  は軸受幅、 $D$  は軸受内径、 $E$  はヤング率である。

$$\bar{\varepsilon}_o = W / (LDE) \quad (9)$$

クランク軸に燃焼圧が作用し、最大荷重が発生する際の軸受

荷重を高精度に推定するため、ペアリングインサート部に相当する角度  $n=9$  および 10 位置の係数  $A$  に着目した。まず、校正試験時における圧縮ひずみ  $(\varepsilon_8^9, \varepsilon_9^9, \varepsilon_{10}^9)$  の平均値  $\bar{\varepsilon}_c$  を算出し、これを用いて式(8)、(9)に基づき角度  $n=9$  位置の係数  $A$  を求めた。同様に、角度  $n=10$  位置についても係数  $A$  を算出した。最終的に、これら2つの係数  $A$  の平均値を、補正係数として採用した。

### 3 荷重予測モデル

#### 3-1. 概要

クランク軸の動的挙動を時々刻々と解き、主軸受に作用する力を予測する解析モデルを構築した。解析モデルの構築には、エンジン機構解析ソフトウェアである EXCITE Power Unit R2023.1 (AVL 社)<sup>[10]</sup> を使用した。

#### 3-2. モデル化詳細

構築した動解析モデルの模式図を図11に示す。また、同モデル内で、有限要素モデル (FEM) により表現した部分を図12、13に示す。モデル化範囲は船外機全体とし、クランク軸挙動に着目する観点から、クランク軸が存在する領域のケース剛性を考慮した。具体的には、エンジンケース、排気管、ガイドエギゾーストを弾性体のソリッドモデルとして構築した。さらに、エンジンの支持剛性を反映するため、船外機のマウント剛性を加味した。マウント剛性値は、マウント単体の加振試験により得られた実測値を基に設定した。その他の部分(船外機ロワー)は、重心位置、質量、慣性モーメントのみを考慮した質点としてモデル化し、剛体要素を介してガイドエギゾーストと接続した。クランク軸は、フライホイールからクランク軸端までをソリッドモデルとして、各 FEM は動的縮退を施したうえで動解析モデルに組み込んだ。

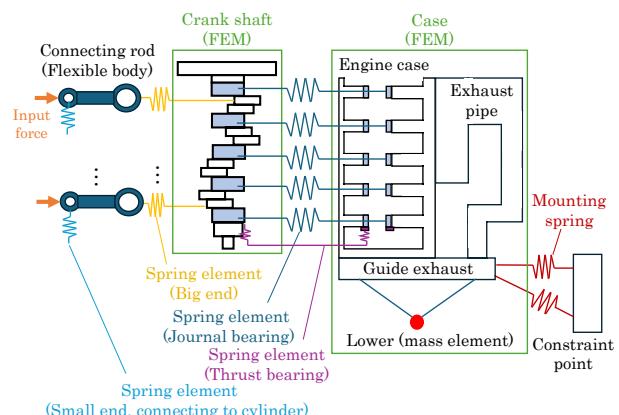


図11 構築したクランク軸動解析モデルの模式図

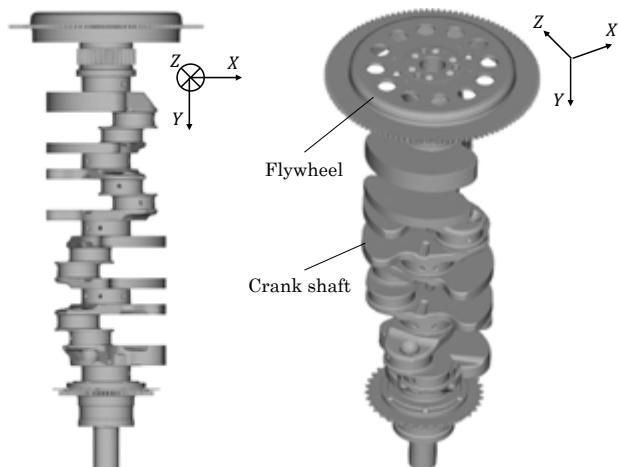


図12 FEMでモデル化したクランク軸

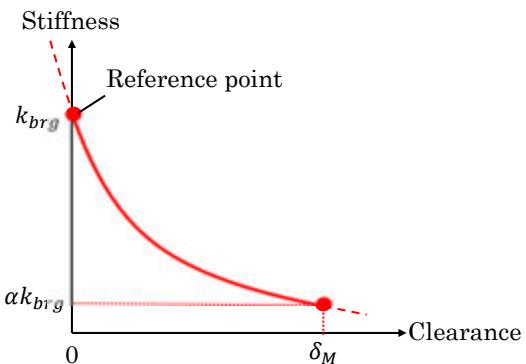


図14 与えた軸受剛性のイメージ

ピストンについては質量のみを考慮し、燃焼圧による力およびピストン質量による慣性力を弾性体であるコンロッド小端部に付与した。燃焼圧には、実稼働ひずみ測定試験時の実測値を使用した。構築した動解析モデルを用いて、エンジンサイクルで5サイクル分のシミュレーションを実施し、収束した最終2サイクル分のデータを用いて実測結果との比較を行った。

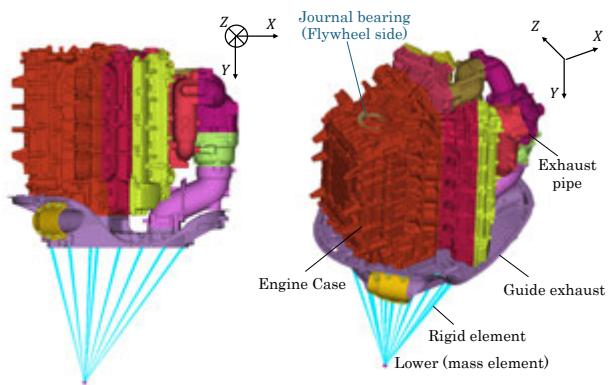


図13 FEMでモデル化したケース部分

主軸受、大端軸受、スラスト軸受には非線形ばね要素を用いて剛性特性を付与し、クランク軸—エンジンケース間およびコンロッド大端—クランク軸間を接続した。これらのばね要素の剛性値は、図14のイメージ図に示すような非線形関数により定義した。最大燃焼圧荷重および想定される最大面圧から得られる荷重を、最大すきまの大きさ  $\delta_M$  で除することで剛性値  $k_{brg}$  を算出した。さらに、接触二面間（軸面と軸受面）のすきまが0のときに剛性値  $k_{brg}$  となるように基準点を設定した。接触二面間のすきまが最大すきま  $\delta_M$  となった場合の剛性値は、 $k_{brg}$  の定数  $\alpha$  ( $0 < \alpha < 1$ ) 倍として与え、この点と基準点を通る非線形関数（すきまが小さいほど剛性が高くなる関数）を剛性特性とした。コンロッド小端—シリンダ間は線形ばね要素で接続し、その剛性値も非線形ばね要素と同様に、想定最大面圧および半径クリアランスに基づいて決定した。

## 4 実測結果と予測結果の比較

合計5つある主軸受のうち、1つの軸受に関する比較検証結果を示す。他の軸受についても同様の検証を行い、同様の結果が得られることを確認している。実測および解析により得られた軸受荷重は、荷重軌跡  $\mathbf{F}(t) = [F_Z(t) \ F_X(t)]^T$  のベクトルを各時刻毎にプロットしたものとして比較した。代表例として、定常回転2,100rpm、3,900rpm、5,500rpmにおける、1サイクル分の荷重軌跡の比較結果を図15に示す。図15においては、すべてのグラフにおいて荷重の大きさのスケールを統一している。いずれのエンジン回転数においても、解析結果は実測結果と同じ第3象限方向に荷重のピークを示し、その荷重の大きさの差は実測結果に対して  $\pm 1.5$  [dB] 以内となった。このことから解析結果は実測と良好な相関を示していることが確認できた。

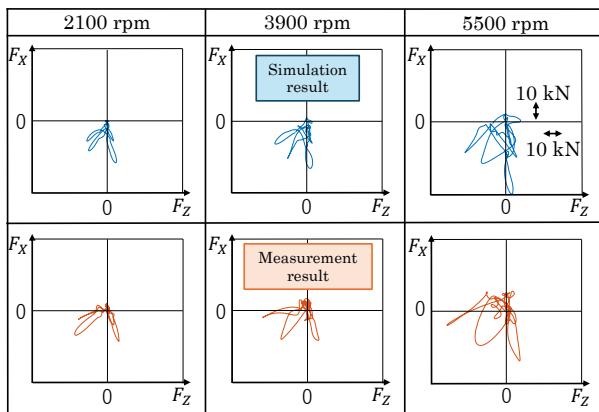


図15 荷重軌跡の比較結果

#### 4-2. 周波数領域での比較

4-1節と同一の軸受に関して、回転スイープ条件下で測定した実測荷重および解析荷重について、次数分析(オーダートラッキング)を実施した。解析荷重は、100rpm 每の定常回転条件にて解析を行い、評価次数に対応する周波数の荷重振幅を抽出して次数分析波形とした。実測および解析の双方において、エンジンケースの加速度応答振幅が大きいエンジン回転1次、4次を比較対象とした。各並進方向の荷重  $F_x, F_z$  について、エンジン回転1次および4次の次数分析(縦軸: 評価次数における荷重の大きさ、横軸: エンジン回転数)の比較結果を図16に示す。評価次数における荷重ピークを示すエンジン回転数について、解析結果の実測結果に対する差は±8[%]以内となった。また荷重ピーク値について、解析結果の実測結果に対する差は±3.5[dB]以内となり、解析結果は実測結果と良好な相関が示していることが確認できた。

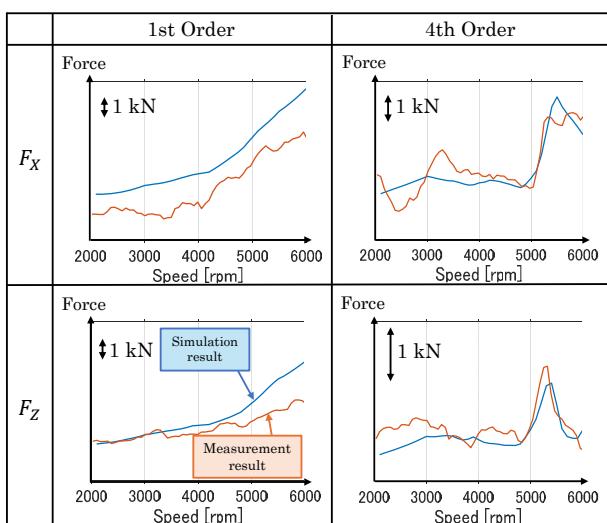


図16 次数分析波形の比較結果

## 5 おわりに

本開発では、船外機エンジンにおいてクランク軸挙動に起因して発生するクランク打音に着目し、その予測の第1ステップとしてジャーナル荷重を予測する解析モデルを構築した。解析モデルの精度検証に際しては、ひずみデータを用いたジャーナル荷重の実測を行い、解析結果との比較を実施した。比較方法としては、荷重軌跡による比較および次数分析による比較を採用した。その結果、実測値と予測解析結果との間に良好な相関が得られていることを確認した。これにより、エンジンケースにおける振動伝達およびケース表面からの放射音予測手法に適用可能な高精度な入力荷重が、本開発により得られるようになった。今後、放射音予測手法の検証を通じて、クランク打音の予測が可能となる。

#### ■参考文献

- [1] 鳥居建史, “重回帰分析を用いた時系列エンジン放射音寄与分離法”, 自動車技術会論文集, Vol. 44, No. 4, pp. 1051-1057(2013).
- [2] 山岸誠弥, 小澤恒, 山下健一, “改良指針検討のための実稼働TPAによるエンジン騒音寄与分析”, 自動車技術論文集, Vol. 53, No. 6, pp. 1227-1232(2022).
- [3] 太田和秀, 入江良彦, 本田巖, 水早純, 福沢順博, 下河内寛子, “回転クランク軸とエンジンブロックの連成振動応答及び放射騒音の予測”, 日本機械学会論文集C編, Vol. 68, No. 671, pp. 1966-1971(2002).
- [4] 大西有希, 山隅允裕, 天谷賢治, “任意形状ホイールに対するひずみゲージを用いた6分力逆解析”, 日本機械学会論文集A編, Vol. 75, No. 759, pp. 1460-1467(2009).
- [5] 荒井政大, 古田剛, 石川智哉, 岡崎羊平, “ひずみ計測に基づく圧力分布測定用ロードセルに関する研究(第1報 測定理論と数値シミュレーション)”, 日本機械学会論文集A編, Vol. 72, No. 713, pp. 69-76(2006).
- [6] 熊谷常雄, “ひずみゲージとブリッジ回路”, 計測と制御, Vol. 45, No. 4, pp. 323-328(2006).
- [7] 山浦義郎, “ひずみ計測の基礎と応用”, 精密工学会誌, Vol. 73, No. 7, pp. 772-775(2007).
- [8] 林洋次, “滑り軸受(1)”, ターボ機械, Vol. 10, No. 6, pp. 366-372(1982).
- [9] 村木正芳, 図解 トライボロジー 摩擦の科学と潤滑技術, 初版(2007), p. 177, 日刊工業新聞社.
- [10] “AVL EXCITE™ | AVL”: <https://www.avl.com/ja-jp/simulation-solutions/software-offering/simulation-tools-a-z/avl-excite-m> (アクセス日: 2025/5/14)

■著者



井上 祐人

Yuto Inoue

技術・研究本部

デジタル開発統括部

MBSE 推進部



高比良 拓人

Takuto Takahira

技術・研究本部

技術開発統括部

エネルギー・マネジメント研究部 MBSE 推進部



藤田 英之

Hideyuki Fujita

技術・研究本部

デジタル開発統括部

エネルギー・マネジメント研究部 MBSE 推進部



小林 光司

Koji Kobayashi

技術・研究本部

技術開発統括部

人間研究部



近藤 拓

Taku Kondoh

技術・研究本部

技術開発統括部

プロジェクト推進部



高橋 徹

Toru Takahashi

マリン事業本部

開発統括部

実験部

当論文は、公益社団法人 土木学会主催 第10回 FRP 複合構造・橋梁に関するシンポジウム(2024年11月)にて発表されたものです。同会の許可を得て転載したものであり、本論文の著作権は公益社団法人 土木学会に属します。無断複製・転載を禁じます。

## 要旨

ヤマハ発動機株式会社は1974年に国内初のオール FRP 製プールを販売し、ニーズや社会変化に応じて多様なプールの開発、事業拡大のための挑戦を続けてきた。現在までに累計6,500件以上の FRP 製プールを納入し、プールの主要材料として FRP が認知されるまでに至った。

ヤマハ FRP プールはプール槽をいくつかのユニットに分割して工場にて製造し、施工現場にてひとつのプール槽にするユニット構造を採用している。FRP の耐腐食性に優れている点や軽量である点がプールの材料として有利である一方、塩素との反応や温度収縮など FRP により生じる問題もあり、構造や工法、原材料に対する改良を行ってきた。本論文ではヤマハ FRP プールの構造的な特徴から施工技術、FRP 製プールの利点や課題を、50年の実績と共に報告する。

## Abstract

Yamaha Motor Co., Ltd. sold Japan's first all-FRP pool in 1974, and has continued to develop a variety of pools in response to needs and social changes, and to take on the challenge of expanding its business. Up to the present, Yamaha Motor has delivered a cumulative total of more than 6,500 FRP pools, and FRP has come to be recognized as a major material for swimming pools.

Yamaha FRP pools adopt a unit structure in which the pool tank is divided into several units, manufactured at the factory, and assembled into a single pool tank at the construction site. While FRP's superior corrosion resistance and light weight make it an advantageous material for pools, there are also problems caused by FRP, such as chlorine reaction and temperature shrinkage, and improvements have been made to the structure, construction method, and materials. This paper reports on the structural characteristics of Yamaha FRP pools, construction techniques, and the advantages and challenges of FRP pools, along with 50 years of achievements.

## 1

## はじめに

ヤマハ発動機株式会社(以下、ヤマハ)は「感動創造企業」を企業目的に、世界の人々に新たな感動と豊かな生活を提供することを目指している。1955年の創業以来、技術と感性で感動を生み出す多彩な製品を世界中へ出荷しており、現在ではランドモビリティ事業やマリン事業、ロボティクス事業からファイナンス事業まで多軸に事業を展開している。そのようなヤマハは、当時ボートの開発で培った FRP (Fiber Reinforced Plastics) の技術を活用し、1974年に国内初となるオール FRP 製プール「ファミリープール」を発売、プールメーカーとしての歴史をスタートさせた。

輸送機器メーカーとしての考え方からプールを「建築物」としてではなく「工業製品」としてとらえ、ユニット式の FRP 製プール

の開発を進めてきた。教育を目的とした幼児用プール(図1(a))や学校向けのスクールプール(図1(b))の開発から、レジャーを目的としたウォーターパーク向けの大型流水プールやウォータースライダー(図1(c))の開発、スポーツ振興を目的とした競泳用公認50m プール(図1(d))の開発、健康増進を目的としたアクウォーカフロア(水中歩行に適したブロックパターン付きフロア)(図1(e))やフラットプールの開発、よりプライベートな利用を目的としたホテル・個人向けのプール(図1(f))の開発など、ニーズや社会変化に対応し、多様なプールの開発、事業拡大のための挑戦を続けてきた。

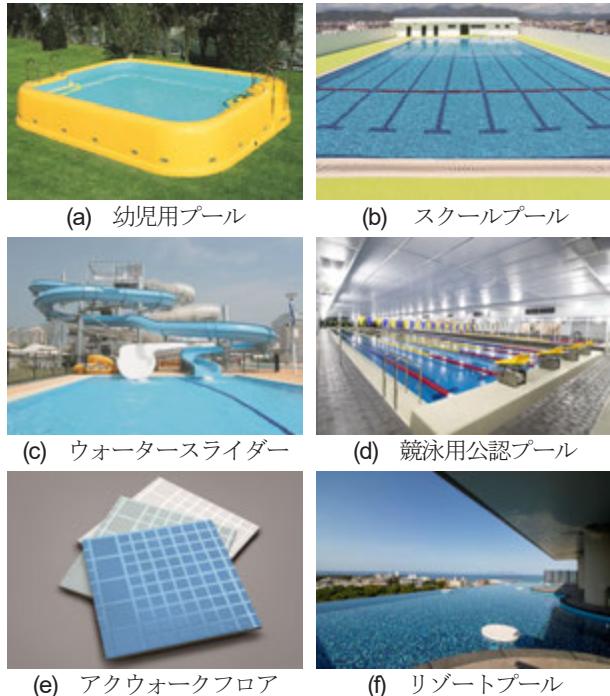


図1 ヤマハ FRP プールの商品事例

プール(幼児用プールを除く)の納入件数は累計で6,500件を超える(図2)、2020年の新規公共スクールプールにおける市場シェア(当社調べ)は、プール全体で約55%、FRP 製プールに限れば約95%に至った。日本国内にとどまらず海外市場への展開を目指し、近年では韓国を中心としたアジア諸国への納入実績も得られた。しかしながら、プール事業50周年を迎えた2024年の3月末日に営業活動を停止し、2027年をもってヤマハはプール事業から撤退することを発表した。ポートフォリオ経営を進める中で、既存事業や新たな成長事業へ経営資源を集中させるべく事業撤退を決定した。

プールの分野において、FRP 製プールはそのシェアを拡大し、広く利用されるようになった。ヤマハがプール事業に参入した当時は RC 製プールが普及していたが、現在では RC、ステンレス鋼に並び、FRP がプールの主要材料として認知されるようになった。そこには FRP の特性とプール槽との親和性の高さに加え、それらを下支えする技術や工夫が関係していると考える。本論は、FRP の新たな活用先を見つける手がかりとなることを期待して、ヤマハ FRP プールの構造的な特徴から施工技術、FRP 製プールの利点や課題を、50年の実績と共に報告するものである。

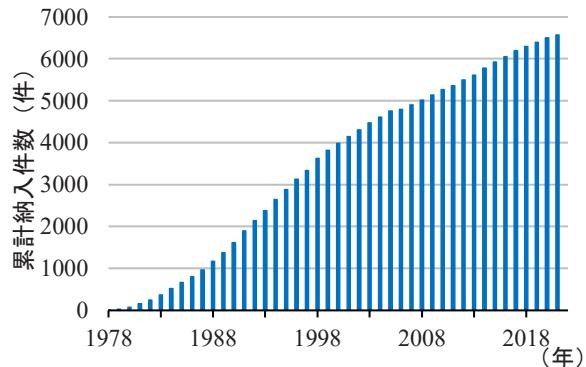


図2 プールの納入件数の推移(累計)

## 2 ヤマハ FRP プールの概要

### 2-1. 全体構造

学校向け25mの標準的なスクールプールの全体図を図3に示す。ヤマハ FRP プールは、RC 布基礎の内側にサンドクッシュン(川砂)を敷き詰め、その上に FRP 製のプール槽を据える構造となっている。一般的な RC 製プールとは異なり、プール槽外周のプールサイドとプール槽は別々の構造となっている。図4 は FRP プールの断面を示している。プール槽はステー・ベース・アンダル(山形鋼)並びに金属系あと施工アンカーを介して RC 布基礎に固定されている。

プール槽は図5に示すように、独自の規格により分割されたユニット構造となっており、サイドユニット、フロアユニット、ガータユニット等から構成される。それぞれのユニットは工場にて製作され、施工現場へと搬入される。その後、ユニット同士を接合することでひとつの25m プール槽が完成する。プール槽を分割されたユニットとして設計し、それらを工場で生産することにより、季節や天候による品質への影響が少ないという長所がある。また、施工現場においては、プールの据付けに必要な期間が一般的には2週間程度であり、他材料と比較して現場工期を短縮できるという長所もある。

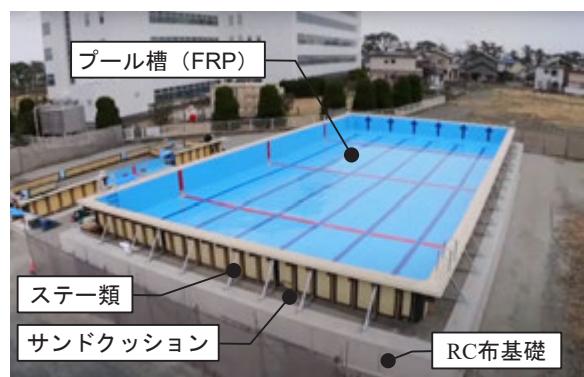


図3 スクールプールの全体図

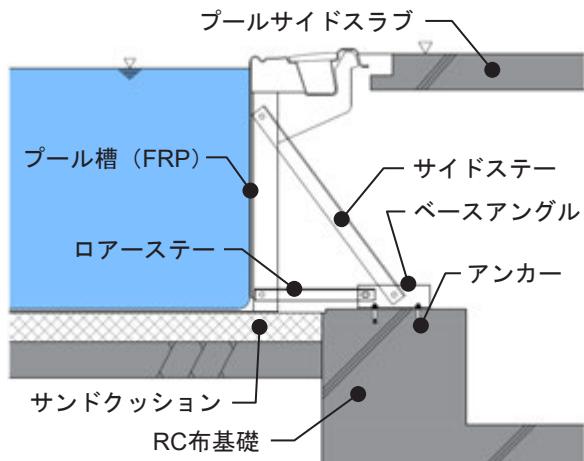


図4 スクールプールの断面図

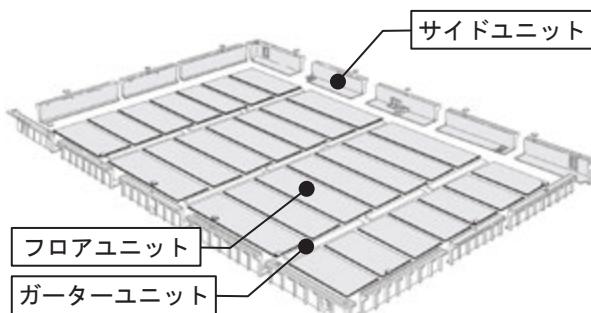


図5 分割されたユニット構造

## 2-2. 各ユニットの構造

サイドユニットはプール槽の側壁部分を構成するユニットであり、ハンドレイアップ法によって、オーバーフロー溝から、側壁、水底の一部までが一体的な構造となっている。ユニットの接合部となる両端部には、接合用のフランジが一体成形されており、一定の間隔で合板や合成木材等による補強用のリブが取り付けられている。フロアユニットはプール槽の水底部分を構成するユニットであり、ハンドレイアップ法による成形の他、インフュージョン法(真空樹脂含浸法)によっても成形されている。

FRP 断面は、GFRP 層(ガラス繊維 + 不飽和ポリエスチル樹脂)とゲルコート層(不飽和ポリエスチル樹脂)によって構成されている。水圧の作用する側壁部、並びに水底部においては、軽量かつ強度・剛性を高めるため硬質発泡材を積層したサンドイッチ構造となっている。

ゲルコート層は、GFRP 層を紫外線等から保護する役割がある。ユニット成形時、成形型にゲルコートを吹き付けた後に樹脂とガラス繊維を積層することにより、ゲルコートと GFRP を一体成形している。また、ゲルコートには GFRP の保護層としての

役割の他、着色の役割がある。一般的な塗料と比較して経年劣化による退色や剥がれが発生しにくく、美しい外観を長期間保つことが可能であると共に、塗替えが不要であるという長所がある。ヤマハ FRP プールの多くに使用されている色「パロスブルー」は、曇天でも栄える水の美しさを目指すとともに、塩素の漂白作用による退色に強い色として開発された。

## 2-3. ユニットの接合方法

ユニットによる分割方式の採用において、ユニット同士の接合方法は非常に重要な課題である。プール槽は漏水が許されない商材であり、ユニット同士の接合部にはその水圧に耐える十分な強度と、高い水密性が要求される。ヤマハ FRP プールでは、ボルト締めによるボルトアップ接続と、現地 FRP 積層による接続(以下、FRP 目貼り接合)の2つの接合方法によりそれらの要求に応えている。

サイドユニット同士の接続には、ボルトアップ接合を採用している。図6にボルトアップ接合の様子を示す。サイドユニットのフランジ面にスポンジパッキンを貼り合わせた後、隣接するサイドユニットのフランジ面と重ね合わせ、ボルトによって締め付けを行う。その後、水側の縫目に化粧コーティング材を塗布して仕上げを行う。ボルト締めによってパッキンが圧縮され、フランジ同士の隙間を埋めることで止水性を確保している。接合部であるフランジ面はユニット成形時の離型面であることから、表面が平滑であり、パッキンとの高い密着性が期待できる。

水底部およびガーターユニットの接続には、FRP 目貼り接合を採用している。図7に FRP 目貼り接合の様子を、図8に FRP 目貼り接合部のモデルを示す。接合部となるユニットの端部はユニット表面より1段(上型)または2段(下型)凹んだ形状となっており、下型の上に接続されるユニットの上型が重なるようにユニットを配置する。ユニット同士をアルミリベットで仮止めした後、接合部の溝面をサンディングし、ガラスマットの配置、樹脂の流し込みと脱泡作業により FRP 積層を行う。硬化後、カラーペースト樹脂を塗布することにより仕上げを行う。FRP 目貼り接合により、ユニット同士が一体的な構造となるため、十分な接合強度により荷重を確実に伝達するとともに、高い水密性を発揮する。

ユニットの分割位置にも工夫がある。サイドユニットは側壁から水底の一部までをひとつのユニットとして成形しており、応力が集中する入隅部を避けた位置にてフロアユニットと接合することができる。接合位置が平面となるため、現場における FRP 目貼り接合の作業性も向上している。

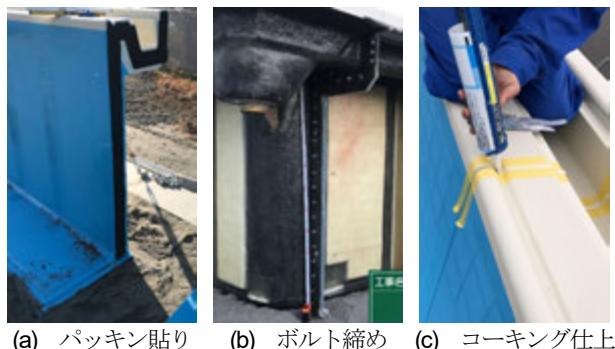


図6 ボルトアップ接合の様子



図7 FRP 目貼り接合の様子

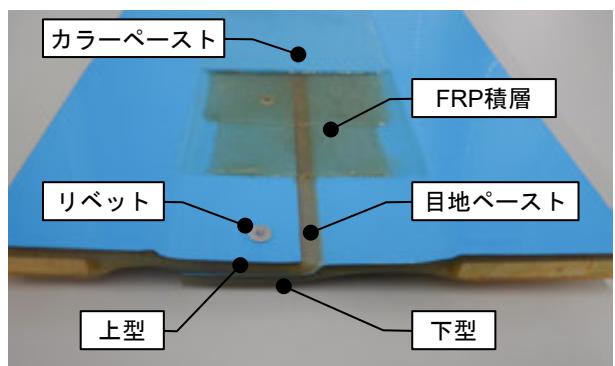


図8 FRP 目貼り接合部のモデル

#### 2-4. 循環方式

プール使用時においては、利用者の安全・衛生が最優先される。プールの運用において必要なのが、ろ過循環システムである。その循環システムによる吸い込まれ事故等により、利用者の安全が脅かされないよう、ヤマハのスクールプールは図9に示すような、ガーターを利用した循環方式を標準採用している。プール水底部に設けられた複数の水底吸込金具および目皿からプール水が吸い込まれ、ガーター、金具、配管を経由し、ろ過機へと運ばれる。ろ過されたきれいな水が、プール側壁に分散して取り付けられた吐出金具よりプール槽へと供給される。主に旧式のプールにおいては、プール槽の最深部に吸込(排水)

部が1箇所のみの事例が多く、それらと比較してヤマハのスクールプールは吸込口を分散させることで、プール槽全体の水の循環効率を向上させるとともに、吸水圧を小さくして吸い込まれ事故の防止を図っている。また図10に示すように、吸込口は目皿のボルト固定に加えて、配管の取付口に吸込防止金具を溶接した、二重安全構造を採用している。文部科学省と国土交通省が策定した「プールの安全標準指針(平成19年3月)」においても、二重構造の安全対策を規定し、複数の吸込口設置を改修事例として記載している。ヤマハにおいてはスクールプール販売開始当初からこの方式を採用していた。

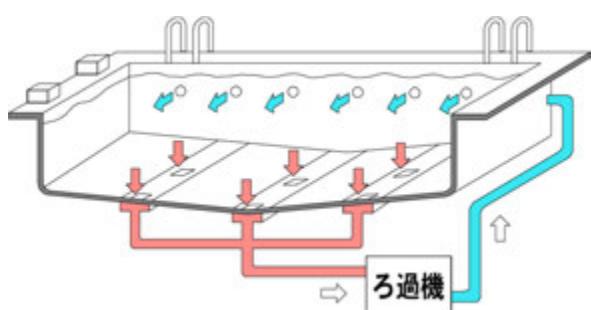


図9 ヤマハスクールプールの循環方式

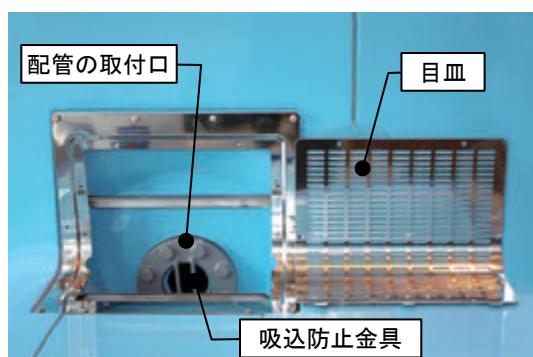


図10 二重安全構造

### 3 ヤマハ FRP プールの構造評価

プールの構造設計にあたり、プール水の水圧に対する各部材の強度評価を行う必要がある。水圧には、側壁に作用する水平方向の水圧と、水底に作用する鉛直方向の水圧(プール水の自重)がある。

#### 3-1. 水平方向の水圧に対する考え方

水平方向の水圧として、長期は静水圧、短期は静水圧に加えて地震による変動水圧を考慮し、プール槽、並びにプール槽を支持するステーや基礎等の強度評価を行う。

サイドユニットの構造検討において、側壁パネルには水圧により曲げ応力が発生する。補強用リブによって側壁パネルが支持されると考え、補強用リブの間隔を狭めることで、側壁パネルに発生する最大曲げ応力度が FRP の許容応力度以下となるよう設計を行う。実際の設計業務においては、プールの水深に応じた補強用リブの間隔の最大値を定めており、その数値以下となるよう補強用リブの配置を検討している。

ステーの部材検討、布基礎の強度検討において、ステーが負担する荷重を求める必要がある。図11に示すように、プール槽を一体化的な構造物ととらえることで、プール側壁に作用する水圧による荷重の一部はフロアユニットがプール両サイドに引っ張られることで負担し、ステーは残りの荷重を負担すると考えられる。実際の構造計算においては、水圧による全荷重の2/3をフロアユニットが負担し、1/3をステーが負担するとして、ステーの構成部材検討、布基礎の強度検討を行っている。合わせてフロアユニットに発生する引張応力が、FRP 目貼り接合部の許容応力以下であることを確認する。

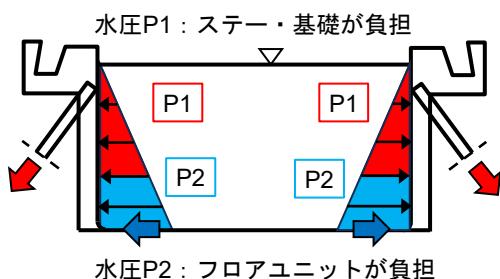


図11 水平方向の水圧に対する考え方

### 3-2. 鉛直方向の水圧に対する考え方

鉛直方向の水圧は、図12に示すように、フロアユニット並びにサンドクッション層を介して地盤が支持すると考え、鉛直方向の水圧に対するフロアユニットの強度照査は行っていない。しかしながら、フロアユニットとサンドクッション層に隙間が生じると、フロアユニットのみで鉛直方向の水圧を負担することになるため、プールユニット据付時にはサンドクッション層の高さに注意するとともに、砂の流出を防止する必要がある。一方、サンドクッション層が浸水した場合、浮力によりフロアユニット全体に大きな曲げ応力が作用し、フロアユニットの破断が発生する可能性がある。そのため、スラブや耐圧盤を支持層とする物件では、布基礎に排水用孔を設けるなどして水が溜まることを防止する必要がある。また、地下水位がプール底より500mm以上低いことを設置の条件としている。ただし、降雨量や地形にも十分配慮する必要がある。

## 4 FRP 製プールの材料的特徴

現在普及しているプールの主要材料としては FRP の他、RC、ステンレス鋼等がある。RC 製プール、SUS 製プールとの比較を含めつつ、各観点における FRP 製プールの利点、課題とその対策について示す。

### 4-1. 耐腐食性

常時水に接するプールにおいて水による腐食への対策が必須であり、腐食しにくい FRP は プール槽という用途との親和性が高い材料と言える。特に、塩害を受けやすい海岸沿いに設置するプールにおいては、耐腐食性の観点で FRP 製プールが有利となる。

一方、プールの消毒剤に含まれる塩素によって FRP 表面が変色する現象が発生する可能性がある。そこで含有物等を調整したプール専用の樹脂やゲルコートを開発し、変色の発生を抑えることに成功した。

### 4-2. 軽量

FRP は比強度に優れた材料であり、RC 製プールや SUS 製プールに比べて、プールの総重量を小さくすることができる。軽量であることは工作物、建築物として構造設計上有利であり、地上階以外にプールを計画する際や既設の建物にプールを設置する際に、重量の観点から FRP 製プールが採用されることがある。また、施工現場におけるプールユニットの据付けの際に大型の重機を必要とせず、人力での運搬も可能であるため作業性に優れている。

### 4-3. 設計自由度

標準的な学校用プールは矩形の平面形状をしているが、プールの幅や水深、水底の勾配など、物件に応じて要求される仕様は異なる。FRP は形状の自由度が高い材料である一方、成形するにあたり型の準備が必要となる。オーバーフロー溝から水底の一部までをひとつのサイドユニットとして成形するヤマハ FRP プールにおいて、プールの幅や水深の違いに応じて都度、成形用の型を製作していくには製造コストや型の維持費が増加してしまう。そこで、標準的な水深・プールの幅に合わせて作られた成形型に対し、型を積み上げて水深や勾配を変える積型作業や、仕切り板を立ててユニットの長さを変更する仕切り作業など、型改造の工程を行うことで、型のコストを抑えつつ様々な要求仕様への対応を可能としている。

スクールプール用の成形型の他、半径の異なる曲線用のサイドユニットの成形型を多数所持しており、これらの型から成形

したユニットを組み合わせることで、流水プール(図13(a))のように複雑な曲線形状をしたプールの製作にも対応可能である。また、水中階段(図13(b))のようにRC やステンレス鋼では設計や加工が複雑になる部分に対しても、FRP ではひとつの成形型から繰り返し成形することが可能であり、FRP の形状の自由度の高さが発揮される。

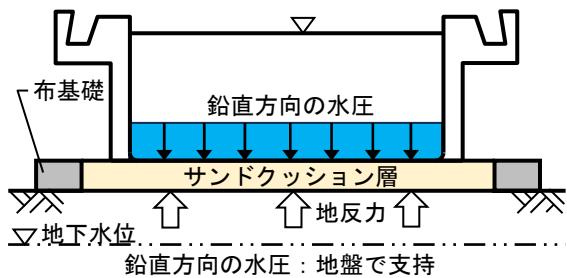


図12 鉛直方向の水圧に対する考え方



図13 形状の自由度を生かした事例

#### 4-4. メンテナンス性

プールのメンテナンスとしては、劣化部の補修の他、劣化した塗装の塗り替え作業が含まれている。プール全体を塗装するRC 製プールやSUS 製プールでは、定期的な塗装の塗り替えが必要となる。一方、FRP 製プールは顔料を混ぜたゲルコート層によって着色を行っているため再塗装が不要であり、長期にわたり美観を保つことができる。距離ラインおよびレンラインのみアクリルウレタン樹脂塗料により現地塗装を行っているため定期的な塗り替えが必要となるが、他の材料と比較して小規模の範囲の塗り替え作業で済む。また、あらかじめゲルコート層によってユニットに距離ラインおよびレンラインの着色をすることで、塗装のメンテナンスが不要な仕様を開発した。

#### 4-5. 温度収縮

FRP 材は温度収縮の大きい材料であり、一般的な鉄やコンクリートの線膨張係数が約 $1.0 \times 10^{-5}/^{\circ}\text{C}$ であることに対し、ヤマハ FRP プールに使用されている FRP の線膨張係数は約 $2.5 \times 10^{-5}/^{\circ}\text{C}$ と高い値を示す。特にプールのように規模の大

きい構造物においては FRP の温度収縮は無視することができず、プールの水を抜いた際には温度差によってフロア部が収縮・変形し、大きく波打つことがあるため、維持管理の方法として非使用時においても満水保管を推奨している。-0mm ~ +10mm の高い寸法精度が求められる競泳用の公認50m プールにおいては、工場でのユニット製作時期と現地施工時期の気温差から FRP の伸縮量を推定し、ユニットの長さを調整することで要求精度を満たすことに成功した。また、フロアユニットにおいては、従来のハンドレイアップ法から、インフュージョン法に切り替えることにより、ガラス繊維含有率を高め、温度収縮の抑制につながった。

#### 4-6. 耐久性

RC はコンクリートの中性化の進行による鉄筋の発錆等から耐用年数を推定することが可能な一方、FRP 製プールの耐用年数の考え方については、現状まだ明確に定められていない。FRP は紫外線等の外的要因によって強度低下する材料であり、ヤマハ FRP プールにおいても長年使用したプールから採取された FRP 板の強度は、成形直後の FRP 板の強度と比較して低下していることが社内試験により分かっている。FRP の強度低下の進行率についてはまだデータが不足しており、加えて FRP 製プールの構造物としての限界をどのように判断するのかという基準も明確に定められていないのが現状である。

ヤマハ FRP プールは30年の使用を目的に開発されており、FRP の設計強度については、実験により得られた材料強度に対し、ばらつき等による安全率とは別に、材料の強度低下を考慮した低減係数を乗じて定めている。なお、強度低下の低減係数は実験等により求めた値ではなく、FRP 水槽の構造計算法<sup>[1]</sup>等を参考に仮定した値である。実績として、1978年に納入されたヤマハ初の学校用スクールプールは、46年経過した2024年現在でも使用されている。

## 5 おわりに

ヤマハ FRP プールの構造的な特徴から施工技術、FRP 製プールの利点や課題を、50年の実績と共に報告した。ボートの開発で培った FRP 技術の活用先としてプールに注目し、国内初となるオール FRP 製プールを開発した。その後もニーズや社会変化に応じて多様な FRP プールの開発を続け、現在では FRP がプールの主要材料として認知されるまでに至った。FRP とプールの親和性の高さに加え、プールとしての要求性能や課題に対して、構造や工法、原材料の開発や改良を続けてきた結果であると考える。

## ■参考文献

[1] 強化プラスチック協会:FRP水槽構造設計計算法(1996年版), p. 20, 1997.

## ■著者



松井 健良  
Takayoshi Matsui  
マリン事業本部  
国内事業推進部  
プール管理部



内山 仁平  
Jimpei Uchiyama  
マリン事業本部  
国内事業推進部  
プール管理部



菊地 秀和  
Hidekazu Kikuchi  
マリン事業本部  
国内事業推進部  
プール管理部



喜田 晃爾  
Koji Kida  
マリン事業本部  
国内事業推進部  
プール管理部



糸永 汐里  
Shiori Itonaga  
マリン事業本部  
国内事業推進部  
プール管理部



萩野谷 健吾  
Kengo Haginoya  
マリン事業本部  
国内事業推進部  
プール管理部

# Development of CO2 emission reduction technology for sport motorcycles

Naoki Makita Masaki Torigoshi Toshihiko Takahashi Hiroki Takase

当論文は、JSAC 2024-096/SAE 2024-32-0096として、SETC2024 (Small Powertrains and Energy Systems Technology Conference)にて発表されたものです。

Reprinted with permission Copyright © 2024 SAE Japan and Copyright © 2024 SAE INTERNATIONAL  
(Further use or distribution is not permitted without permission from SAE.)

## 要旨

気候変動に対する世界的な関心が高まる中、二輪車(MC)においてもカーボンニュートラル(CN)を実現することが課題であり、CNを実現するためには様々なアプローチが必要である。内燃機関(ICE)を用いたパワートレインでは、e-fuel やバイオ燃料などのCN燃料を採用することでCNを達成できるが、コストや供給面を考慮すると、ICEのCO2削減技術の開発が重要である。MCは四輪自動車と比較して、高出力、軽量、コンパクト、長距離走行が要求されるが、採用できるCO2削減技術は動力性能とCO2削減のトレードオフになりがちであり、両方の要求を高いレベルで実現することが課題である。そこで、特に高い動力性能が要求されるミドルクラスのスポーツMCに着目し、CO2削減技術を開発することにした。技術開発目標としては、スポーツMCに求められる動力性能を維持しつつ、CO2排出量を worldwide-harmonized motorcycle test cycle (WMTC) クラス3-2で65g/kmとした。目標達成に必要な技術の組み合わせをシミュレーションし、技術実証のコンセプトを選定した。その結果、電動アシストターボ(E-Turbo)付ダウンサイ징コンセプトを選定し、試作エンジンを用いてCO2排出量と動力性能をダイナモで検証した。本稿では、これらのシミュレーションに基づく技術選定の考え方と、試作エンジンを用いた実証結果について述べる。

## Abstract

With growing global concern about climate change, the challenge is to achieve carbon neutrality (CN) in motorcycles (MCs) as well, and various approaches are needed to achieve CN. For powertrains using internal combustion engines (ICEs), CN can be achieved by adopting CN fuels such as e-fuel and biofuel, but considering cost and supply, it is important to develop CO2 reduction technologies for ICEs. Compared with 4-wheel vehicles, MCs are required to be powerful, lightweight, compact and capable of travelling long distances, the CO2 reduction technologies that can be adopted tend to be a trade-off between dynamic performance and CO2 reduction, and a challenge is to achieve a high level of both requirements. We decided to focus on middle-class sports MCs, which require particularly high dynamic performance, and to develop CO2 reduction technologies. As a technology development target, CO2 emissions were set at 65 g/km in the worldwide-harmonized motorcycle test cycle (WMTC) class 3-2, while maintaining the dynamic performance required for sports MCs. The combination of technologies required to achieve the target was simulated and a concept was selected for technology demonstration. As a result, the downsizing concept with electrically assisted turbocharger (E-Turbo) was selected and CO2 emissions and dynamic performance were verified on a dynamometer using a prototype engine. This paper describes our approach to selecting the technologies based on these simulations and the demonstration results using the prototype engine.

## 1

## INTRODUCTION

In recent years, global concern about climate change has increased and efforts are being made in various fields to reduce CO2 emissions to achieve CN. In the transport

equipment industry, CN compliance is also a major challenge. For vehicles using ICEs, CN can be achieved by adopting CN fuels such as e-fuel and biofuel, but there are also issues in terms of cost and supply<sup>[1][2]</sup>, so not only replacing fuel types but also improving fuel

consumption is important as a value provided to customers. To maintain and improve the value provided to customers while meeting social demands for reducing environmental impact, it is necessary to develop ICEs technologies to reduce CO2 emissions. Various CO2 reduction technologies are also being researched and developed for MCs, with the use of electric components<sup>[3][4]</sup> and the adoption of turbochargers to reduce CO2 emissions<sup>[5]</sup>. On the other hand, sports MCs with medium to large displacement engines tend to require dynamic performance such as high power, light weight and compactness. Figure 1 shows the relationship between power weight ratio (PWR) and CO2 emissions for MCs, four-wheel vehicles with ICEs and four-wheel hybrid electric vehicles (HEVs), showing that MCs are a product group with a smaller PWR than four-wheel vehicles. However, CO2 reduction technologies tend to require various additional devices for existing conventional MCs, which do not directly lead to improved power performance or impair this. We have recognized that the compatibility between low CO2 emissions and competitive dynamic performance is a major issue in sports MCs and have conducted technical verification of CO2 reduction technologies suitable for sports MCs, which can achieve a top-class CO2 emission of 65 g/km for four-wheel vehicles.

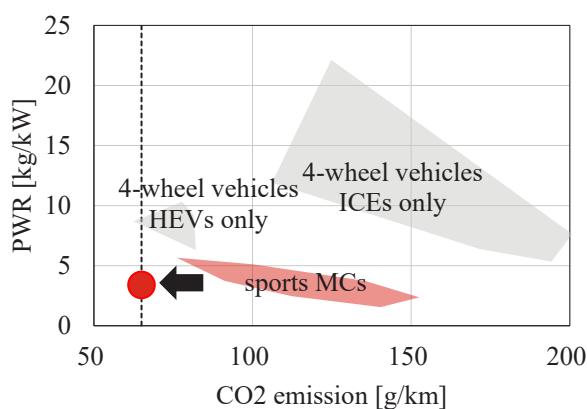


Fig. 1 PWR as a function of CO2 emission

## 2

## CONCEPT RESEARCH

To narrow down the technologies required to achieve a CO2 emission to 65 g/km, a driving simulation was used to calculate the CO2 emission level. The driving simulation used was a one-dimensional (1D) simulation by Gamma Technologies' GT-SUITE. An overview of the calculation method is shown in figure 2. The main input parameters are driving pattern, gear position schedule, gear ratio, vehicle weight, displacement, brake specific fuel consumption (BSFC) map, friction mean effective pressure (FMEP) map, cooling loss map, FMEP increase factor at cold-start, and fuel injection increase factor at cold-start. The engine operating points are calculated from the driving pattern and gear position schedule to obtain fuel consumption, which is converted into CO2 emissions during WMTC mode driving. The oil and water temperature are calculated from the cooling loss map to determine the amount of heat received and is reflected in the increase in oil and water temperature. The FMEP increase factor at cold-start and fuel injection increase factor at cold-start are used to calculate the fuel consumption deterioration at low engine temperatures.

For the input parameters, a virtual BSFC map created with reference to data from previously developed engines and high thermal efficiency engines for four-wheel vehicles<sup>[6][7][8]</sup> was used. The specification table for each calculation model is shown in table 1 and the BSFC maps created are shown in figure 3.

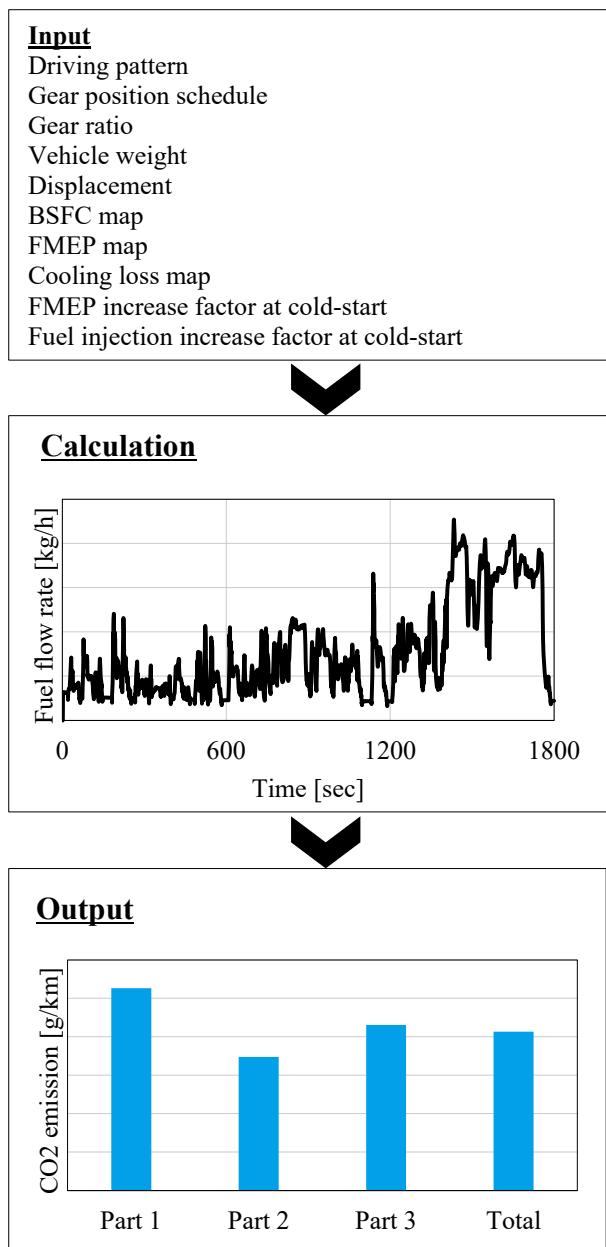


Fig. 2 Outline of calculation logic

Table 1 Specifications of calculated engine model

	Base engine	Engine A	Engine B	Engine C
Displacement	689 cc	689 cc	412 cc	689 cc
CR	11.5	13.6	13.6	13.6
High tumble port		✓	✓	✓
VVT		✓	✓	✓
Cooled EGR		✓	✓	✓
Lean burn		✓		
DI-Turbo			✓	
Cylinder deactivation				✓

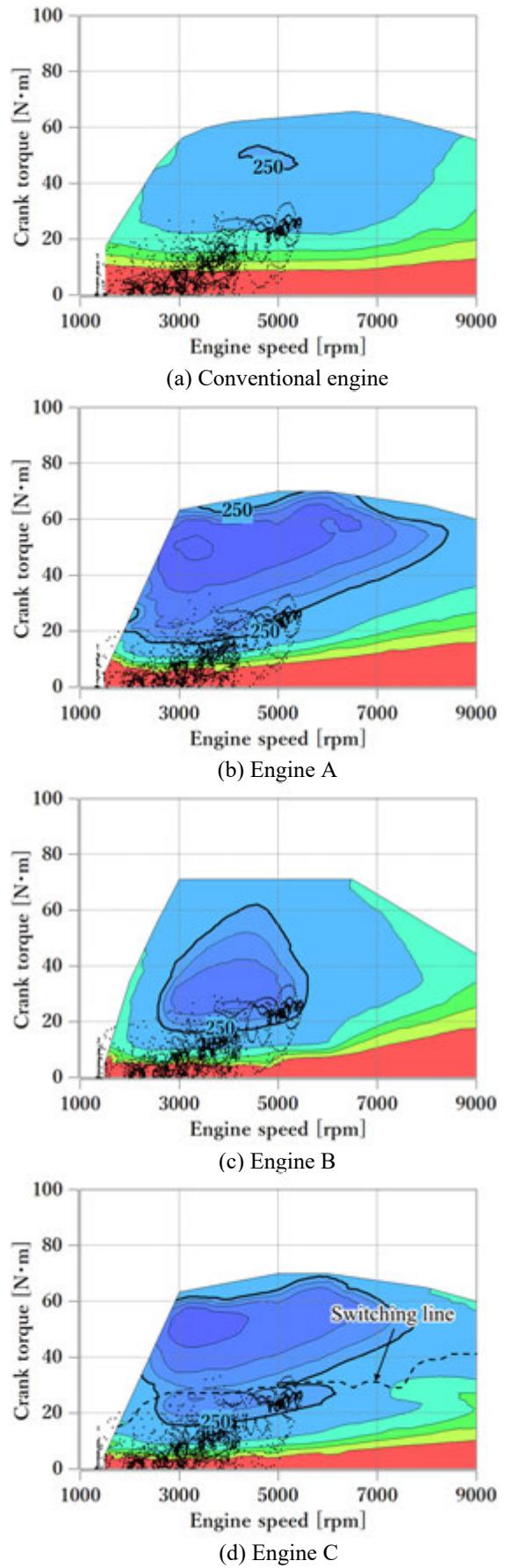


Fig. 3 BSFC map of calculation model

The Base engine is a middle class general sports MC naturally aspirated (NA) engine; the BSFC best point indicates 250 g/kWh around an engine speed 5000 rpm and a crank torque 50 N·m. The engine operating points during WMTC mode driving uses engine speeds up to 5500 rpm and crank torque up to 33 N·m. The engine operating points are same for Engine A, Engine B and Engine C as they have the same gear ratios.

Engine A is a NA engine aiming for high thermal efficiency as combustion enhancement technologies such as high compression ratio (CR), high tumble port, variable valve timing (VVT), cooled exhaust gas recirculation (EGR), and lean burn. The BSFC best point is around an engine speed 3500 rpm and a crank torque 50 N·m, which is on the lower engine speed than the base engine. The main concept is to increase the required throttle opening at low loads and reduce pumping losses through lean or dilute combustion, and the engine is characterized by a wide low BSFC range centered on the BSFC best point.

Engine B is a downsized turbocharged engine with displacement downsized to 412 cc, high CR, high tumble port, VVT, cooled EGR and turbocharger. The fuel injection system adopted direct injection (DI) to counter knock caused by boosting and to suppress fuel blow-off during intake valve and exhaust valve overlap. The BSFC below 10 N·m is improved compared to Engine A. The main concept is to reduce pumping losses like Engine A. Turbocharger adopted in engine B is able to supply a larger mass of air than the same displacement NA engines by compressed air and that enables output targets to be achieved with a relatively small displacement compared to NA engines. In the low-load range, where a large driving force is not required, the required throttle opening can be increased, and the pumping loss can be reduced compared to a large-displacement NA engine.

Engine C combines a high CR, high tumble port, VVT, cooled EGR and cylinder deactivation with the base engine and operates with a displacement of 344.5 cc during cylinder deactivation. The dashed line indicates

the cylinder deactivation on and off boundary. In the analysis, cylinder deactivation is mechanically switched on and off with reference to the boundary line, irrespective of the operating conditions. The cylinder deactivation used in Engine C is a mechanism that operates at a cylinder number with high engine efficiency according to the required driving force. In general, cylinder deactivation is used in the low-load range to reduce pumping losses by increasing the required throttle opening by operating with a smaller displacement than the actual total displacement.

The CO2 emissions in the calculation results are shown in figure 4. According to the calculation results, engine A showed CO2 emissions of 78.8 g/km, falling short of the target value of 65 g/km. Engine B achieved the target value of CO2 emissions of 64.3 g/km. Engine C showed CO2 emissions of 66.1 g/km, which was short of the target value, but confirmed its potential to achieve the target value.

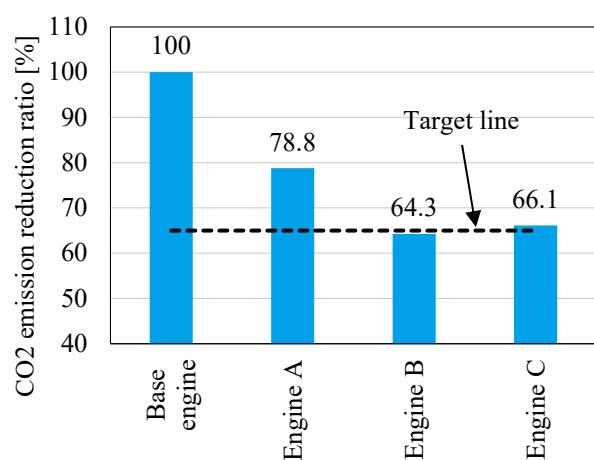


Fig. 4 Calculated CO2 emission

All concepts aim to improve BSFC overall, and reduce pumping losses especially at low loads, but the analysis results show that mechanisms such as Engine B and Engine C, which run at low displacement at low loads, are effective in significantly reducing CO2 emissions. These are concepts that can be called displacement-on-demand, and it was found that a reduction in displacement at low loads is necessary to reduce CO2 emissions to the 65 g/km in sports MCs.

### 3 ENGINE DESIGN

Based on the results of the concept research and our previous development experience<sup>[5]</sup>, we decided to select downsized turbocharging as one of the technology options to achieve CN for verification. In conducting a detailed study of the downsizing turbocharger, the target values were set in table 2. In addition to the CO2 emission targets, these target values were set to ensure the dynamic performance of the sports MCs. The CO2 emissions in the demonstration were only measured in a hot-start due to the workloads required engine calibration. The difference in CO2 emissions between for a cold-start and a hot-start was calculated and 61 g/km was set as the target value for the demonstration.

Table 2 Target value

CO2 emission in WMTC class3-2	65 g/km (simulation at cold-start) 61 g/km (demonstration at hot-start)
Maximum crank power	62 kW 7000 rpm
Maximum crank torque	93 N·m 3000 to 6400 rpm
Vehicle weight	210 kg PWR = 3.39
Throttle response	Less than 1000 msec at 3rd gear Road load
Acceleration performance Time to travel 200 m	From 80 km/h 6.6 sec at 6th gear Road load
	From 100 km/h 5.8 sec at 6th gear Road load
	From 120 km/h 5.2 sec at 6th gear Road load

#### 3-1. Selection of displacement

To select the displacement, the CO2 emissions for the combination of displacement and gear ratio were calculated. The BSFC map of the engine for the study is shown in figure 5, which is the previously prototyped NA engine with high CR, high tumble port, VVT and cooled EGR. The fuel supply system for this engine is port fuel injection, but the demonstration engine is adopted DI based on the concept of the engine B in Table 1. The BSFC best point indicates 220 g/kWh around an engine speed 3500 rpm and crank BMEP 9 bar, and a wide low BSFC area centered on the best point is characteristic. Driving simulations were conducted by changing the combination of displacement and gear ratio based on this map.

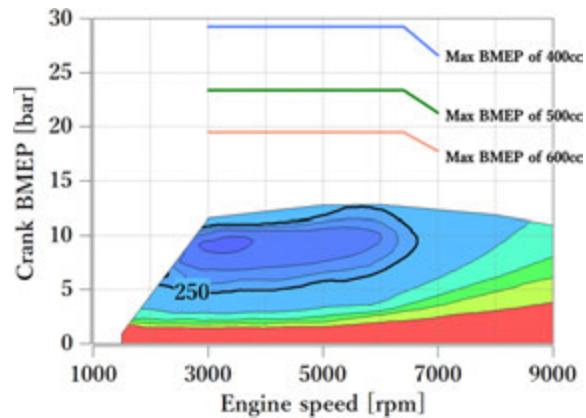


Fig. 5 BSFC map for the study

Figure 6 shows the CO2 emissions in the analysis results. The analysis results show that there is a CO2 emission target limit between 470 cc and 560 cc depending on the combination with the overall gear ratio at 6th gear. Further downsizing is expected to further reduce CO2 emissions, but excessive downsizing increases the required higher brake mean effective pressure (BMEP) and there are concerns about pressure and reduced thermal efficiency due to the reduction of compression ratio as a countermeasure will increase<sup>[9]</sup>.

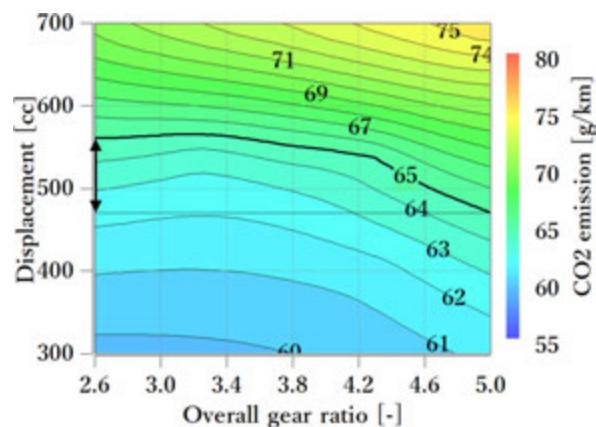


Fig. 6 Calculated CO2 emissions for combinations of displacement and Overall gear ratio at 6th gear

In this study, downsizing is kept to the minimum necessary to minimize the impact on dynamic performance. The bore was set at 70 mm in consideration of fuel adhesion to the cylinder from the DI penetration. The stroke bore ratio should be around 1.0 to ensure turbulence flow in the cylinder, so the stroke was set at 70.2 mm and the total displacement was downsized to 540 cc.

### 3-2. Turbocharger matching

Selection was conducted by creating an engine performance analysis model using 1D simulation with Gamma Technologies' GT-SUITE. Result of conventional turbocharger analysis showed that it was possible to achieve maximum power output, but there were concerns that low end torque (LET) would not be achieved and acceleration performance would deteriorate due to insufficient exhaust gas volume in the low engine speed range. As a countermeasure, it was decided to adopt an E-Turbo provided by Garrett Motion as shown in figure 7. E-Turbo has a permanent magnet synchronous motor as motor generator (MG) integrated in the center housing between the turbine and compressor stage, and by supplying power to this MG, driving force can be applied to the turbine shaft, which can assist in driving the compressor. As supercharging is possible without waiting for an increase in exhaust energy, an increase in torque and improved throttle response, especially in the low engine speed range, are expected. The E-Turbo also has a harvesting capability, which was not part of these investigations. In line with the adoption of an E-Turbo, the power generation system is based on the 12 V alternating current motor-generator used on existing motorcycles, with modified magnet and winding specifications and a 48 V - 2 kW class MG.



Fig. 7 Structure of the Garrett E-Turbo

The results of the wide-open throttle (WOT) performance analysis are shown in figure 8. It was confirmed that the crank torque at engine speed 3000 rpm did not reach the target value with the conventional turbocharger, but with the addition of 0.88 kW of electric assist (E-assist) by the

E-Turbo, the target torque was achieved. The results of the rapid throttle opening response analysis conducted using a same calculation model are shown in figure 9. The calculations were conducted with the engine speed fixed at 3000 rpm and the waste gate actuator (WGA) fully closed. As WGA control is not implemented, the torque continues to increase after the target torque is reached. In this paper, time to torque is defined as the time from the start of acceleration until 90% of the maximum torque is reached, and this is evaluated. By performing E-assist, the torque rises more steeply than without E-assist, and the target time to torque of 1.0 sec can be achieved. As shown in figure 8, w/o E-assist, the target LET is not reached even in the steady state, so the target LET is not reached in the transient analysis.

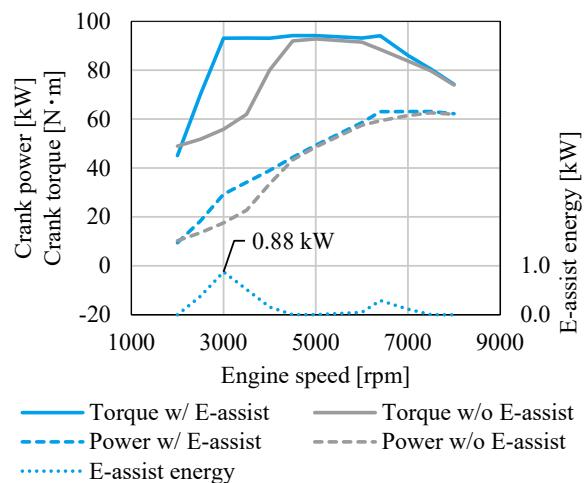


Fig. 8 Calculated crank power and torque as a function of engine speed

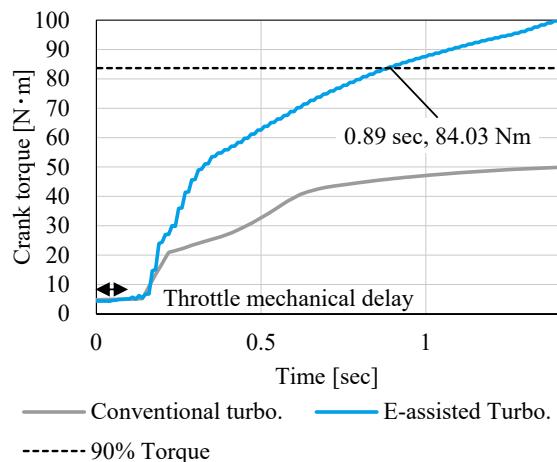


Fig. 9 Calculated crank torque of acceleration

### 3-3. Selection of gear ratio

The CO2 emission calculation results in figure 6 show that for a displacement of 540 cc, over all gear ratio at 6th gear of 2.993 to 4.240 satisfies the CO2 emission target. From this range, over all gear ratio at 6th gear was selected to 3.972 by balancing the vehicle's acceleration performance. Together with the Automatic transmission (AT) mode, this setting allows the selective use of a low engine speed, high engine load region with good thermal efficiency in WMTC mode. Figure 10 shows the calculated overtaking acceleration performance from the driving simulation. The overtaking acceleration performance is evaluated in terms of the time to travel 200 m after starting acceleration at 80, 100 and 120 km/h. The results of the simulation showed that the target acceleration performance could be achieved with an over-all gear ratio of 3.972 or higher.

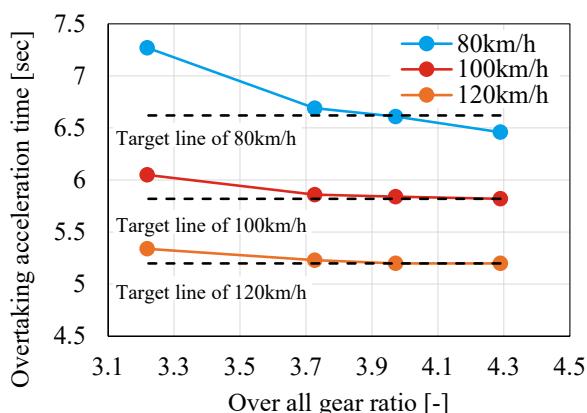


Fig. 10 Calculated overtaking acceleration time

Figure 11 shows the electrical energy used for E-assist during acceleration. The smaller the overall gear ratio, the more electrical energy is required. An overall gear ratio of 3.972 used around 4 kJ of electrical energy for all acceleration conditions. E-assist is assumed to be powered by external power supply such as a battery or capacitor, but it may also be possible to generate the power using an MG mounted on the crankshaft and use it immediately as assist power without charging the power source, which is called "on demand assist". The feasibility of on demand assist will be tested in demonstrations.

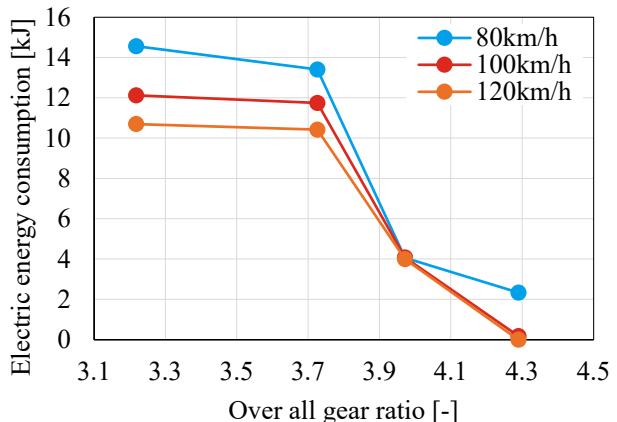


Fig. 11 Calculated electric energy consumption

### 3-4. Decision specifications

Specifications of the demonstration engine incorporating these design concepts are shown in table 3.

Table 3 Fixed specifications

Displacement [cc]	540	
Cylinder number	2	
Bore [mm]	70	
Stroke [mm]	70.2	
Stroke/Bore ratio	1.00	
Compression ratio	13.6	
Intake system	E-Turbo	
Injector type	DI	
Maximum power [kw]/[rpm]	62 / 7000	
Maximum torque [N·m]/[rpm]	93 / 3000 to 6400	
Gear ratio	Primary	1.681
	1st	2.667
	2nd	1.833
	3rd	1.421
	4th	1.200
	5th	1.037
	6th	0.879
	Secondary	2.688
Final gear ratio at 6th gear	3.972	
Other items	Tumble port	
	VVT	
	Cooled EGR	
	48V MG	
	AT	

An example vehicle layout for the demonstration engine is shown in figure 12. As an example, a mass-produced sports MCs frame was used, the demonstration engine is certified to be able to mount on an existing motorcycle body. The vehicle weight was 207.7 kg at the designed value, including the exterior, and the target vehicle weight was achieved.



Fig. 12 Example of a vehicle layout

## 4 DEMONSTRATION

The engine investigated in the section “ENGINE DESIGN” was prototyped and evaluated. The turbocharger was prototyped at Garrett Motion. An overview diagram of the prototype engine and evaluation system is shown in figure 13. The evaluation system used a virtual real simulator (VRS) dyno, which enables the engine to be evaluated as a stand-alone unit simulating vehicle driving. The prototype engine has equivalent to those assumed to be installed in a vehicle except for a cooling system and a power supply system. The cooling system is connected to an external heat exchanger for water temperature control, and the power supply system is connected to an external power supply for testing under various power supply conditions.

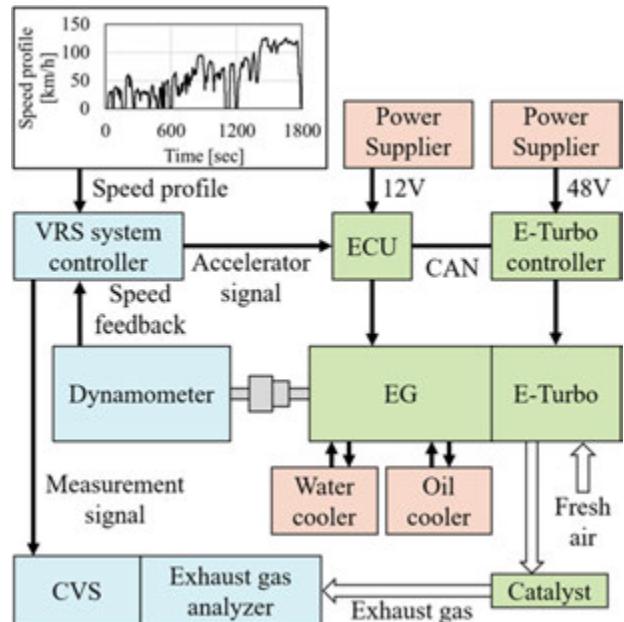


Fig. 13 Outline of dyno system

### 4-1. Static performance

The BSFC map for the demonstration engine is shown in figure 14. Compared to Figure 3 (a) base engine, the low BSFC region has expanded: the best BSFC point indicated approximately 220 g/kWh around an engine speed 3500 rpm and crank torque 50 N·m, which is equivalent to the best BSFC point in figure 5. Due to the vibration of dyno system, BSFC was not able to be measured below 2000 rpm.

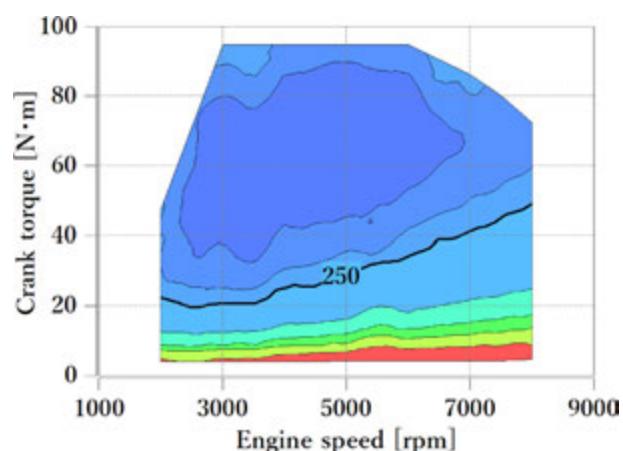


Fig. 14 Measured BSFC map of prototype engine

Figure 15 shows the power performance. The crank power and crank torque achieve the target value, the target torque is not reached in the low engine speed range without E-assist as mentioned in turbo matching.

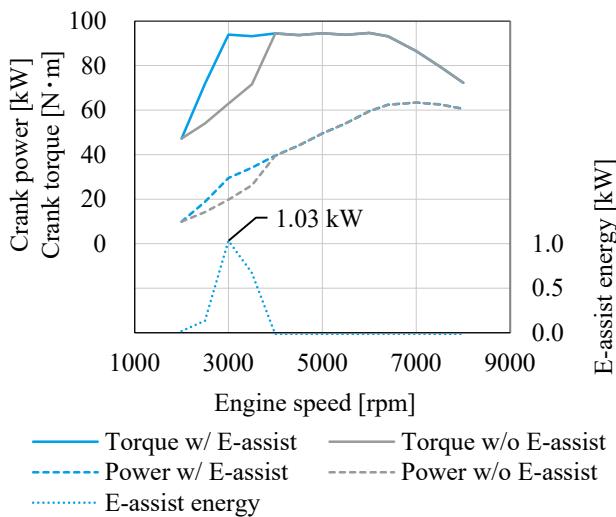


Fig. 15 Measured WOT performance

The knock index and exhaust gas lambda are shown in figure 16. The knock index is within the threshold throughout the all engine speed range. The exhaust gas lambda is measured at the exhaust pipe collective area and is adapted to target stoichiometry for all engine operating points. At engine speed 2500 rpm, the exhaust gas lambda is leaner because the intake and exhaust valve overlap section is enlarged by the VVT and the scavenging effect is used.

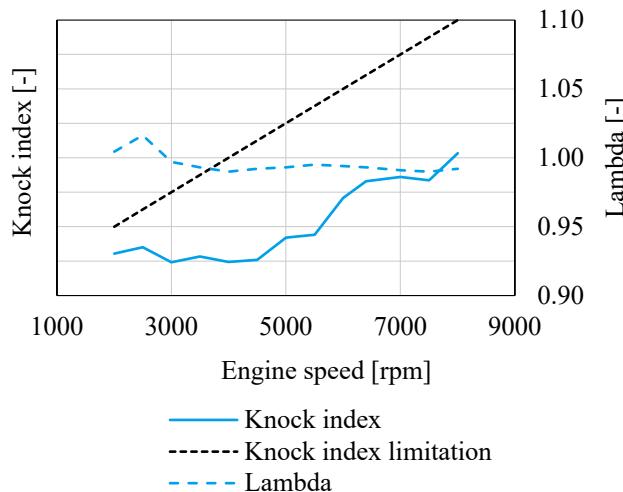


Fig. 16 Measured knock index and exhaust gas lambda

Figure 17 shows the dynamic CR and mass fraction burned 50% (MFB50%). The CR adopted is 13.6, which is a high setting for a supercharged engine, but by adjusting the dynamic CR with VVT, MFB50% could be

set to a maximum of about 25 degree crank angle (degCA).

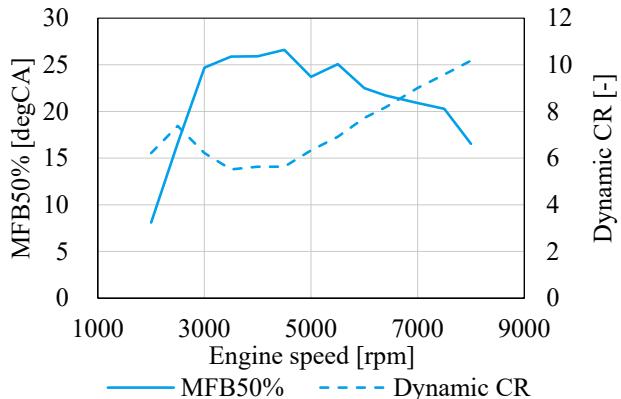


Fig. 17 Measured MFB50% and dynamic CR

#### 4-2. Throttle response

E-assist which uses an external power supply and on demand assist which utilizes the MG were evaluated. On demand assist was assisted by an external power supply due to the number of workloads required for engine calibration, and the generating loss was reproduced by increasing the driving resistance. Figure 18 shows the results of the rapid throttle opening test. Without E-assist, the torque rise was slow and increased with increasing engine speed, with a time to torque of 3.08 sec. External assist, the torque rise was steep above 45 Nm. The time to torque was 0.91 sec, which is close to the analysis results shown in figure 9. In on demand assist, the throttle response improved to 1.12 sec compared with no E-assist, although it did not reach the target value. Figure 19 shows the electric power and energy used by E-assist. External assist gives a maximum power output of more than 2.5 kW and uses 1.81 kJ of energy when the torque is reached at 90% of the maximum torque. On demand assist gives a maximum output of around 1.8 kW and uses 1.38 kJ of energy when the torque is reached at 90% of the maximum torque. The electric power for E-assist of on demand assist is fixed by the MG specification, so a higher capacity MG can increase the electric power for E-assist and shorten the time to torque.

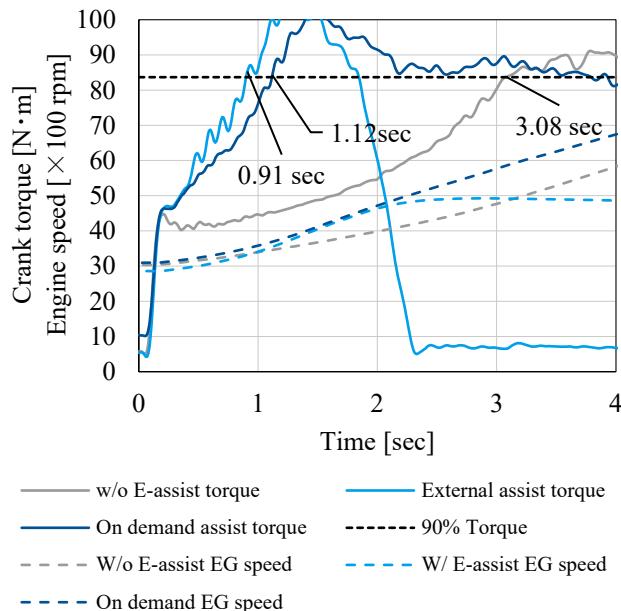


Fig. 18 Measured crank torque of acceleration

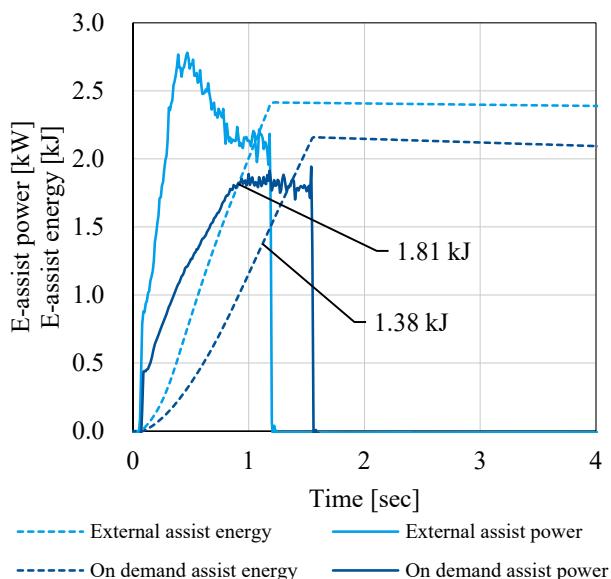


Fig. 19 Measured E-assist power and energy of acceleration

#### 4-3. CO2 emission

The CO2 emissions in WMTC mode are shown in figure 20 and the engine operating points on the BSFC map in figure 21. The CO2 emissions target for demonstration of 61 g/km was achieved. According to the simulation calculations, the CO2 emissions deterioration due to cold-start is 3.8 g/km, which means that the cold-start target of 65 g/km is also expected to be achieved. As in the simulation described in the previous section, the AT is

used to actively use the high thermal efficiency area with low engine speed and high engine load, and the shift schedule is programmed to use as high a number of gear steps as possible within the range where engine vibration is tolerated. In this demonstration, regeneration and assist by crank-mounted MG, harvesting capability by E-Turbo and idling stop have not yet been implemented. Further reductions in CO2 emissions can be expected by incorporating these systems.

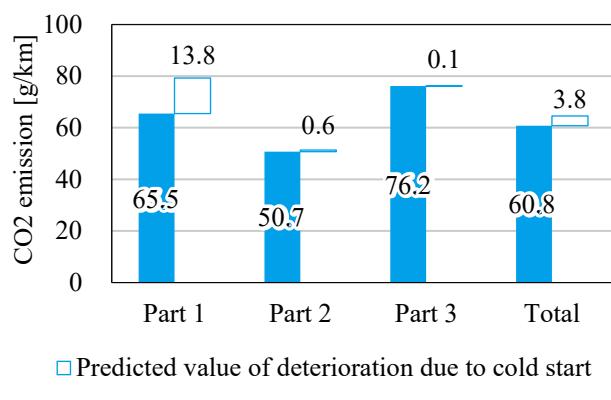


Fig. 20 Measured CO2 emission in WMTC

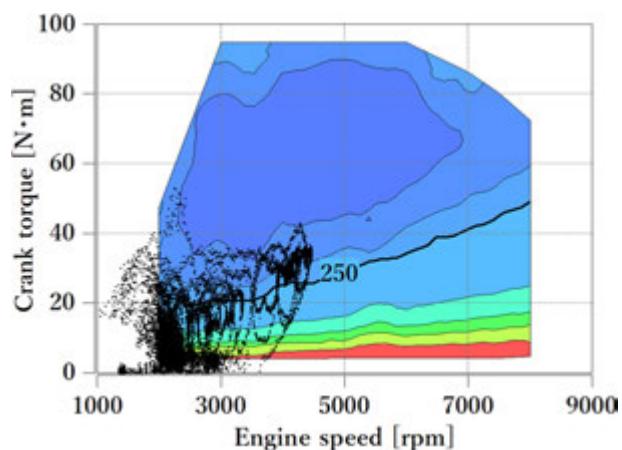


Fig. 21 Measured engine operating points on BSFC map

#### 4-4. Overtaking acceleration

As in the throttle response evaluation, external assist and on demand assist were evaluated. Overtaking acceleration times are shown in figure 22. The time of external assist was achieving the target time, and on demand assist was almost on target; the difference between external assist and on demand assist is greater at lower initial vehicle speeds at the start of acceleration. As a representative

example of time continuous data, data at an initial speed of 80 km/h is shown in figure 23. The torque of external assist increases more quickly after 50 N·m than on demand assist. At this time, external assist is assisted by the allowable transient electric power in about 1 sec from the start of acceleration, and then the allowable steady-state electric power is continuously assisted. The crank torque is controlled by opening the WGA, as the setting is such that the assist continues to be applied even after the crank torque reaches the upper limit. On demand assists at a maximum of about 1.8 kW, as the assist power is limited by the MG specifications. After the crank torque is reached to upper limit, the system is set to no E-assist. The electric energy for E-assist used at the end of overtaking acceleration was 13.45 kJ for external assist and 4.79 kJ for on demand assist. Although the overtaking acceleration time can be reduced by continuing to E-assist with external power supply, the fact that the target value can be satisfied almost entirely with on demand assist means that an electric energy around 5 kJ is sufficient to achieve the effect of the E-assist. In addition, because the electric power for E-assist is output from the crankshaft in on demand assist, the electric energy is less than 4.79 kJ when aiming for a time equivalent to the target value with external power supply. Acceleration from 80 km/h is the condition under which the engine speed at the start of acceleration is around LET, and therefore the most benefit from the assist can be obtained. At 100 and 120 km/h, when the engine speed is higher at the start of acceleration, the turbo drive power from the exhaust gases increases, so the effect of the E-assist becomes relatively small.

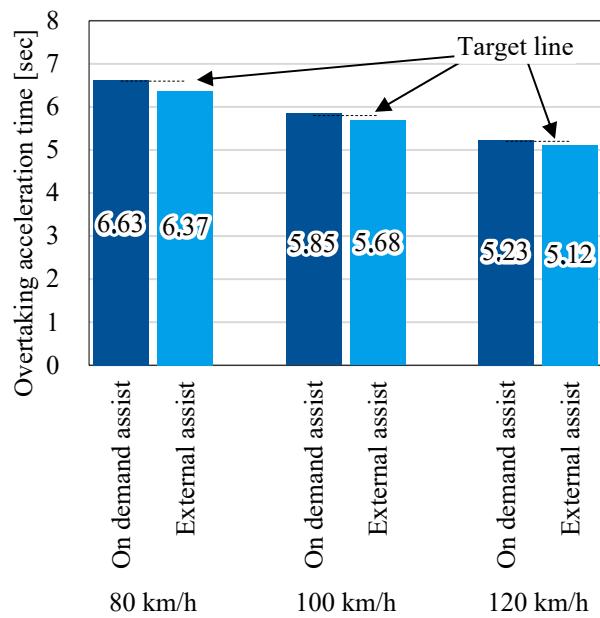


Fig. 22 Measured overtaking acceleration time

## 5 SUMMARY/CONCLUSIONS

The CO2 emissions of 65 g/km are achieved by downsizing engine with turbocharger or cylinder deactivation, making it clear that displacement-on-demand concepts are effective to 65g/km level of CO2 emissions reduction.

The concept of a highly thermally efficient downsized turbocharged engine combined with E-Turbo has proven that the engine can meet CO2 emission and dynamic performance targets and can be installed in MCs.

An evaluation of the prototyping engine has demonstrated that the selected concept can achieve the target values for CO2 emissions and dynamic performance.

The downsized turbocharged engine with E-Turbo enabled an elevated level of compatibility in terms of both CO2 emissions and dynamic performance and demonstrated one of the directions in the evolution of MCs towards achieving carbon neutrality.

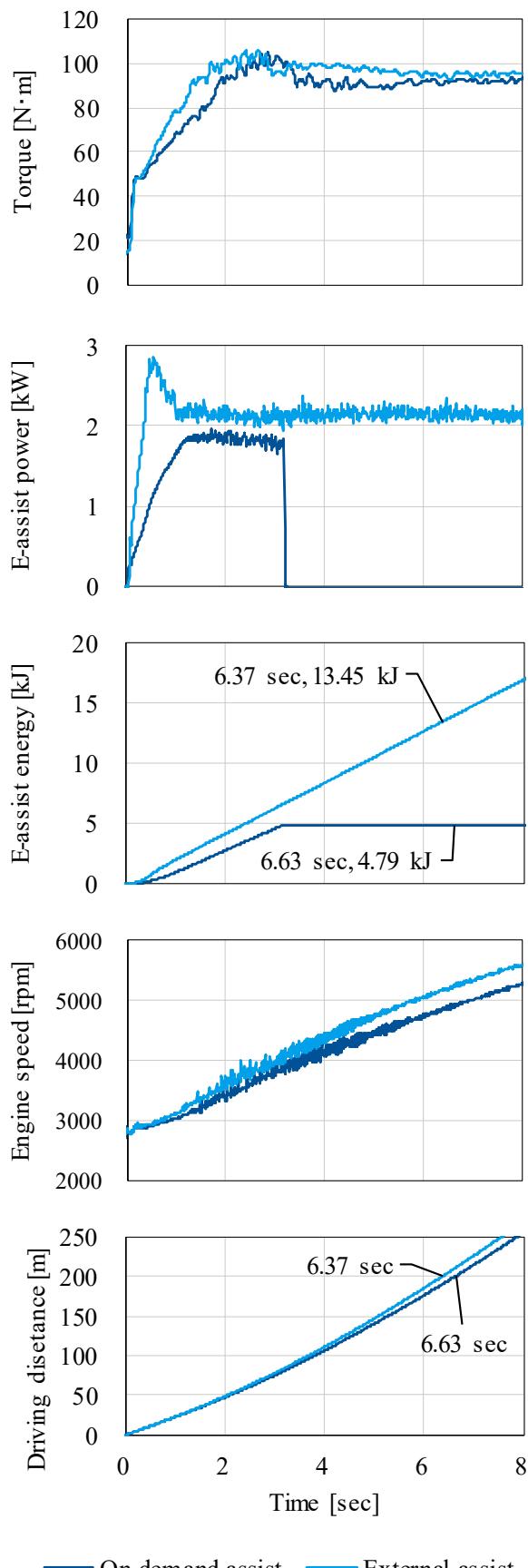


Fig. 23 Measured driving data of overtaking acceleration from 80 km/h

## ACKNOWLEDGMENTS

The authors are grateful to Garrett Motion and Vitesco Technologies France S.A.S for kind collaboration and support on this study.

## REFERENCES

- [1] 経済産業省(Ministry of Economy, Trade and Industry), “合成燃料研究会 中間とりまとめ,” [https://www.meti.go.jp/shingikai/energy\\_environment/gosei\\_nenryo/pdf/20210422\\_1.pdf](https://www.meti.go.jp/shingikai/energy_environment/gosei_nenryo/pdf/20210422_1.pdf), accessed March 2024.
- [2] 経済産業省(Ministry of Economy, Trade and Industry), “合成燃料(e-fuel)の導入促進に向けた官民協議会 2023年中間とりまとめ,” [https://www.meti.go.jp/shingikai/energy\\_environment/e\\_fuel/pdf/2023\\_chukan\\_torimatome.pdf](https://www.meti.go.jp/shingikai/energy_environment/e_fuel/pdf/2023_chukan_torimatome.pdf), accessed March 2024.
- [3] Furuta, H. and Yoshida, J., “Hybrid Electric Two-Wheeled Vehicle Fitted with an EVT System (Electrical Variable Transmission System),” SAE Technical Paper 2023-01-1853, 2023, doi:10.4271/2023-01-1853.
- [4] Matsuda, Y., “Hybrid-MC as a Solution in the Transit Stage for the Carbon Neutral Society,” SAE Technical Paper 2023-01-1854, 2023, doi:10.4271/2023-01-1854.
- [5] Sato, H., Torigoshi, M., Takase, H., and Makita, N., “Feasibility study of boosted DI technology for sport motorcycle,” SAE Technical Paper 2022-32-0079, 2022, doi:10.4271/2022-32-0079.
- [6] Kargul, J., Stuhldreher, M., Barba, D., Schenk, C. et al., “Benchmarking a 2018 Toyota Camry 2.5-Liter Atkinson Cycle Engine with Cooled-EGR,” SAE Int. J. Advances & Curr. Prac. in Mobility 1(2):601-638, 2019, doi:10.4271/2019-01-0249.
- [7] Claus, G., Achim, K., and Ingo, H., “The Boosting System – A Key Technology also for Electrified Combustion Engines?,” presented at 28th Aachen Colloquium Automobile and Engine Technology 2019, DEU, October 7-9, 2019.
- [8] Urata, Y., Kondo, T., and Takabayashi, T., “Gasoline Engine Combustion Technology in Honda,” Journal of the Combustion Society of Japan, 60 (191): 18-26, 2018.
- [9] Michael, B., Benjamin, H., and Jonathan, H., “Dynamic Downsizing for Gasoline Engines,” presented at 24th

Aachen Colloquium Automobile and Engine Technology  
2015, DEU, October 5-7, 2015.

■著者



牧田 直希  
Naoki Makita  
パワートレイン開発本部  
先行企画開発統括部  
先行企画開発部



鳥越 昌樹  
Masaki Torigoshi  
パワートレイン開発本部  
先行企画開発統括部  
先行企画開発部



高橋 俊彦  
Toshihiko Takahashi  
パワートレイン開発本部  
先行企画開発統括部  
先行企画開発部



高瀬 浩規  
Hiroki Takase  
パワートレイン開発本部  
先行企画開発統括部  
先行企画開発部

## DEFINITIONS/ABBREVIATIONS

CN	Carbon Neutrality
ICEs	Internal Combustion Engines
WMTC	Worldwide-harmonized Motorcycle Test Cycle
E-Turbo	Electrically assisted Turbocharger
PWR	Power Weight Ratio
HEVs	Hybrid Electric Vehicles
1D	1 Dimensional
BSFC	Brake Specific Fuel Consumption
FMEP	Friction Mean Effective Pressure
NA	Naturally Aspirated
CR	Compression Ratio
VVT	Variable Valve Timing
EGR	Exhaust Gas Recirculation
DI	Direct Injection
BMEP	Brake Mean Effective Pressure
LET	Low End Torque
MG	Motor Generator
WOT	Wide-Open Throttle
E-assist	Electric assist
WGA	Waste Gate Actuator
AT	Automatic Transmission
VRS	Virtual Real simulator
MFB50	Mass Fraction Burned 50%
degCA	degree Crank Angle

# Human model on multi-body dynamics simulation of Motorcycle

Motohito Ueki Akihiro Takayama Noboru Yabe

当論文は、JSAE 20249005/SAE 2024-32-0005として、SETC2024 (Small Powertrains and Energy Systems Technology Conference)にて発表されたものです。

Reprinted with permission Copyright © 2024 SAE Japan and Copyright © 2024 SAE INTERNATIONAL  
(Further use or distribution is not permitted without permission from SAE.)

## 要旨

二輪車の動力学シミュレーションへのライダーモデル追加に関する可能性と課題を検討した。

本研究では、二輪車の動きを運動方程式で再現する10自由度(10-DoF)モデルに加えて、人体の各部位をマスとジョイントで構成した人体モデルを追加した。人体モデルは、腕と肩の角度を制御することで操舵角を操作可能とした。

このモデルの効果を検証するために、固有値解析、定常円旋回、スラローム走行の3つのシミュレーションを実施した。

固有値解析においては、両手をハンドルバーに固定した状態で、ウォブルモードの固有値が根軌跡上で安定側へとシフトすることが確認された。

走行シミュレーションでは、人体モデルの有無による走行影響に加え、人体モデルを介したハンドル操作と直接ハンドルにトルクをかける操作による走行影響が見られた。特にスラローム走行では人体モデルを介したことでの操作遅れが確認された。

これらの結果から、人体モデルがシミュレーションの振動および走行軌跡に影響を及ぼすことが示唆された。

今後はシミュレーションに応じた、人体モデルパラメータの適切な調整が求められる。

## Abstract

The possibilities and challenges of adding a rider model to the motorcycle dynamics simulation were investigated for the future planning of a full virtual test.

The human model was added to a multi-body dynamics model that reproduces the equations of motion of a motorcycle, called the 10 degrees of freedom (10-DoF) model. The human model is composed from multiple masses and joints, and the steering angle can be controlled by determining the angle of the arms and shoulder. To study the effect of this model, three distinct simulations were carried out: 'the eigenvalue analysis', 'the steady-state circular test simulation' and 'the slalom running simulation'.

In the eigenvalue analysis, the eigenvalues of the wobble mode shifted to a stable side in the root locus when both hands were fixed on the handlebars.

As a result of the slalom running simulation, the response of the handlebar control through the human model produced a more convex trajectory than a direct control of the steering angle.

For a full virtual test in the future, the human model has some effects to the vibration and trajectory modes of a running simulation. Hence, depending on the purpose of the simulation, the parameters of the human model should be calibrated to fit that purpose.

## 1

## INTRODUCTION

The electrification of vehicles, including motorcycles, is an effective way of achieving a carbon-neutral society. However, an electric unit is a power source with different

characteristics from a conventional engine, which also affects the kinematic characteristics of the vehicle. To design motorcycles that take advantage of the characteristics of electric power units, simulation and drivability evaluation technologies that consider the

kinematic characteristics are necessary. In addition, motorcycle customers use motorcycles not only for daily use, but also for hobbies such as touring on holidays. Such customers prefer gentle and smooth acceleration and deceleration in their daily use, and in touring situations they prefer vehicles that are powerful yet easy to handle with no unnecessary body movements.

In addition, leaning and turning motorcycles that are lightweight will have a large impact on the body of the vehicle from the human weight transfers and steering maneuvers. Vibrations such as chattering and shift shock are also easily transmitted to the human body, causing unpleasant feelings.

In order to develop products in a short period of time while meeting these complex product requirements, it is necessary to improve the level of simulation technology at the planning stage.

Accordingly, considering prior research, a human model generated with the Biomotion module was incorporated into the motorcycle vehicle motion analysis model on SIMPACK, a multi-body dynamics tool. This allowed us to identify challenges associated with incorporating the human physical model into the analysis.

## 2 METHOD

In this paper, motorcycle and rider models were prepared for the simulation and assembled to form one model for the calculations. The respective models and modelling methods are presented as follows.

### 2-1. Motorcycle model

The motorcycle model is based on the 10-DoF equations of motion.

The equations of motion for a motorcycle model with 10-DoF are referenced as follows; the 10-DoF model includes the four degrees of freedom(4-DoF); steering rotation, lateral displacement, vehicle body yaw, and vehicle body roll, that are similar to Sharp's 4-DoF model<sup>[1]</sup>.

The frame consists of a front fork, main frame, and rear swing arm. The degrees of freedom for torsion and lateral bending are represented by lumped stiffness elements. Therefore, six more degrees of freedom are considered in addition to Sharp's 4-DoF model as shown in Fig. 1.

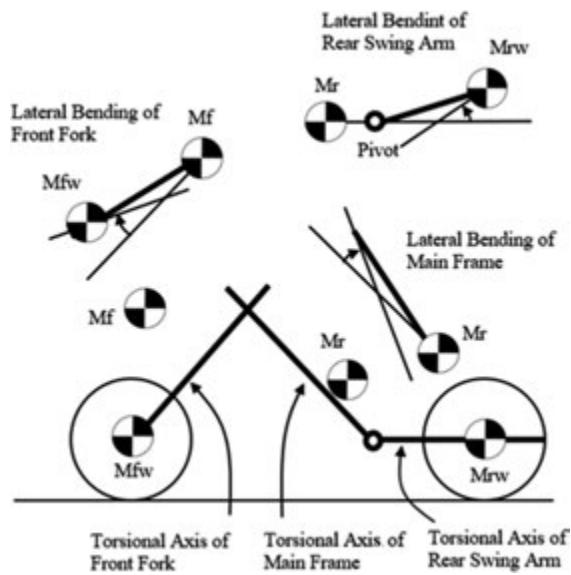


Fig. 1 Geometry of 10-DoF Equations Model "Aoki, et al.; Analysis of the effect of frame rigidity on the straight-line stability of motorcycles, Japan Transactions of the Society of Mechanical Engineers of Japan ©, Vol. 64, No. 625 (1998-9) Paper No.97-1"99"

Aoki et al.<sup>[2]</sup> reported the effect of frame rigidity on the straight-line stability of motorcycles together with the results of actual motorcycle verification.

Terayama et al.<sup>[3]</sup> reproduced a 10-DoF model with multi-body dynamics, which is used in this study.

### 2-2. Rider model

The rider model consists of a human model and a maneuvering controller that was created using the Biomotion module.

#### 2-2-1. Human model

The Biomotion human model is constructed from joints and spring elements connecting several masses. This model rides on the motorcycle model as shown in Fig. 2. The connections are shown in Fig. 3.

Hands-Handle bar: The hands are connected using joints and by aligning the coordinates of the hands with the coordinates of the handlebar position.

Pelvis-Frame: The pelvis is aligned with the seat position coordinates on the frame and connected with a joint.

Legs-Frame: The legs are aligned to the step position coordinates on the frame and are constrained.

The mass and spring characteristics of the human model are shown in Tables 1 and 2. Table 3 shows the characteristics of the connections between the human body parts and the motorcycle.

The parameters in this study used the default values available in Biomotion.



Fig. 2 Rider model and motorcycle model

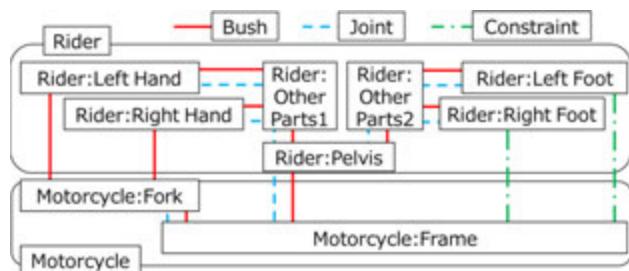


Fig. 3 Relationship between motorcycle model and the human body parts

Table 1 Mass properties of the humanbody parts

Part	Mass [kg]	I <sub>xx</sub> [kg m <sup>2</sup> ]	I <sub>yy</sub> [kg m <sup>2</sup> ]	I <sub>zz</sub> [kg m <sup>2</sup> ]
Head	4.814	$2.730 \times 10^{-2}$	$2.452 \times 10^{-2}$	$2.030 \times 10^{-2}$
Neck	1.712	$6.387 \times 10^{-3}$	$6.387 \times 10^{-3}$	$3.008 \times 10^{-3}$
Thorax	15.78 <sup>1</sup>	$1.760 \times 10^{-1}$	$7.242 \times 10^{-2}$	$1.491 \times 10^{-1}$
Lumbus	10.60	$1.324 \times 10^{-1}$	$8.468 \times 10^{-2}$	$5.362 \times 10^{-1}$
Upper arm	2.056	$1.278 \times 10^{-2}$	$1.145 \times 10^{-2}$	$3.966 \times 10^{-3}$
Forearm	1.262	$6.546 \times 10^{-3}$	$6.079 \times 10^{-3}$	$1.299 \times 10^{-3}$
Hand	0.486	$1.326 \times 10^{-3}$	$8.840 \times 10^{-4}$	$5.417 \times 10^{-3}$
Pelvis	9.658	$6.669 \times 10^{-2}$	$5.378 \times 10^{-2}$	$6.060 \times 10^{-2}$
Thigh	7.700	$2.027 \times 10^{-1}$	$2.027 \times 10^{-1}$	$4.260 \times 10^{-2}$
Shank	3.215	$3.883 \times 10^{-2}$	$3.738 \times 10^{-2}$	$6.562 \times 10^{-3}$
Foot	0.999	$4.444 \times 10^{-3}$	$2.346 \times 10^{-2}$	$1.483 \times 10^{-2}$

Table 2 Spring constants and damping coefficients between the human body parts

Part	Spring constant [Nm/deg]	Damping coefficient [Nms/deg]
Head-Neck	7.058	$4.458 \times 10^{-1}$
Neck-Thorax	9.849	$6.220 \times 10^{-1}$
Thorax-Lumbus	34.47	2.176
Lumbus-Pelvis	47.60	3.007
Upper arm-Thorax	6.566	$4.147 \times 10^{-1}$
Forearm-Upper arm	2.299	$1.451 \times 10^{-1}$
Hand-Forearm	2.627	$1.659 \times 10^{-1}$
Thigh-Pelvis	5.238	$3.318 \times 10^{-1}$
Shank-Thigh	3.939	$2.489 \times 10^{-1}$
Foot-Shank	3.283	$2.074 \times 10^{-1}$

Table 3 Spring constants and damping coefficients between the human model and the motorcycle

Human model-motorcycle	Spring constant		Damping coefficient	
	Translation [N/m]	Rotation [Nm/deg]	Translation [Ns/m]	Rotation [Nms/deg]
Handlebar-Hand	$6.0 \times 10^4$	5.236	$1.0 \times 10^2$	$1.745 \times 10^{-1}$
Frame-Pelvis	$3.0 \times 10^4$	8.727	$3.0 \times 10^3$	$1.745 \times 10^{-1}$

## 2-2-2. Maneuvering controller

The maneuvering controller uses part of the Biomotion module. It is equipped with acceleration and deceleration controls in the longitudinal speed and the lateral direction, which are controlled by steering angle. To assess the impact of the human model, two distinct control modes were prepared, the SAD mode and the HC mode.

**SAD (Single Axis Direct):** To match the target steering angle, the steering torque is controlled by direct input to the steering axis.

**HC (Handlebars Control):** To match the target steering angle, the steering torque is input to the handlebar through the human body elements, which controls the angle of the forearm and the upper arm.

Both control modes use look-ahead models as shown in Fig. 4. This model checks the trajectory and steers the motorcycle towards the target trajectory in both the steady state circular test simulation and in the slalom running simulation. The rider model looks at the estimated position after 1.1 sec from the present time. In this study, this parameter was tuned to avoid falling, but it has not been validated by any test or verification.

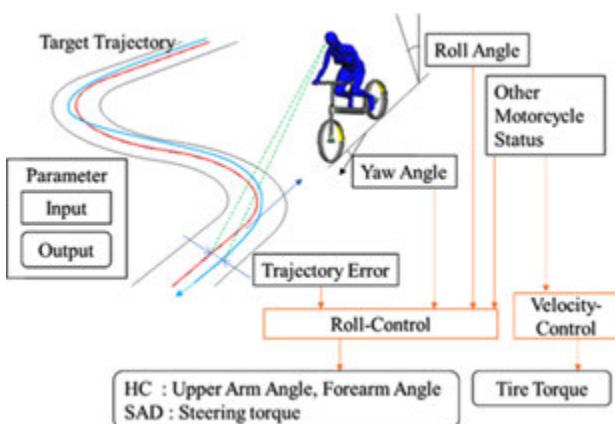


Fig. 4 Maneuvering image

The 'Trajectory Error' is the distance QP, which is the distance in the normal direction between point P, which is a point extended from the current position by the distance of the look ahead on the 'target trajectory', and point Q, which is the extension point of the 'direction of motorcycle

travel', as shown in Fig. 5. The steering angle of SAD or the angles of human parts of HC are calculated by each PID controller. As long as the human body model is represented by a spring mass model, some of the response delay will be included in the result of maneuvering model or in the elements of the human body parts.

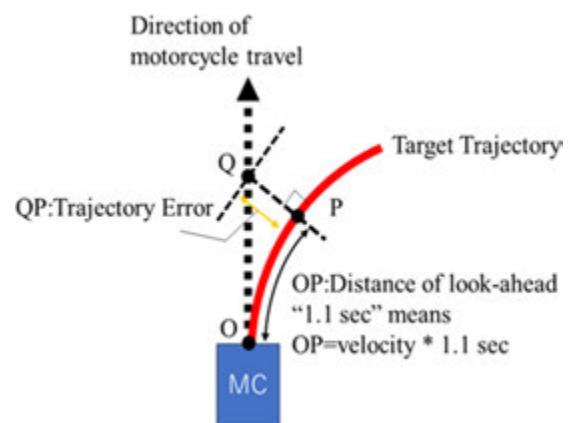


Fig. 5 Trajectory error image

## 2-3. Analysis and Simulation

To study the influence of the human model and the maneuvering model to the motorcycle behavior, the following is the list of analyses or simulations used.

Table 4 List of analyses and simulations

Target	State	Analysis/Simulation
Static characteristics (Eigenvalue Analysis)	Static	Frequency response
Dynamic characteristics (Driving Simulation)	Steady	Steady state circular test
	Transient	Slalom running

To study the influence of maneuvering and dynamic characteristics, the driving simulation was provided with two types of maneuvering controllers, the SAD mode and the HC mode in Biomotion.

The MF-Tire was used as the tire model for all simulations.

### 2-3-1. Eigenvalue Analysis

To study how a rider influences the motorcycle vibration characteristics, two motorcycle models were prepared; one with a rider and the other without a rider. The

wobble and weave modes are typical motorcycle vibration modes with frequencies under 10 Hz. In the past study, we know that the handlebar operation produced different results from the no rider model. Therefore, the frequency analyses were conducted on the motorcycle-only model and the combined rider model.

- Handle-free: The rider's hands are not connected to the handlebars.
- Fixed handlebar: The handlebars are connected to the hands of the human model.
- STD 10-DoF: The rider was represented by a mass property fixed to the motorcycle mass property.

### 2-3-2. Steady state circular test

The radius of the circle was set to 150 m, and motorcycle speed was simulated from 80 km/h to 120 km/h in 10 km/h increments as shown in Table 6. Fig. 6 shows one of the driving simulation cases.

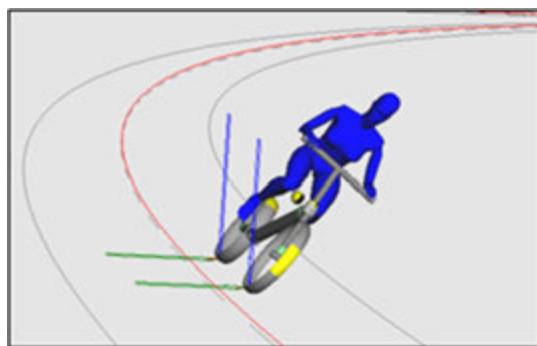


Fig. 6 Driving simulation representing the steady state circular test

To assess the impact of the human model, the two distinct control modes, the SAD mode and the HC mode were differentiated, as detailed in chapter 2-2-2.

### 2-3-3. Slalom running

The slalom test was simulated to study a maneuver controller in a transient state. Fig. 7 shows one of the driving simulation cases. Three rider models were compared as shown in Table 5. The slalom course was defined as sinusoidal, and the amplitude was gradually increased until it reached 200 m. From 200 m to 500 m, the amplitude takes a constant value is shown in Table 6.

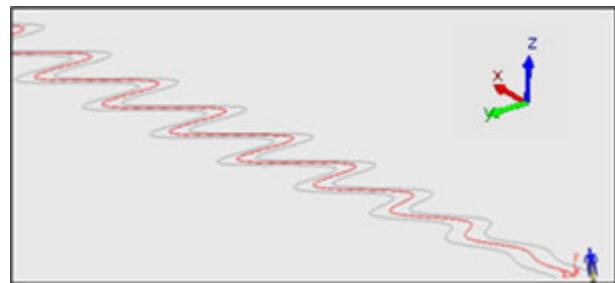
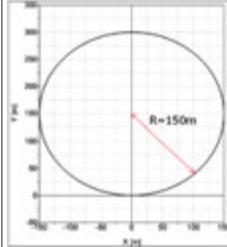
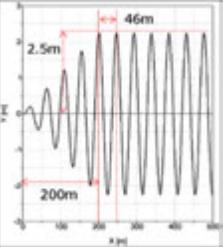


Fig. 7 Driving trajectory in the slalom running simulation

Table 5 Rider models for the slalom running simulation

	Human model	Maneuvering controller
Human model with SAD	Human (Biomotion)	SAD
Human model with HC		HC
STD 10-DoF with SAD	Rigid (STD 10-DoF)	SAD

Table 6 Driving simulation conditions

	Steady state circular test	Slalom Running
Driving course		
Motorcycle speed	80~120km/h (in 10km/h increments)	40km/h

## 3 RESULT

### 3-1. Eigenvalue Analysis

The wobble and weave modes in each condition are shown in the root locus plot as shown in Fig. 8.

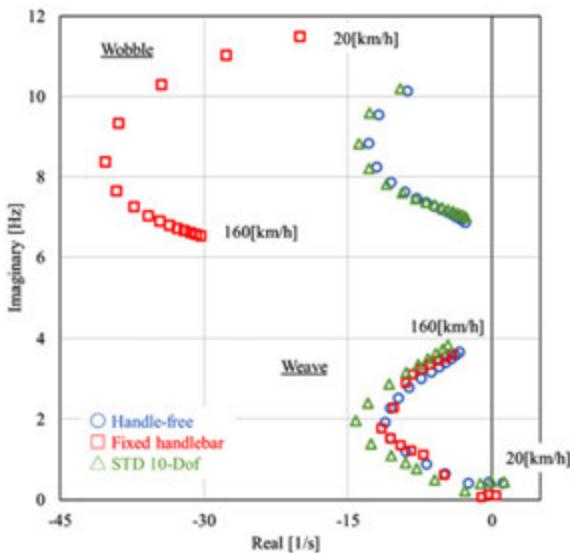


Fig. 8 Root-locus plot for the eigenvalue analysis

The locus of the wobble mode shifted to the stable side in the 'Fixed handlebar' model compared to the 'Handle-free' model. When a human body was attached to the handlebar in the 'Fixed handlebar' model, the steering angle movement was suppressed by a spring and damper system as represented by the human body parts. In the previous paper<sup>[4]</sup>, the force of a handle push increased the number of wobble vibrations and moved the locus to the stable side.

However, in this paper, the results were calculated using the default parameters of Biomotion. These parameters need to be validated by test measurements for a more accurate analysis result.

### 3-2. Steady state circular test

Under the steady state conditions, the criteria for the orientation of the angle between the forearm and the upper arm are shown in Fig. 9. The direction of each arrow shows the axis of rotation.

The simulation results for the steering angle, roll angle, steering torque, front/rear tire lateral forces and

forearm/upper arm angle at different motorcycle speeds are shown in Fig. 10. The legend in Fig. 10 indicates the maneuvering methods (SAD, HC).

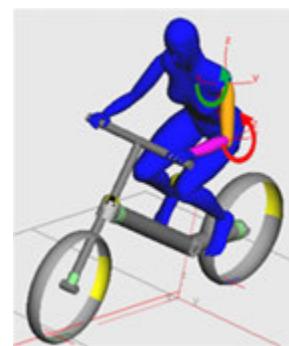


Fig. 9 Reference axes for the Forearm/Upper arm angles on the human model

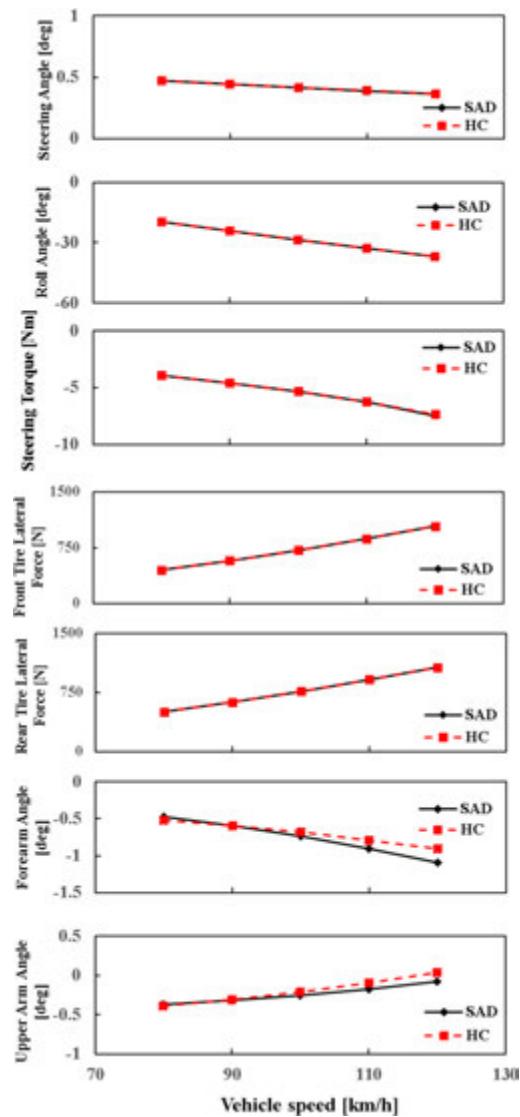


Fig. 10 Results of the 'Steady state circular test' simulation

### 3-3. Slalom Running

The results of the travel trajectory as shown in Fig. 11 show that the motorcycle follows the runway with a larger amplitude in the Y direction than the target trajectory. It was found that sinusoidal running in constant amplitude can itself be performed in the region beyond 200 m of the X-axis distance, although it has moved away from the target. In addition, the trajectory did not shift even if the control method was changed in the same way as for the steady state circular test. Since, it was found that the trajectory was almost the same even if there was a difference in the presence or absence of a human model.

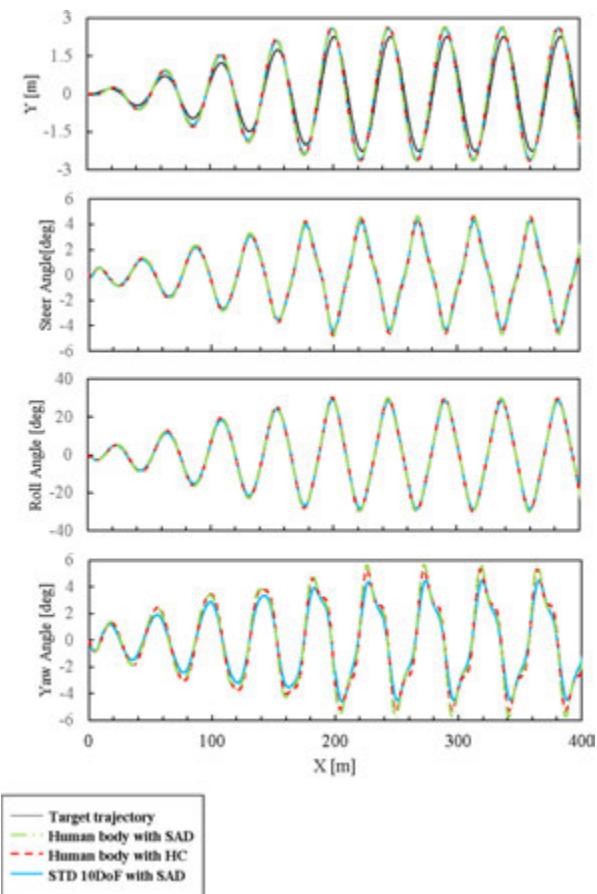


Fig. 11 Results of the 'Slalom running' simulation in 40 km/h

The trajectory of the enlarged chart is shown in Fig. 12, which indicates that the 'STD 10-DoF with SAD' was tracing in the closest proximity to the target trajectory because this had been directly connected to the controller without any delayed movement and effect from

behavior of human parts. The trajectories of the 'Human body with HC' and 'Human body with SAD' models exhibited slightly more convex shapes, with the former turning later than the latter.

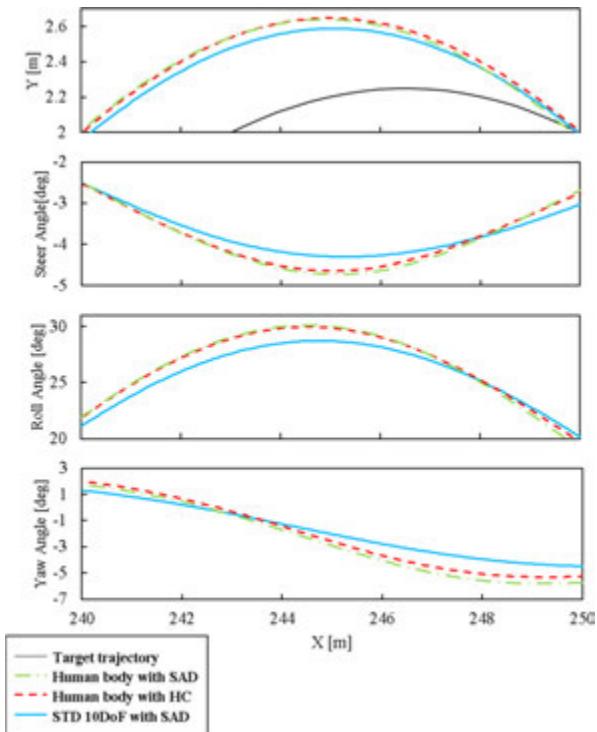


Fig. 12 Enlarged chart of the 'Slalom running' simulation from 240 m to 250 m

Due to the delay in control through the arms, the steering angle of the 'Human body with HC' was smaller than that of the 'Human body with SAD' at 242-246 m as shown in Fig. 11. Also, the roll angle in HC mode was smaller than in SAD mode at 244-248 m due to the delay in the maneuver.

It was explained in chapter 2.2.2 that 'Human body with SAD' determines the angles of the Forearm and Upper arm by directly controlling the steer angle with respect to the target steer angle, whereas 'Human body with HC' controls the steer angle by adjusting the Forearm and Upper arm to the target steer angle and operating the steering. These delays were caused by this control difference.

## 4 CONCLUSION

By using Biomotion on SIMPACK, the human model could be used in the motorcycle multibody dynamics simulation.

The human model affected the eigenvalues of the wobble mode in the same way as in the previous paper. In the slalom running, the human model caused a delay in the steering angle control and the trajectory convex shape.

In this paper, all parameters of the human parts were set to their default values. However, the results indicated that, depending on the purpose, the validation of a spring constant and a damping coefficient is crucial for an accurate simulation. Therefore, one solution is to use the machine learning method<sup>[5]</sup> to calibrate these parameters in the short term.

### ■著者



**植木 幹人**  
Motohito Ueki  
技術・研究本部  
デジタル開発統括部  
MBSE 推進部



**高山 照浩**  
Akihiro Takayama  
技術・研究本部  
デジタル開発統括部  
MBSE 推進部



**矢部 犀**  
Noboru Yabe  
技術研究本部  
デジタル開発統括部  
MBSE 推進部

## REFERENCES

- [1] Sharp, R. S. : The Stability and Control of Motorcycles, Journal of Mechanical Engineering Science, Vol. 13, No. 5, 316-329, 1971.
- [2] Aoki, A., Nishimi, T., Okayama, T., Katayama, T.: Analysis of the effect of frame rigidity on the straight-line stability of motorcycles, Transactions of the Japan Society of Japan Mechanical Engineers (C), Vol. 64, No. 625 (1998-9), Paper No., 97-1799.
- [3] Terayama, T., Yamaguchi, S., Kitagawa, H., Yabe, N., Reproducing the Motorcycle Equation of Motion with Lumped Stiffness on Multi-Body Dynamics Model, SAE Technical Paper 2023-01-1828, 2023.
- [4] Kageyama, I., Kogo, A.: Human factors in the steering system of two wheeled vehicles, JSME 28 (240), 1233-1239, 1985.
- [5] Mitsuhashi, Y., Takeshita, H., Momiyama, Y., Yabe, N., Study on Motorcycle Rider Model using Reinforcement Learning—Basic Research to Represent the Behavior according to the Rider Proficiency—, JSME Technical Paper, 559-564, Vol. 55, No. 3, 2024.

当論文は、「A Study of Path Prediction Algorithm for Camera-Based Rider Assistance System」として、ifz (Institut für Zweiradsicherheit) 主催 第15回 IMC (International Motorcycle Conference) にて発表されたものです。

Reprinted with permission Copyright © 2024 ifz (Further use or distribution is not permitted without permission from ifz.)

## 要旨

近年、二輪車向け運転支援システムの研究・開発が進められている。我々は先行車への接近に対しライダーへ注意を促す、カメラを用いたシステムの研究に取り組んでいる。研究中のシステムは物体検出機能、自車と周辺車両の進路予測機能、対象選択機能、出力判定機能で構成する。四輪車と異なる二輪車特有の車両挙動に対応するため、自車と周辺車両の進路予測アルゴリズムを提案した。しかし、これまでにはシンプルなシーンに限定した効果確認しかできていない。そこで、本稿ではアルゴリズムの説明と、シーンを拡充し提案したアルゴリズムの性能評価を行う。

アルゴリズムは自車の進路予測機能、周辺車両の進路予測機能、対象選択機能で構成する。自車の進路予測機能はカメラ画像上で自車の進路を予測する。周辺車両の進路予測機能はカメラ画像上で周辺車両の進路を予測する。対象選択機能は自車の予測進路と周辺車両の予測進路に基づいて注視すべき対象を選択する。

性能評価は、アルゴリズムを組み込みカメラシステムに実装し、外部から入力するテストコース走行データに対し、机上で実施した。評価の結果、直進中や旋回中に先行車両を対象選択できていることが確認された。ただし、車線変更中は先行車両を対象選択できないタイミングもあることが確認された。今後、アルゴリズムを実環境に適用する場合には、動きが複雑な場面や複数車両と混走する場面での処理改善が課題となる。

## Abstract

In recent years, there has been a growing focus on developing rider assistance systems for powered two-wheeler (PTW). We have been studying a camera-based system that informs rider when a preceding vehicle approaches. The system under development consists of an object detection function, a path prediction function for the ego vehicle and other vehicles, and an output judgment function. To address the unique vehicle dynamics of PTW compared to four-wheelers, we proposed an algorithm for predicting the paths of both the ego vehicle and other vehicles. However, thus far, we have only been able to confirm the effectiveness of this algorithm in limited, simplistic scenarios. Therefore, this paper presents the proposed algorithm and evaluates its performance in more varied scenes.

The proposed algorithm comprises functions of the ego vehicle's path prediction, the other vehicle's path prediction, and a target selection. The ego vehicle's path prediction function and other vehicle's path prediction function predict and maintain the paths of the ego vehicle and other vehicle as seen in the camera image. The target selection function selects targets to be observed based on the predicted path of the ego vehicle and other vehicles.

We conducted performance evaluation in a lab test environment, using an embedded camera system with the algorithm implemented and recorded data from a test track with inputs from external sources. The evaluation results revealed the system's ability to select the preceding vehicle as a target during straight and turning maneuvers. However, it also identified instances where the preceding vehicle could not be selected as a target during lane changes. In the coming years, improving algorithmic processing in more complex scenarios, including those with multiple vehicles, will be crucial for real-world deployment.

## 1 INTRODUCTION

The development and adoption of driving assistance systems, similar to those used in four-wheel vehicles, are now expanding to PTW<sup>[1]</sup>. Among rider assistance systems, camera-based systems feature object recognition and have the potential to respond more accurately to traffic conditions. Consequently, the present research focuses on a motorcycle riding assistance system that uses a camera to inform the rider when the preceding vehicle is close<sup>[2]</sup>. The behavior of motorcycles differs from that of four-wheel vehicles, which must be taken into account during function development. One such behavior is banking when cornering, which results in the preceding vehicle appearing inclined in the camera image. Accordingly, a function that appropriately detects the preceding vehicle based on the bank angle has been envisioned. Another motorcycle-specific behavior is the high degree of flexibility with respect to the driving trajectory within a lane. Therefore, relative vehicle positions based on lane markings cannot be used to distinguish the relevant preceding vehicle from other vehicles. This has led to envisioning a function that identifies oncoming or other non-preceding vehicles based on the behavior of other vehicles. Both the ego vehicle's and other vehicle's path prediction algorithms required to realize those functions have already been proposed<sup>[3]</sup>. This study examines whether the proposed algorithms select the appropriate object and avoid incorrect selections in straight-line, cornering, and lane-changing scenarios.

## 2 FUNCTIONAL OVERVIEW

This section provides an overview of the path prediction algorithm functionality<sup>[3]</sup>.

### 2-1. System functional structure

Figure 1 shows the functional structure of the rider assistance system used in this study. The system takes camera images, as well as acceleration, angular velocity, and vehicle speed as inputs, and uses a buzzer sound and a flashing LED as outputs. The functional structure of the

system consists of object detection, ego vehicle path prediction, target selection, time-to-collision (TTC) estimation, and human machine interface (HMI) output control. Of the potential targets detected in the camera images, only the one selected as a focus based on path prediction is subject to TTC estimation and output decision. This paper describes both the ego and other vehicle path prediction algorithm functions, as well as the target selection function.

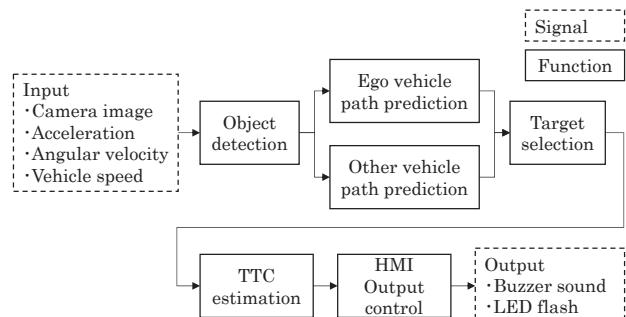


Fig. 1 System functional structure

### 2-2. Ego vehicle path prediction function

This function for the ego vehicle takes angular velocity and vehicle speed as inputs, calculates the turning radius, and then predicts the path of the ego vehicle on the camera image. Figure 2 shows the predicted path of the ego vehicle as it follows the preceding vehicle and goes by an oncoming vehicle. Using the estimated roll angle to correct the path enables appropriate prediction that accounts for the banking of the vehicle body even during cornering.

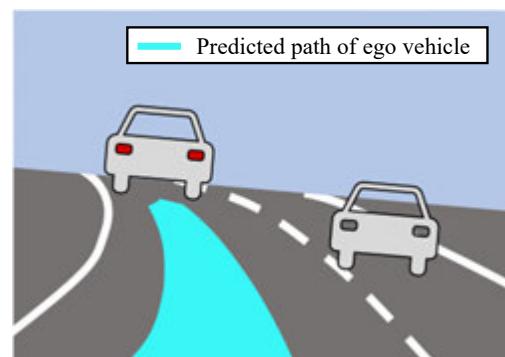


Fig. 2 Predicted Path of Ego Vehicle

## 2-3. Other vehicle path prediction function

This function calculates the movement vector (direction, magnitude) on the image based on the bounding box (BB) coordinates of the detected object and its previous coordinates, and predicts the path of other vehicles on the camera image as shown in Fig. 3. Limiting the input to the BB data makes it possible to predict paths irrespective of white lines, and without depending on the ego vehicle's position.

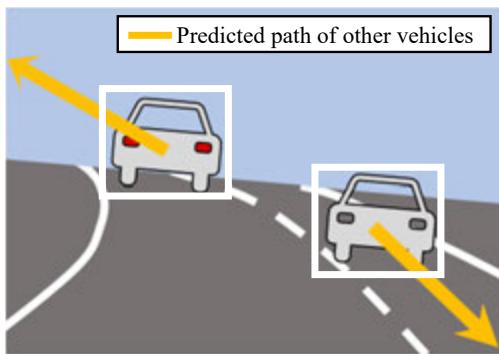


Fig. 3 Predicted Path of Other Vehicles

## 2-4. Target selection function

This function assigns a selection priority to target vehicles along the predicted path in order of proximity to the ego vehicle. It then calculates the probability of a collision and of a change in path based on the predicted path of other vehicles, and excludes vehicles determined to present no risk of collision with the ego vehicle. That process leads to selecting the target to focus on, as shown in Fig. 4.

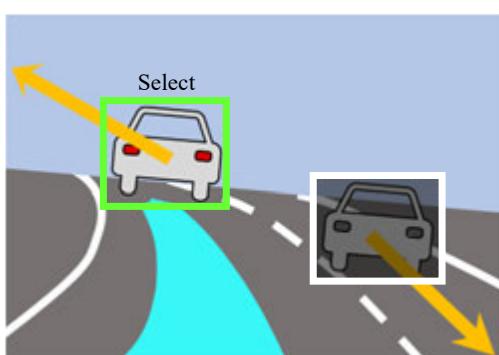


Fig. 4 Selected Vehicle Based on Predicted Path

## 3 THE ALGORITHMS

### 3-1. Ego vehicle path prediction algorithm

This algorithm involves calculating the path, converting it to camera coordinates, and correcting the roll. As shown in Fig. 5, the algorithm establishes a geographic coordinate system using the position of the camera on the ego vehicle as the origin, and setting the course of the ego vehicle  $x$  axis, and the left side of the course parallel to the ground as the  $y$  axis. In addition,  $(x_g, y_g)$  is used to indicate an arbitrary point in the geographic coordinates. As shown in Fig. 6, the algorithm establishes a camera image coordinate system using the top left of the image as the origin, using the width direction as the  $x$  axis and the height direction as the  $y$  axis. In addition,  $(x_s, y_s)$  is used to indicate an arbitrary point in the camera image coordinates. The path of the ego vehicle is calculated in the geographic coordinate system shown in Fig. 5 with the equation below, which uses the  $r[m]$  turning radius obtained from the vehicle speed and yaw rate.

$$y_g = r - \sqrt{r^2 - x_g^2} \quad (1)$$

The following equation is used to convert that path into coordinates in the camera coordinate system from Fig. 6 and calculate the predicted path.

$$x_s = \frac{W_s}{2} - y_g \frac{W_s}{2x_g \tan\left(\frac{\phi}{2}\right)} \quad (2)$$

In this equation,  $W_s[\text{pix}]$  is the screen width, and  $\phi[\text{deg}]$  is the horizontal angle of view. At the same time, the roll angle  $\theta_x[\text{deg}]$  estimated from factors such as the roll rate is used to rotate the coordinates to the center of the screen and correct the predicted path. The coordinates after correction are indicated by  $(x_s', y_s')$ .

$$\begin{pmatrix} x_s' \\ y_s' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \frac{W_s}{2} \\ 0 & 1 & \frac{H_s}{2} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta_x & \sin\theta_x & 0 \\ -\sin\theta_x & \cos\theta_x & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -\frac{W_s}{2} \\ 0 & 1 & -\frac{H_s}{2} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_s \\ y_s \\ 1 \end{pmatrix} \quad (3)$$

In this equation,  $H_s$ [pix] is the screen height. To avoid being unable to capture the preceding vehicle due to excessive variation in the path, the straight line state is determined and the predicted path is fixed to straight ahead of the camera when moving in a straight line. The yaw rate is used to determine the straight line state.

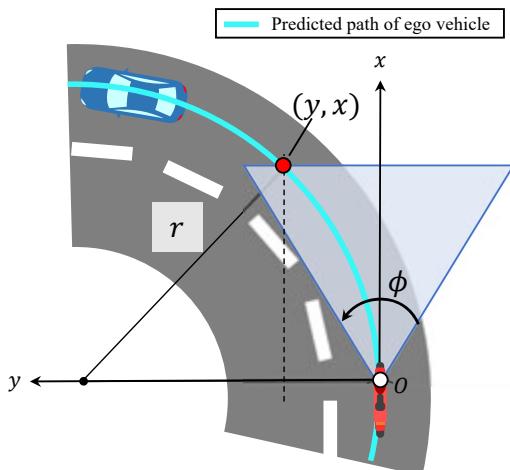


Fig. 5 Top View of Ego Vehicle's Predicted Path

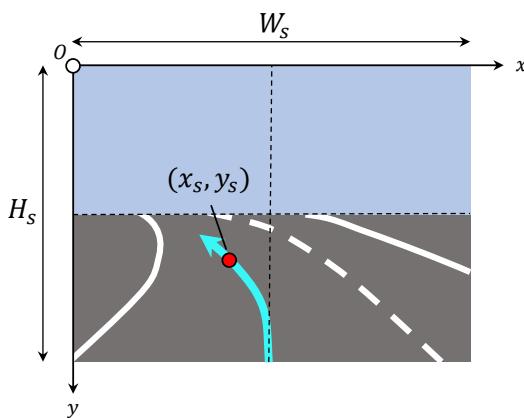


Fig. 6 Camera image of Ego Vehicle's Predicted Path

### 3-2. Other vehicle path prediction algorithm

This algorithm involves calculating the movement vector and deriving path coordinates. As shown in Fig. 7, the direction and magnitude of vehicles are, respectively, obtained by detecting the BB coordinates of the current frame and from the difference with the detected BB coordinates of a given previous frame.

$$\angle \mathbf{A} = \tan^{-1} \left( \frac{y_t - y_{t-L}}{x_t - x_{t-L}} \right) \quad (4)$$

$$|\mathbf{A}| = \sqrt{(x_t - x_{t-L})^2 + (y_t - y_{t-L})^2} \quad (5)$$

In these equations,  $\angle \mathbf{A}$  is the direction of the movement vector,  $|\mathbf{A}|$  is its magnitude,  $L$  is the difference in the number of frames,  $(x_t, y_t)$  represents the coordinates of the center of the base of the detected BB at time  $t$  [frame], and  $(x_{t-L}, y_{t-L})$  represent the coordinates for  $L$  [frame] earlier. In addition, the BB coordinates used to calculate the movement vector cancel out the ego vehicle movement component to extract only the other vehicle movement component. The predicted path of other vehicles is calculated on the premise that it will keep extending in a straight line in the direction of the vector on the screen.

$$y_s = (x_s - x_c) \frac{\sin(\angle \mathbf{A})}{\cos(\angle \mathbf{A})} + y_c \quad (6)$$

In the equation,  $(x_c, y_c)$  represents the center coordinates of the detected BB.

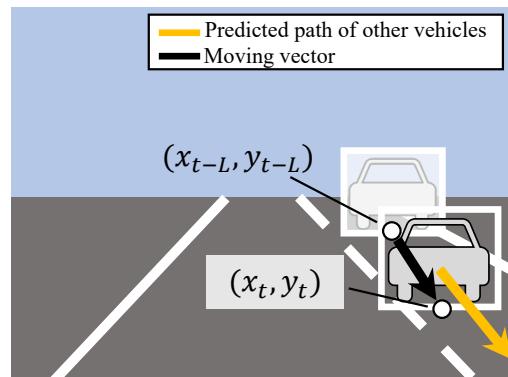


Fig. 7 Movement Vector and Other Vehicle's Predicted Path

### 3-3. Target selection algorithm

This algorithm consists of determining target selection priority based on the predicted path of the ego vehicle, calculating the probability of collision based on the predicted path of other vehicles, calculating the probability of a change in the trajectory of other vehicles, and selecting the target vehicle.

#### 3-3-1. Determining the target selection priority

Figure 8 illustrates how target selection priority is determined. After narrowing down potential targets in the selection range, the target selection is set in order of the smallest to largest values of  $P$  obtained from the equation

below.

$$P = k_P d_y + (1 - k_P) d_x \quad (7)$$

In the equation,  $d_x$  [pix] is the distance in the  $x$  axis direction between the BB base center coordinates and the ego vehicle's path,  $d_y$  [pix] is the distance in the  $y$  axis direction between the BB base center coordinates, and  $k_P$  is a weight coefficient that determines whether to prioritize proximity to the vehicle or proximity to the predicted path. The target selection range is defined as a range consisting of the width of the ego vehicle centered on the predicted path combined with a range extended vertically from the end of the predicted path. Differences in target vehicle sizes are taken into account by considering objects that overlap with the target selection range and the BB base to be within the target selection range.

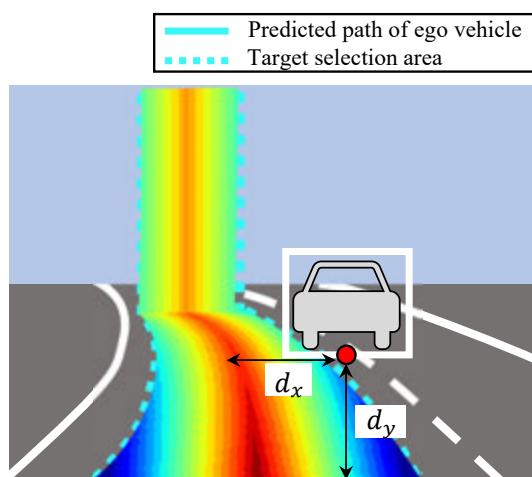


Fig. 8 Target Selection Priority Determination

### 3-3-2. Calculating the collision probability

Figure 9 illustrates how the probability of a collision is calculated. The Gaussian distribution obtained from the predicted path and the integral of the  $W$  ego vehicle width range are used to calculate the collision probability. The average  $\mu$  of the Gaussian distribution represents the  $x$  coordinates at the intersection of the predicted path and the bottom edge of the screen, while the standard deviation  $\sigma$  defines the reciprocal of the magnitude of the movement vector based on the BB center coordinates.

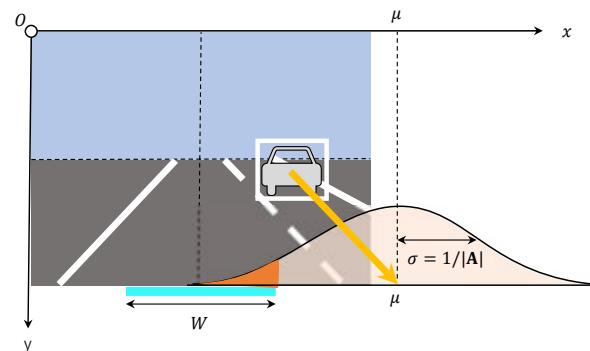


Fig. 9 Collision Probability Calculation

### 3-3-3. Calculating the probability of a change in trajectory

Two probabilities of a change in trajectory are calculated based on the magnitude of the movement vector and on the continuity of its direction. The probability of a change in trajectory based on the magnitude of the movement vector decreases according to the premise that a change in trajectory is unlikely as the magnitude of the vector increases. The probability of a change in trajectory based on the continuity of the direction of the movement vector decreases according to the premise that a change in trajectory is unlikely while the variation in direction remains within a certain range.

### 3-3-4. Selecting the target

Vehicles requiring caution are selected in order of highest priority among the vehicle inside the target selection range. In addition, vehicles with both a collision probability below the threshold and one of the two probabilities of a change in trajectory below the threshold are excluded from the selection targets.

## 4 TESTING

Table 1 shows the items that were validated for the path prediction algorithm.

Table 1 Validation Items

No.	Validation item
A	The correct object is selected as the target vehicle.
B	No incorrect object is selected as the target vehicle.

#### 4-1. Test device

The experiment was conducted using a camera system implementing all of the functions shown in Fig. 1. The device was equipped with a camera and an inertial measurement unit (IMU) and installed on a motorcycle, enabling the real time execution of the proposed algorithm. Individual video frames, angular velocity, and vehicle speed can also be retrieved ahead of time, making it possible to enter that information from an external source and validate it analytically. In this experiment, the validation scenarios were conducted by driving on a closed course and the information retrieved was verified through analysis.

#### 4-2. Test conditions

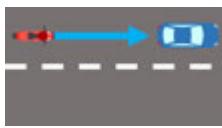
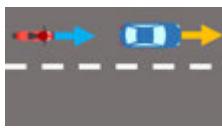
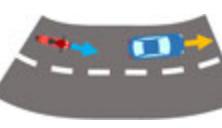
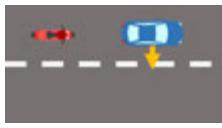
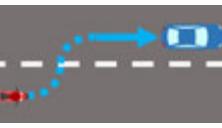
The validation scenarios, validation sections, index, and collection method used as test conditions are described below.

##### 4-2-1. Validation scenarios

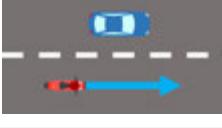
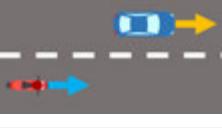
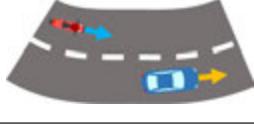
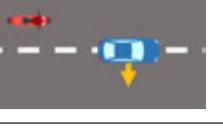
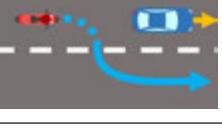
Table 2 shows the validation scenarios for validation item A.

Table 3 shows the validation scenarios for validation item B.

**Table 2 Scenarios to Confirm the Correct Object Is Selected (Validation Item A)**

Scenario No.	A00	A01	A10
Details	Approaching a stationary vehicle ahead	Following the preceding vehicle	Following the preceding vehicle while cornering
Overview			
Scenario No.	A20	A21	A22
Details	The preceding vehicle cuts into the lane	The preceding vehicle moves into the other lane	The ego vehicle cuts into the lane
Overview			

**Table 3 Scenarios to Confirm No Incorrect Object Is Selected (Validation Item B)**

Scenario No.	B00	B01	B02	
Details	Overtaking a stationary vehicle in the other lane	Maintaining distance from the vehicle in the other lane while driving	Moving past a vehicle cornering ahead while driving straight	
Overview				
Scenario No.	B10	B20	B21	B22
Details	Maintaining distance from the vehicle in the other lane while cornering	The preceding vehicle cuts into the lane	The preceding vehicle moves into the other lane	The ego vehicle moves into the other lane
Overview				

#### 4-2-2. Validation sections

The validation sections were defined as a frame in which the ego vehicle drives in accordance with the conditions and the target vehicle is correctly detected. However, in the lane changing scenes, different validation sections were set based on the driving state, as shown in Figs. 10 and 11.

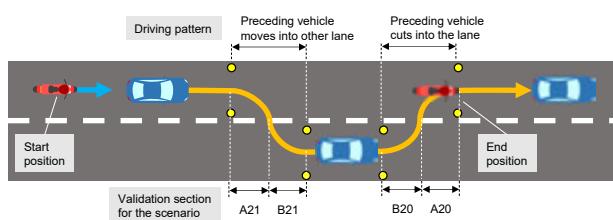


Fig. 10 Validation Section for the Preceding Vehicle Lane Change Scenario

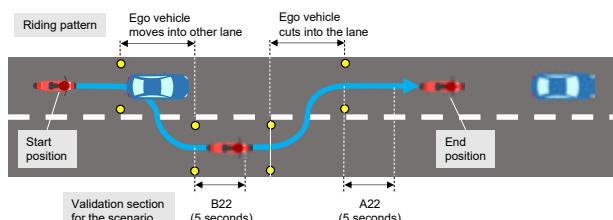


Fig. 11 Validation Section for the Ego Vehicle Lane Change Scenario

#### 4-2-3. Index

The index represents the correct vehicle selection ratio. That ratio consists of the proportion of frames where the algorithm selects the intended frame within a validation item.

#### 4-2-4. Collection method

The correct vehicle selection ratio was calculated using the number of frames in the validation section after counting the number of correct frames. The number of correct frames was counted by determining the correct vehicle selection from the images output by the system. The criterion for correctness depends on the scenario. Table 4 shows decision examples for each validation item. In the validation A scenarios where the vehicle is in the same lane as the ego vehicle, correctness was defined as the ability to select the preceding vehicle in the same lane. In the validation B scenarios where the vehicle is not in the same lane as the ego vehicle, correctness was defined

as the ability to select the vehicle in the other lane. The collection process was designed to retrieve at least 10 seconds of data for each scenario.

Table 4 Examples of Correct Vehicle Selection Decision

No.	Validation item A	Validation item B
Correct	Selected	NOT Selected
Incorrect	NOT Selected	Selected

#### 4-3. Results

This section presents the collected results for the correct vehicle selection ratio. Table 5 shows the collected results for validation item A. The A00 and A01 results confirm that the vehicle is correctly selected when driving in a straight line. While the A10 results show that the correct selection ratio drops during cornering when there is a lot of distance between the vehicles. The A20 scenario also demonstrated a drop in the correct vehicle selection when the preceding vehicle cuts into the lane. In contrast, the A21 scenario showed that when the preceding vehicle moves into the same lane as the ego vehicle, the system can continue to make the selection while that vehicle is still in the same lane as the ego vehicle. In addition, the A22 confirms that the target vehicle is selected correctly after the preceding vehicle has fully entered the ego vehicle's lane.

Table 5 Collected Results for Selecting the Correct (Validation Item A)

Scenario	A00	A01	A10	A20	A21	A22
Vehicle speed	50 km/h					
Distance between vehicles	—	12m	40 m	12 m	40 m	—
Turning radius	—	—	—	425 m	425 m	—
Correct selection ratio	96.7 %	100 %	100 %	98.2 %	66.0 %	56.1 %
					87.8 %	82.9 %

Table 6 shows the collected results for validation item B. The B00 and B01 results confirm that the other vehicle in the other lane is not incorrectly selected when driving in a straight line, while the B02 results confirm that going past an oncoming vehicle in a curve ahead does not incorrectly select that vehicle. In those scenarios, the correct answer cannot be obtained using only ego vehicle information, but using the predicted path of other vehicles made it possible to exclude them from the selection. The B10 results confirm that other vehicles ahead of the ego vehicle are not incorrectly selected during cornering. The B20 results show that the preceding vehicle cutting into the lane is not selected while it is still in the other lane. Similarly, the B21 results show that the preceding vehicle moving into the other lane is not selected after it enters that lane. In addition, the B22 results confirm that the preceding vehicle is not selected when the ego vehicle finishes moving into the other lane.

**Table 6 Collected Results for Not Making an Incorrect Selection (Validation Item B)**

Scenario	B00	B01	B02	B10	B20	B21	B22
Vehicle speed	50 km/h						
Distance between vehicles	—	12 m	40 m	—	40 m	—	—
Turning radius	—	—	—	80 m	425 m	—	—
Correct selection ratio	98.9 %	100 %	100 %	84.9 %	86.0 %	89.1 %	98.6 %
							96.6 %

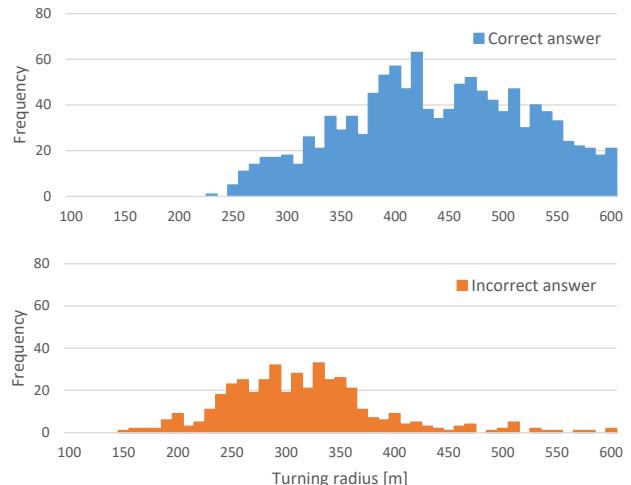
#### 4-4. Observations

This section examines the factors behind incorrect selections and their impact on the system.

##### 4-4-1. Scenario A10: Target selection during cornering

We consider the factors that lead to incorrect answers. Table 5 shows that during cornering, the correct selection ratio is lower at a vehicle distance of 40 m than at 12 m. This is expected to be due to either the ego vehicle path prediction algorithm or to the actual driving during the test. The algorithm was assessed first. The ego vehicle path prediction algorithm estimates the turn and predicts the path based on the current turning radius of

the ego vehicle. Consequently, the further the distance the greater the span of variation, making it more difficult to capture the preceding vehicle in the path of the ego vehicle. The reason for the drop in the correct selection ratio at further distances is thought to be that the larger span of variation in the path of the ego vehicle relative to the variation in the turning radio prevented the capture of the preceding vehicle in the path of the ego vehicle. Next, the driving during the test was examined. Sections with an incorrect selected were observed to generally involve driving at a smaller turning radius than the path. Figure 12 shows the turning radius distribution for correct and incorrect selections for the data from a single trial on a road with a 425 m turning radius and a vehicle distance of 40 m. The graph makes it clear that that the ego vehicle was driving at a turning radius close to 425 m when correct selections were made, but at a turning radius around 250 m to 350 m when incorrect selections were made. The drop in the correct selection ratio is believed to stem from temporarily driving at a smaller turning radius to adjust the driving position, leading to predicting a path inward of that of the preceding vehicle.



**Fig. 12 Turning Radius Distribution for Correct and Incorrect Selections in the Following while Cornering Scenario**

##### 4-4-2. Scenario A20: Preceding vehicle cutting into the lane

The A20 scenario was examined next. As seen in Table 5, this scenario has a poorer correct selection ratio than the other lane change scenarios. That drop is attributed to the

path prediction algorithm for other vehicles. Figure 13 shows the camera images for the A20 scenario. The light blue solid lines represent the predicted path of the ego vehicle, while the orange solid lines represent the predicted path of other. In addition, the green box around the preceding vehicle indicates it has been selected as the target vehicle, and the orange box indicates the other vehicle has been excluded from selection based on its predicted path. The number at the upper left is the frame number.

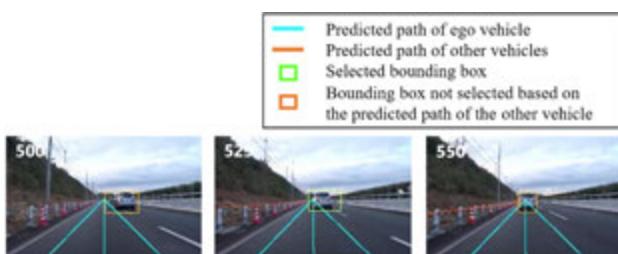


Fig. 13 Target Selection when the Preceding Vehicle Cuts into the Lane Cornering Scenario

The images demonstrate that the other vehicle was excluded from the selection based on its predicted path. The same phenomenon was observed in other data for the same scenario. In the preceding vehicle lane change scenarios, that vehicle only moves toward the side in the image. It seems that the algorithm therefore predicts a path that does not collide with the ego vehicle, and excludes the vehicle from the selection. This makes it a scenario that is not covered by the current algorithm.

Furthermore, the impact of this phenomenon on the system was examined. The durations of the delay in selecting the target were collated to assess the impact of that delay due to the exclusion of the vehicle from the selection. The selection delay was obtained using the A20 test conditions by collating the time from a stable continuous full second until the first frame of the section in which the vehicle was selected after the preceding vehicle entered the ego vehicle's lane. The results are shown in Fig. 14, and indicate that the target vehicle is chosen after a maximum delay of 2.2 seconds. There are concerns that during that selection delay, the system would not react to approaching the preceding vehicle and fail to activate.

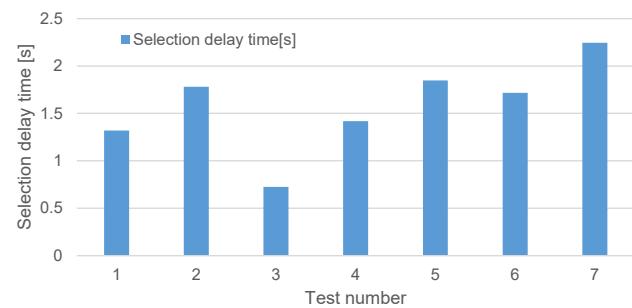


Fig. 14 Duration of Target Selection Delay When the Preceding Vehicle Cuts into the Lane

#### 4-4-3. Scenario A22: Ego vehicle cutting into the lane

The impact on the system of the delay in selecting the target vehicle after changing lanes was examined for the A22 scenario. The selection delay was obtained using the A22 test conditions by collating the time taken for the system to stabilize and select the target vehicle after the ego vehicle finishes its lane change. The results are shown in Fig. 15, and indicate that the target vehicle is chosen after a maximum delay of 1.6 seconds. There are concerns that during that selection delay, the system would not react to approaching the preceding vehicle and fail to activate.



Fig. 15 Duration of Target Selection Delay When the Ego Vehicle Cuts into the Lane

## 5 CONCLUSION

This study analyzed and examined the effectiveness of the ego vehicle and other vehicle path prediction algorithms that account for ego vehicle behavior in a rider assistance system intended for two-wheeled vehicles. The following results were obtained.

- The proposed algorithm was confirmed to select the

correct target object in six of eight scenarios involving straight-line driving, cornering, and lane changes.

- The proposed algorithm was confirmed to avoid selecting an incorrect target object in all eight scenarios involving straight-line driving, cornering, and lane changes.
- In the scenario involving the preceding vehicle cutting into the lane, it was found that the correct selection rate of the preceding vehicle was limited to 56% of the frames. At the same time, incorrect selections were attributed to the other vehicle path prediction algorithm, and the current algorithm was determined to be unable to address that scenario.

To achieve the eventual adoption of the algorithms in rider assistance systems for two-wheeled vehicles, the next step will be to address the issues brought to light in this study and conduct evaluations in situations that resemble real-world usage conditions.

### ■著者



長谷川 翔馬

Shoma Hasegawa

技術・研究本部  
技術開発統括部  
知能化研究部



武田 拓己

Takumi Takeda

技術・研究本部  
技術開発統括部  
知能化研究部



尾上 太郎

Taro Onoue

技術・研究本部  
技術開発統括部  
知能化研究部



品川 晃徳

Akinori Shinagawa

技術・研究本部  
技術開発統括部  
知能化研究部

## REFERENCES

- [1] 2023 TRACER 9 GT+ European Model Exhibited at EICMA - Equipped with the World's First UBS Linked to the Millimeter Wave Radar, Adaptive Cruise Control, and More  
<https://global.yamaha-motor.com/jp/news/2022/1108/tracer9.html>
- [2] DMP: DMP, Launch External Sales of Camera System, the Result of Collaboration with Yamaha Motor, DMP News, <https://www.dmpf.com/en/news/dmp-launch-external-sales-of-camera-system-the-result-of-collaboration-with-yamaha-motor/>
- [3] S. Hasegawa, T. Takeda, T. Onoue, A. Shinagawa: "Path Prediction Algorithm for Motorcycle Rider Assistance System with Camera", Transactions of Society of Automotive Engineers of Japan, Vol. 55, No. 5, pp. 967-971, 2024

# 沖永良部島の脱炭素・持続可能なモビリティ社会実現へ向けた実証研究

Verification Study toward Realizing a Decarbonized and Sustainable Mobility Society on Okinoerabu Island

和田 朋智毅 川口 雅弘 中村 俊之 福田 晋平 稲波 純一 山本 俊行  
桑原 昌広 乾 大樹

本稿は、第72回土木計画学研究発表会・秋季大会(2025年)での発表内容を基にしています。

## 要旨

沖永良部島の知名町および和泊町は、「国立大学法人東海国立大学機構名古屋大学(以下名古屋大学)」、「ヤマハ発動機株式会社(以下当社)」とともに、2023年8月10日に沖永良部島をモデルとした脱炭素・持続可能なモビリティ社会の構築に向けた連携協定を締結した。両町が開示する2018年の部門別CO<sub>2</sub>排出量データによれば、運輸部門の割合が38%と最も大きく、当該部門における脱炭素推進の重要性は高い。効果的な排出量削減には、自動車の移動実態に基づいた施策提案が不可欠である。そこで、名古屋大学の情報通信技術( ICT )を活用し、自動車の移動実態調査を実施した。調査により、運輸部門からのCO<sub>2</sub>排出量を、自動車種別ごとの移動実態に基づくデータとして提示することが可能となった。また、火力発電が主な電力供給源である沖永良部島において、電動化による排出削減効果と、発電源を再生可能エネルギーに置き換えた場合の効果を比較した。さらに、島内のモビリティ課題を広く把握し、脱炭素・持続可能なモビリティ社会の実現に向けた戦略について議論を行い、提案をまとめた。

## Abstract

China Town and Wadomari Town on Okinoerabu Island concluded a partnership agreement on August 10, 2023, with Nagoya University, part of the Tokai National Higher Education and Research System (hereafter "Nagoya University"), and Yamaha Motor Co., Ltd. (hereafter "the Company"), aimed at building a decarbonized and sustainable mobility society using Okinoerabu Island as a model.

According to 2018 sectoral CO<sub>2</sub> emission data disclosed by the two towns, the transportation sector accounts for the largest share, at 38%, underscoring the importance of advancing decarbonization in this field. To achieve effective emission reductions, policy proposals based on actual vehicle usage patterns are essential. Therefore, a vehicle movement survey was conducted using Nagoya University's Information and Communication Technology (ICT).

The survey made it possible to present CO<sub>2</sub> emissions from the transportation sector as data categorized by vehicle type and actual driving behavior. In addition, on Okinoerabu Island—where thermal power generation is currently the main source of electricity—the project compared the emission reduction effects of vehicle electrification under the current conditions with those that would result if renewable energy were introduced as the power source.

Furthermore, by identifying the island's mobility challenges, the partners held discussions on strategies for realizing a decarbonized and sustainable mobility society, and compiled a set of proposals based on these findings.

## 1 はじめに

日本のCO<sub>2</sub>排出量のうち約2割は運輸部門によるものであり<sup>[1]</sup>、輸送機器を製造する企業にとって排出量削減につながる活動を推進する社会的責任は大きい。ヤマハ発動機グループは“環境計画2050”において、モビリティの小型化および電動化を脱炭素に向けた戦略として提示している<sup>[2]</sup>。一方、電動モビリティの市場拡大は限定的であり、四輪市場においても電動

化の進展が鈍化しているとの指摘がある。電動化が進まない背景にある課題を把握することは、今後の戦略立案において重要なである。

本研究では、脱炭素およびモビリティ課題の調査対象として、鹿児島県の奄美群島の一つである沖永良部島を選定した。沖永良部島は以下の3点において研究フィールドとして適している。第一に、2022年4月に環境省より第1回脱炭素先行地域に指定されており、島内の知名町および和泊町は先進的な脱炭

素活動に積極的である。取り組みの一例として、2022年には高校生向けに“EVバイク貸出による町内実証実験”が実施された。島内高校生の約9割が通学に原動機付自転車を利用していることが背景にある。実証実験に「E-Vino」を活用したことが、当社と両町との関係構築につながった。第二に、離島であるため調査対象エリア内外との自動車の往来がなく、調査範囲を明確に定義できる。第三に、両町の課題と研究内容が一致している点である。両町が開示する2018年の部門別CO<sub>2</sub>排出量データによれば、運輸部門の割合が38%と最も高く<sup>[3][4]</sup>、当該部門における脱炭素推進の重要性は高い。効果的な排出量削減には、自動車の移動実態に基づいた施策提案が不可欠である。また、現状の排出量データは、鹿児島県全体の運輸部門からの排出量を、島内の車種別車両保有台数割合で按分した値が用いられている。島内の移動実態を調査することで、より実態に即した排出量の算出が可能となる。

以上の背景を踏まえ、本研究は“沖永良部島をモデルとした脱炭素・持続可能なモビリティ社会の構築”を目的とし、以下2点を目標とした。

1. 島内の自動車種別ごとの移動実態を調査し、CO<sub>2</sub>排出量の現状をより詳細に把握する
2. 島内のモビリティ課題および電動化に関する課題を把握し、脱炭素・持続可能なモビリティ社会に向けた提言をまとめる

本活動は公益的価値が高く、ICTの活用により学術的な知見が求められるものであった。このため、知名町・和泊町・名古屋大学・当社の4者による連携協定の下、推進体制を構築した(図1)。名古屋大学は、Wi-Fiパケットセンサを活用した人流把握(2019)<sup>[5]</sup>、モバイル空間統計データを活用した人口流動調査(2020)<sup>[6]</sup>など、ICT活用実績を有している。



図1 連携協定締結の様子

## 2 沖永良部島の基礎情報

沖永良部島は、鹿児島県大島郡に属する奄美群島の南西部

に位置する。隆起サンゴ礁から成る島で、西部中央には標高240mの大山があるが、それ以外の土地は緩やかな丘陵地である。島の総面積は93.6km<sup>2</sup>であり、沿岸を走る道路を一周すると約55kmとなる(図2)。島の一部は奄美群島国立公園に指定されており、迫力ある断崖絶壁が続く海岸線や大規模な鍾乳洞群など、豊かな自然が広がっている。

島内には知名町と和泊町の2町があり、2025年5月時点の人口は知名町が5,305人、和泊町が5,905人である。人口は減少傾向にあり、就職や大学進学による島外流出が顕著である。

島の基幹産業は農業であり、さとうきび、じゃがいも、花卉、繁殖牛による子牛生産が主要な産物である。

島外との交通は空路と海路の両方が利用可能である。空路は沖永良部空港から鹿児島へ1日3便、徳之島および沖縄へは1日1便が運航されている。海路は鹿児島港から奄美群島を経由し那覇港に向かうフェリーが毎日往復している。島内交通は自家用車による移動が大半を占める。バスも複数路線があり島内を巡ることは可能であるが、主要路線でも1日9本と便数は少ない。島内に1校ある鹿児島県立沖永良部高校の生徒の約9割が原動機付自転車で通学している。

島内の電力は、知名町に所在する九州電力送配電の火力発電所から供給されている。再生可能エネルギーの導入も一部で進められているが、後述する離島特有の課題により、大幅な拡大には至っていない。



図2 沖永良部島地図

## 3 移動実態と運輸部門からのCO<sub>2</sub>排出量の把握

### 3-1. 島内のモビリティの分類とその調査方法

本研究では、島内のモビリティを移動実態が異なるパターンになると推察される8種に分類し、種別ごとに車両台数および移動実態を把握し、それらの情報に基づいてCO<sub>2</sub>排出量の調査を実施した。

分類項目は、“軽自動車”、“小型・普通自動車”、“原動機付自転車”、“レンタカー”、“物流トラック”、“タクシー”、“路線バス”、“運転代行”とした。なお、小型および普通自動車と軽自動車については、貨物と乗用の区別は行っていない。

台数が多い軽自動車および小型・普通自動車のCO<sub>2</sub>排出量の把握には、車両に設置したBluetooth® Low Energy(以下BLE)タグが発するパケットを活用し、移動量を把握する方式を採用した。この方式は、車両側にはBLEタグを設置するのみで済むため、サンプル数の確保が容易である。また、本方式では移動量だけでなく移動軌跡や移動時間帯などのデータを併せて入手することが可能である。CO<sub>2</sub>削減にこれらのデータを活用することも視野に方式を選定した。原動機付自転車については、BLEタグの防水性など設置に制約があったため、軽自動車の移動実態データを基に算出した。日本自動車工業会の調査結果<sup>[7][8]</sup>には、軽自動車および原動機付自転車の年間平均走行距離が示されており、それらの比率と本調査に基づく軽自動車の年間平均走行距離を用いて算出した。レンタカーについては、島外からの来訪者による仕事・観光利用の移動実態を把握するため、GPSロガーを設置して計測したデータを活用した。台数が少ない路線バス、タクシー、物流トラックについては、事業者が管理する車両運行情報(走行距離、燃料消費量等)を参照した。運転代行については、運転手へのヒアリングに基づき走行距離を算出した(表3-1)。

BLEタグおよびGPSによる移動実態の把握方法の詳細については、次節3-2にて記述する。

表3-1 調査した車両の種別、台数とデータ取得方法

利用者	種別	台数	移動実態データ取得方法
島内住民	軽自動車	7,650	パケットセンサ・BLEタグ設置(計1,088台)
	小型・普通自動車	2,526	
	原動機付自転車	2,226	軽自動車データの換算
島外からの訪問者	レンタカー	68	GPSロガー(延べ385台/123日)
島内事業者	物流トラック	112	事業者から10台の走行距離情報取得
	タクシー	22	事業者から3台の走行距離情報取得
	路線バス	11	事業者から全数の走行距離情報取得
	運転代行	7	事業者から3台の走行距離情報取得

### 3-2. ICTを活用した移動実態の調査方法

軽自動車および小型・普通自動車の移動実態調査においては、ICTによる流動把握手法の一つであるパケットセンサを活用した(パケットセンサ技術の詳細は西田<sup>[9]</sup>を参照)。パソコンやスマートフォンなどの電子機器が発信するプローブリクエストを捕捉することで、位置情報をもとに流動を把握する仕組みである。プローブリクエストを捕捉するためには専用機器の設置が必要となるが、パケットセンサはリアルタイムでの観測

が可能である。なお、スマートフォンから発信されるプローブリクエストであるMACアドレスは、ランダマイズ化の影響を受ける。ランダマイズ化とは、セキュリティおよびプライバシー保護の強化を目的としてMACアドレスがランダムに発信される事象である。したがって、スマートフォンを利用した調査では抽出サンプルが断片的な流動となることが課題である。そこで、今回の調査では、ランダマイズの影響を受けないBLEタグを利用した。

パケットセンサは島内35箇所に設置した。主な設置箇所は空港、港、町役場などの主要施設に加え、移動実態に基づくCO<sub>2</sub>排出量算出を目的として、主要道路上に設置した。特に、島民が自動車で移動する際に分岐が生じる地点を中心に配置した。センサの設置にあたっては常時電源が必要であり、店舗や自動販売機の電源を用いた(図3-1)。



図3-1 パケットセンサの設置箇所・構成と設置状況

軽自動車および小型・普通自動車の移動実態を把握するためには、BLEタグを町民に配布する必要がある。BLEタグは、沖永良部島で開催される大規模イベント(港まつり、夏祭り、体育大会)にて配布した。知名町および和泊町で開催されるこれらのイベントには、島民の多くが自動車で来場する。島民の約9%に相当する1,088個のタグを島民の車両に設置した(図3-2)。居住地により移動実態が異なることが想定されたため、CO<sub>2</sub>排出量の算出にあたっては、年代、性別、居住地(字単位)、車種を聞き取り、センサ間で取得されたデータに対して距離計算を行い、車籍地の車両台数に応じて拡大処理を施した。



図3-2 乗用車への BLE タグの設置と配布の様子

BLE タグからリアルタイムに収集されたデータに基づく移動実態の一例を図3-3に示す。BLE タグを設置した車両がセンサ付近を通過した際にデータが収集され、それらを接続することで車両の移動実態を把握することが可能となる。ただし、センサ間の経路詳細は把握できないため、CO<sub>2</sub>排出量の算出に必要な移動距離については、センサ間の最短経路距離を用いて算出した。



図3-3 BLE タグを用いた自動車の移動軌跡例

次に、レンタカーの移動実態調査について述べる。レンタカーアーの協力のもと、スマートフォンにアプリケーションをインストールした GPS ロガー（図3-4）をレンタカーに設置し、データの取得を行った。GPS ロガーは、貸し出し時に利用者の同意を得たうえで車内に設置したものであり、利用者による操作を必要とせず、常時位置情報を取得する仕組みである。GPS ロガーは、レンタカーの貸出から返却までの間、GPS データを継続的に収集する。

収集されたデータは1秒間隔で取得され、車両 ID、緯度・経度、時刻、GPS 精度、速度、機器の充電残量で構成される。

収集されたレンタカーの移動軌跡は1秒ごとに記録されるため、図3-5に示すように点列データを道路にマッチングすることで、1日1台ごとの移動距離を算出することが可能である。この移動距離をもとに CO<sub>2</sub> 排出量を算出した。



図3-4 レンタカーデータ収集の機器

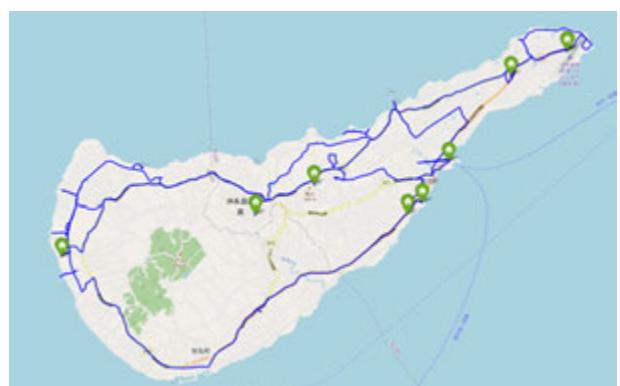


図3-5 レンタカーデータ収集の例

### 3-3. ICT を活用した移動実態の基礎分析結果

3-2節の方法により収集したデータを用いた、移動実態に関する基礎分析結果の一部を報告する。

図3-6は、島民に配布した BLE タグが島内のセンサでどの程度捕捉されたかを、2025年8月の日別で示したものである。この図では、1台の車両（1つの BLE タグ）が複数回カウントされても1回として集計している。すなわち、各日において、どれだけの車両が島内を移動し、センサで捕捉されたかを示す数値である。8月3日に1,000台を超えて捕捉されているのは、この日に夏まつりが開催され、BLE タグを配布したことが影響している。

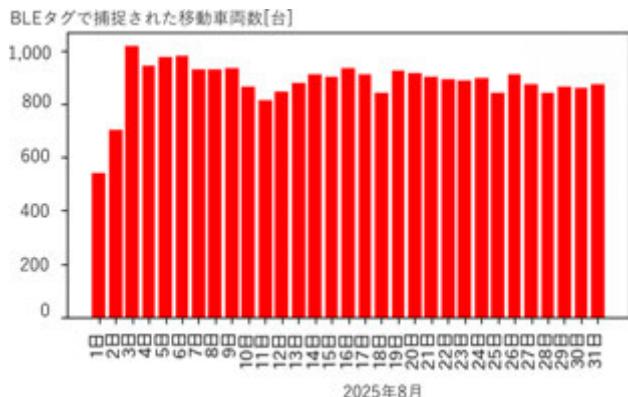
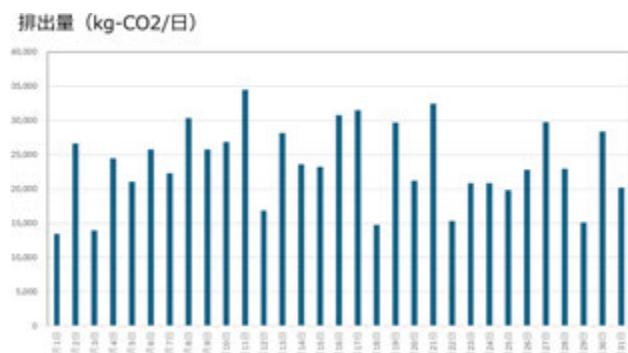


図3-6 島民の車両が日別にどれくらい動いているのか

各車両のセンサ間移動距離の算出結果と、CO<sub>2</sub>排出原単位2.29kg/L<sup>[10]</sup>、燃費12.9km/L<sup>[11]</sup>を用いてCO<sub>2</sub>排出量を算出した。これに、BLEタグ配布時に把握した居住字に基づく車籍地から拡大した車両台数を乗じて、CO<sub>2</sub>排出量を求めたものが図3-7である。本事例では、2025年8月の日別で算出している。日によって排出量が異なるのは、稼働している車両の有無や走行ルートの違いが調査結果に反映されているためである。

図3-7 BLEタグに基づく乗用車の日別CO<sub>2</sub>排出量

レンタカーの移動実態の調査結果を図3-8に示す。これは、1日1台あたりの移動距離の分布を示したものである。多くのレンタカーは、1日の移動距離が100km以下である。沖永良部島内を1周走行した場合でも約50kmであることから、すべての利用者が島内を1周しているとは考えにくく、多くの利用者が50km以内に留まっていることが確認された。また、図3-9は滞在地点のヒートマップであり、和泊町および知名町の市街地を中心に分布していることが確認された。

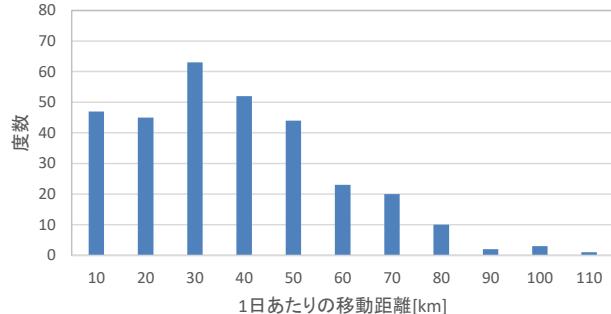


図3-8 レンタカー1日あたりの移動距離分布

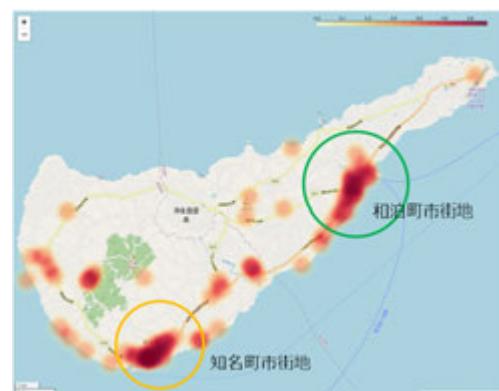


図3-9 レンタカー利用者の滞在地点分布

### 3-4. ICT以外での移動実態の調査結果

各種別の1台当たり年間平均走行距離を表3-2に示す。軽自動車、小型・普通自動車、レンタカーについては、3-2節で示したICTによる調査結果に基づき、2024年8月から11月の1台当たりの平均的な移動距離から、年間平均走行距離に換算して算出した。

表3-2 種別ごとの1台当たり平均年間走行距離

種別	年間平均走行距離 [km/(台・年)]	年間総走行距離 [km/年]
軽自動車	11,013	84,248,991
小型・普通自動車	14,922	37,692,341
原動機付自転車	4,380	9,749,880
レンタカー	25,877	1,759,636
物流トラック	5,850	655,200
タクシー	24,860	546,920
路線バス	32,386	356,246
運転代行	41,229	288,603

年間総走行距離は、年間平均走行距離に各種別の車両台数を乗じて算出した。年間平均走行距離が最も大きいのは運転代行であるが、台数が少ないため、年間総走行距離は最も短くなっている。島内住民の生活の足となっている軽自動車、小型・

普通自動車、原動機付自転車は、その他の種別と比較して車両台数が非常に多く、年間総走行距離が長くなる傾向にある。

### 3-5. 移動量を基にしたCO<sub>2</sub>排出量のまとめ

CO<sub>2</sub>排出量の算出にあたり、以下の3項目を用いた。各項目の算出方法については朴ら<sup>[12]</sup>を参照とした。

#### 1. Tank to Wheel(以下 TtoW)

内燃機関車両(以下 ICE)が走行する際に、車両から直接排出されるCO<sub>2</sub>の排出量である。算出式は以下のとおりである。

$$M_{\text{TtoW}} = \lambda L_y F_e$$

ここで、 $M_{\text{TtoW}}$  は TtoW[t/年]、 $\lambda$  は排出係数 [t-CO<sub>2</sub>/L]、 $L_y$  は年間走行距離 [km/年]、 $F_e$  は燃費 [L/km] である。

#### 2. Well to Wheel(以下 WtoW)

ICE の場合は、TtoW に加え、燃料の採掘・製造・輸送・供給に伴うCO<sub>2</sub>排出量を加えたものである。算出式は以下のとおりである。

$$M_{\text{WtoW}} = \eta M_{\text{TtoW}}$$

ここで、 $M_{\text{WtoW}}$  は WtoW[t/年]、 $\eta$  は WTT 係数 [t-CO<sub>2</sub>/t-CO<sub>2</sub>] である。WTT 係数は、燃料種別や地域ごとに設定された数値である。

電動車両の場合は、走行時に消費した電力の発電・送電に伴うCO<sub>2</sub>排出量である。

$$M_{\text{WtoW}} = \varepsilon L_y E$$

$\varepsilon$  は電力 CO<sub>2</sub>排出係数 [t-CO<sub>2</sub>/kWh]、 $L_y$  は年間走行距離 [km/年]、 $E$  は走行時電力消費量 [kWh/km] である。

#### 3. Life Cycle Assessment(以下 LCA)

LCA は TtoW に加え、車両の原料調達、製造、廃棄、リサイクルまでのライフサイクル全体における CO<sub>2</sub>排出量である。ただし、朴ら<sup>[12]</sup>にも示されるように、廃棄時の排出量は総量に対して微量であるため、ここでは製造時の排出量のみを考慮する。算出式は以下のとおりである。

$$M_{\text{LCA}} = M_{\text{WtoW}} + \frac{M_{\text{manu}}}{Y}$$

$M_{\text{LCA}}$  は LCA[t/年]、 $M_{\text{manu}}$  は車両製造時の CO<sub>2</sub>排出量 [t]、 $Y$  は車両の想定寿命年数 [年] である。

上述の方法により算出した ICE の CO<sub>2</sub>排出量の結果を表3-3

に示す。沖永良部島における電動車両の保有比率は、自動車の7割以上を占める軽自動車でさえ1%未満であり、本結果は概ね島内の運輸部門からの CO<sub>2</sub>排出量と捉えることができる。島内の全車両による1年間の CO<sub>2</sub>排出量は、TtoW で 21,399t となつた。これは、2020年に各町役場が試算した数値<sup>[3][4]</sup>の合計である 20,773t と比較して大きな差異はない。

表3-3 ICE 車両からの年間 CO<sub>2</sub>排出量

種別	ICE		
	TtoW[t-CO <sub>2</sub> /年]	WtoW[t-CO <sub>2</sub> /年]	LCA[t-CO <sub>2</sub> /年]
軽自動車	13,511	15,875	18,571
小型・普通自動車	6,691	7,862	8,666
原動機付自転車	373	439	754
レンタカー	116	137	153
物流トラック	415	447	501
タクシー	97	114	123
路線バス	150	163	168
運転代行	46	54	59
合計	21,399	25,091	28,996

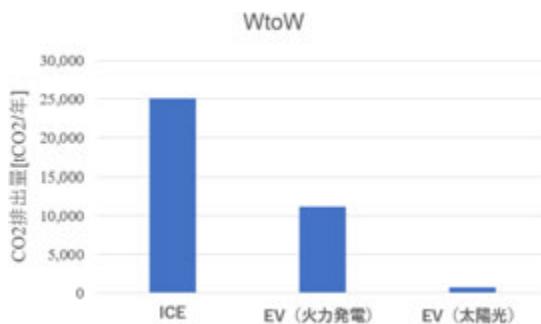
## 4 考察、脱炭素へ向けた提言

### 4-1. モビリティ電動化によるCO<sub>2</sub>削減効果

島内のモビリティを電動化した場合の CO<sub>2</sub>削減効果を表4に示す。なお、算出に際しては、島内の発電方法を100%火力発電として取り扱った。実際には、島内電力の約15%が再生可能エネルギーで賄われており、そのほとんどは太陽光発電である<sup>[13][14]</sup>。島内の主な産業が一次産業であり、車両の充電が帰宅後の夜間に行われることを想定すると、充電電力を太陽光発電で賄うことは困難である。したがって、本節では火力発電による電力供給を前提として算出した。電動化により、CO<sub>2</sub>排出量は WtoW で 13,917t 減少(ICE 比 -55%)、LCA で 10,948t 減少(ICE 比 -38%) することが確認された。

表4 モビリティを電動化した場合の年間 CO<sub>2</sub>排出量

種別	EV	
	WtoW[t-CO <sub>2</sub> /年]	LCA[t-CO <sub>2</sub> /年]
軽自動車	6,749	11,804
小型・普通自動車	3,376	4,884
原動機付自転車	225	344
レンタカー	59	98
物流トラック	443	559
タクシー	49	65
路線バス	250	262
運転代行	23	32
合計	11,174	18,048
ICEとの差	-13,917	-10,948

図4-1 電動化・再エネ化によるCO<sub>2</sub>削減効果

さらに、島内電力を太陽光発電に置き換えた場合のCO<sub>2</sub>排出量をWtoWで算出した。その結果を図4-1に示す。電動車両のTtoWはゼロであるが、消費する電力の発電および送電によりCO<sub>2</sub>が排出される。したがって、電動化によるCO<sub>2</sub>排出量削減効果をTtoWで算出すると過大評価となるため、本節ではWtoWにより島内の排出量削減効果を評価する。ICEと比較すると、単純なモビリティの電動化でもCO<sub>2</sub>の排出量を大きく削減できることが確認された。しかし、同じ電動化であっても、火力発電による電力供給では排出量が多くなる。その効果を最大化するためには、火力発電からの脱却を進め、再生可能エネルギーへの移行が不可欠である。

また、電動化には一般的に航続距離の課題がある。島内は1周約55km、軽自動車の1日当たりの平均走行距離は約30kmであり、自動車による長距離移動の機会が相対的に少ない沖永良部島は、電動化が進みやすい環境にあるといえる。

## 4-2. エネルギーとの連携

前述のとおり、脱炭素の実現にはモビリティの電動化とあわせて、再生可能エネルギーの導入が不可欠である。一方で、実現には大きな課題がある。九州電力株式会社が2014年に実施した、再生可能エネルギーの接続出力が3,224kWの場合のシミュレーション結果によれば、再生可能エネルギーの発電出力が最大となる正午前後の時間帯に、火力発電の最低出力と再生可能エネルギーの発電出力の合計が需要を上回る能性があることが示されている<sup>[13]</sup>。2025年3月末時点での沖永良部島における再生可能エネルギーの接続出力は3,639kWであり<sup>[14]</sup>、すでに電力不安定化のリスクを抱えている状態にある。このため、これ以上の接続は保留せざるを得ない。さらに、電動化が進展した場合には、島内の電力使用量が増加する。すべてのモビリティが電動化された場合、約15,500MWh/年の電力量が必要となる。本調査では、自動車種別ごとの年間平均走行距離に、車種別の走行時電力消費量を乗じて合計し、必要電力量を算出した。これは定常的な発電出力に換算すると約1.8MW

に相当する。仮に、小型・普通自動車および軽自動車の半数が夕方帰宅後に、自宅充電で一般的な3kWの充電を同時に行つた場合、15.3MWの電力が必要となる。これに、2011年の夕方に記録された最大電力13.2MWを加えると、28.5MWとなり、島内火力発電設備の最大出力19.1MWを大きく上回る。

このように、電動化の課題にはエネルギー課題が密接に関係している。これらの課題に対しては、モビリティの視点からアプローチも可能である。その一例が、充電タイミングの制御である。例えば、前述の仮定のように一斉に充電が開始されることを防ぐ方法が挙げられる。図4-2は、軽自動車および小型・普通自動車の移動実態調査から得られた1日の自動車稼働率を示しており、日中でも約9割の車両が停車している。これを踏まえ、再生可能エネルギーの発電量が多い昼間の時間帯に充電を促す方法が有効と考えられる。充電制御システムの導入には多額の投資が必要となるため、外出先での充電設備の整備など、複数の視点から施策を検討する必要がある。

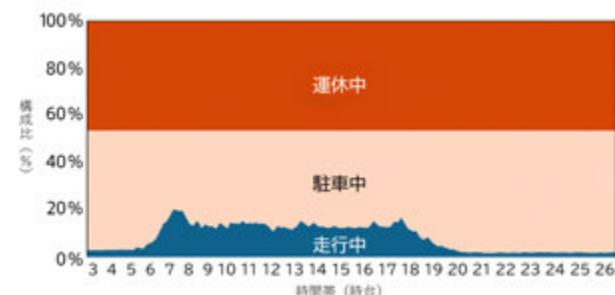


図4-2 1日の自動車稼働率

## 4-3. 沖永良部島のモビリティ社会ビジョン

これまで、定量的なデータに基づいて課題を把握してきた。効果の高い施策や優先度は提示可能であるが、実際に施策を実行するのは島民一人ひとりであり、行動変容を促していくことの重要性は高い。例えば、モビリティの電動化は既に購入補助金や体験会などにより推進されてきたが、普及が進んでいないのが実態である。島民が目指す沖永良部島のモビリティ社会ビジョン(図4-3)を設定することで、行動変容につながる施策をバックキャストにより立案することとした。ビジョンの作成にあたり、島の20代を中心に構成される一般社団法人シマスキと3回のワークショップを実施した。第1回・第2回では島のモビリティ課題の抽出を行い、“農業の活性化”、“高齢者の移動”、“安全”、“自然保護”の4つの重要課題テーマを特定した。第3回では、これらの課題を乗り越えながら実現したいモビリティ社会ビジョンを作成した。単なる課題解決にとどまらず、“夢が広がるエラブのモビリティ社会”と題し、島の自然や産業、人の魅力に寄り添う“ワクワクする”アイディアを盛り込むことをコンセプトとしている。



図4-3 モビリティ社会ビジョン

#### 4-4. 脱炭素へ向けた提言

これまでの研究結果およびモビリティ社会ビジョンを踏まえ、脱炭素に向けた提言を図4-4に示す。“交通全体の課題”、“電動化における課題”、“エネルギー課題”、“脱炭素へ向けた選択肢”の4つの視点から、課題と施策を整理した。

本研究でも試算したとおり電動化と電力の再生可能エネルギー化が提言の主軸である。一方で、モビリティ社会の課題はビジョン検討でも明らかになった高齢ドライバーの増加や利用率の低い公共交通機関など様々な課題が絡んでいる。公共交通の改善により利用率を高めることは、交通課題と脱炭素同時にアプローチすることが可能である。このような相乗効果の高い取り組みは電動化よりも優先的に取り組むべき事項とした。また、住民が楽しみながら各施策に参画できる要素を戦略に組み込んだ。例えば、電動車により電気を持ち運んで使えるようになると、畠仕事の休憩時に冷たい飲み物や氷菓子を楽しむことができる。やってみたくなるアイディアを提案・実践・発信し続けていくことも電動化を推進する上での重要な活動の1つである。電動車両メーカーとして、電動車らしい新しい使い方を提案していくことも重要な役割である。

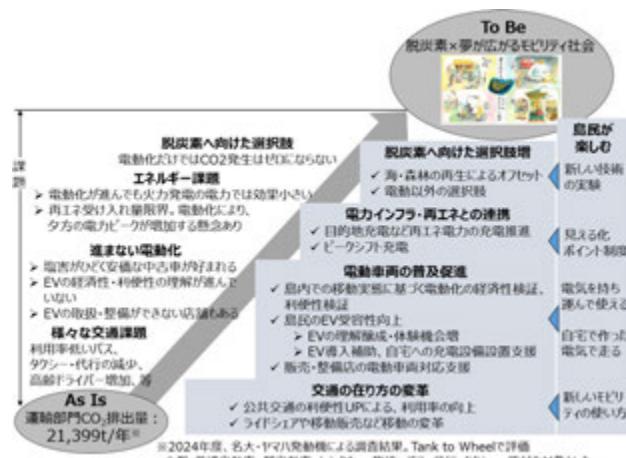


図4-4 脱炭素へ向けた提言

## 5 まとめ

1. ICTの活用により、島内の自動車種別ごとの移動実態に基づいたCO<sub>2</sub>排出量を把握することができた。軽自動車、小型・普通自動車などの自家用車による排出量の割合が最も大きく、施策の優先度が高いことが示された。
2. 移動実態およびCO<sub>2</sub>排出量の把握結果と、島内の交通課題を踏まえ、脱炭素に向けた議論と提言につなげることができた。
3. 移動実態調査ではBLEタグを活用し、町民の移動を把握したが、センサ間のデータ取得間隔が大きい場合の補完処理などに課題があり、今後はその改善を進める必要がある。
4. 移動実態データは、エネルギー利用実態と組み合わせることで、より有効な脱炭素施策の提案が可能となる。本事例の提示により、データ活用を島内で推進していく重要性を示した。

## 6 参考事例: 関連実証と共創活動の紹介

本報告で紹介した取り組みは、沖永良部島の知名町および和泊町と名古屋大学、および当社による2023年8月10日に締結された連携協定に基づき開始されたものである。連携協定は、沖永良部島をモデルとした脱炭素・持続可能なモビリティ社会の構築を目的としている。この取り組みの過程において、名古屋大学が代表を務める「国立研究開発法人科学技術振興機構(JST)」のCOI-NEXT共創の場形成支援プログラム(地域共創分野令和4(2022)年度採択:地域を次世代につなぐマイモビリティ共創拠点)に、名古屋大学、岐阜大学、鹿児島大学、知名町、一般財団法人サステナブル経営推進機構、当社が共同で応募し、採択された。これにより、脱炭素およびモビリティに関する取り組みの拡大を進めている。

ここではモビリティに関する取り組みのうち、電動車両の自動運転を活用した実証について紹介する。名古屋大学が開発した自動運転技術の横展開として、観光地における移動課題の解決と地域の賑わい創出を目的に、観光鍾乳洞「昇竜洞」において電動車両の自動運転実証を令和6(2024)年3月15日から6月9日にかけて実施した。昇竜洞では入口と出口が離れており、観光後に出口から入口まで約600mの緩やかな登り坂を歩いて戻る必要がある。この移動を自動運転車両により代替し、名古屋大学で講習を受けた地元住民がオペレータとなって、週末に運用した。実証では、当社のグリーンスローモビリティ「AR-

05」に名古屋大学の自動運転プログラム「ADENU」を搭載し、自動運転(レベル2)にて約13km/hで昇竜洞出口から入口駐車場まで約600mを約5分で結んだ(図6-1)。



図6-1 沖永良部島観光昇竜洞における自動運転実証

この実証により、グリーンスローモビリティに自動運転技術を搭載することで、観光地の移動課題の解決に寄与することが関係者間で確認された。一方で、移動課題の解決に加え、国立公園の一部である観光昇竜洞周辺の景観との調和を図るなど、さらなる魅力向上の必要性が課題として残る結果となった。

グリーンスローモビリティによる移動体験が昇竜洞観光と一体化した魅力となるよう、島の自然と島民との調和を目指し、沖永良部島高校の生徒と車両・ガレージをデザインした。車両だけでなく、走行空間もデザインし、高校生のコンセプトを活かした沖永良部島専用車両を制作した。さらに、知名町および民間企業による再生エネルギーを活用したガレージの導入も検討した(図6-2)。

令和7(2025)年度も、名古屋大学 COI-NEXT「地域を次世代につなぐマイモビリティ共創拠点」の取り組みは継続して推進される予定である。



図6-2 高校生との車両・ガレージのデザイン

## ■謝辞

本研究の実施に際し、調査にご協力いただいた知名町役場および和泊町役場の皆様に対し、心より御礼を申し上げる。また、ビジョンの共同策定において多大なご助力をいただいた一般社団法人シマスキーの皆様にも、合わせて深く御礼を申し上げる。

## ■参考文献

- [1] 国土交通省開示情報:[https://www.mlit.go.jp/sogoseisaku/environment/sosei\\_environment\\_tk\\_000007.html](https://www.mlit.go.jp/sogoseisaku/environment/sosei_environment_tk_000007.html)
- [2] ヤマハ発動機グループ環境計画2050:<https://global.yamaha-motor.com/jp/profile/csr/environment/plan-2050/>
- [3] 知名町地域再エネ導入戦略策定報告書(2021):<https://www.town.china.lg.jp/kikakushinkou/kurasu/zero/documents/kuikisesakuhen.pdf>
- [4] 和泊町地球温暖化防止実行計画(2024):<https://www.town.wadomari.lg.jp/documents/295/kuikisesakuhen.pdf>
- [5] 大野 沙知子, 中村 俊之, 薄井 智貴, 手嶋 茂晴:人流把握のためのWi-Fiパケットセンサー調査手法に関する研究, 土木学会論文集D3, Vol. 75, No. 5(2019)
- [6] 川上 陸, Jan-Dirk SCHMÖCKER, 宇野 伸宏, 中村 俊之:モバイル空間統計のデータ特性を考慮したOD推計手法:京都観光地間流動におけるケーススタディ, 土木学会論文集D3, Vol. 75, No. 6(2020)
- [7] 一般社団法人日本自動車工業会:軽自動車の使用実態調査報告書(2024)
- [8] 一般社団法人日本自動車工業会:二輪市場動向調査報告書(2024)
- [9] 西田 純二:Wi-Fiパケットセンサ商用化に至る課題克服の歩み—産学官連携が生んだ交通流動解析システム—, デジタルプラクティス, Vol. 11, No. 3(July 2020)
- [10] 環境省 温室効果ガス算定方法及び排出係数一覧:[https://policies.env.go.jp/earth/ghg-santeikohyo/files/calc/itiran\\_2023\\_rev4.pdf](https://policies.env.go.jp/earth/ghg-santeikohyo/files/calc/itiran_2023_rev4.pdf)
- [11] 奥嶋 政嗣, 石井 亜也加:社会的相互作用を考慮したCEV普及シミュレーション, 土木学会論文集D3, Vol. 70, No. 5(2014)
- [12] 朴 秀日, 加藤 博和, 大石 直毅:中大量輸送機関の端末交通としてのパーソナルモビリティツール導入に関する低炭素性評価, 土木学会論文集D3, Vol. 78, No. 5(2023)
- [13] 九州電力株式会社開示情報:<https://www.kyuden.co.jp/library/pdf/press/2014/wo8r9ykne.pdf>
- [14] 九州電力送配電株式会開示情報:<https://www.kyuden.co.jp/td/renewable-energy/remote-island-connection.html>

■著者



和田 朋智毅

Tomoki Wada

技術・研究本部

共創・新ビジネス開発部



川口 雅弘

Masahiro Kawaguchi

技術・研究本部

共創・新ビジネス開発部



中村 俊之

Toshiyuki Nakamura

国立大学法人

東海国立大学機構

岐阜大学



福田 晋平

Shimpei Fukuda

技術・研究本部

共創・新ビジネス開発部



稻波 純一

Junichi Inami

技術・研究本部



山本 俊行

Toshiyuki Yamamoto

国立大学法人

東海国立大学機構

名古屋大学



桑原 昌広

Masahiro Kuwahara

国立大学法人

東海国立大学機構

名古屋大学



乾 大樹

Daiki Inui

国立大学法人

東海国立大学機構

名古屋大学

# [NEDO 助成事業] ポスト 5G 向けチップオンウェハダイレクト接合 3D 積層統合技術開発

Development of Chip-on-Wafer Direct Bonding 3D Stacking Technology for Post-5G

菊地 広 菱沼 隼 李 瑾 永田 憲雅 尾川 謙一

## 要旨

「ヤマハロボティクス株式会社(以下当社または YRC)」は、「ポスト5G向けチップオンウェハダイレクト接合3D積層統合技術開発」を NEDO 助成事業として実施した。本技術の実現に必要な要素技術である、①異物除去洗浄・前処理技術、②半導体チップ非接触ハンドリング技術、③異物検査技術、④接合技術の開発成果について報告する。

## Abstract

“Yamaha Robotics Co., Ltd. (hereafter “the Company” or “YRC”)” carried out the *Development of 3D Integrated Chip-on-Wafer Direct Bonding Technology for Post-5G Applications* as a project subsidized by the New Energy and Industrial Technology Development Organization (NEDO). This report presents the development results of the key technologies necessary to realize this technology: (1) foreign particle removal and surface pre-treatment technology, (2) non-contact handling technology for semiconductor chips, (3) foreign particle inspection technology, and (4) bonding technology.

## 1 実施計画

### 1-1. 事業目的および目標

#### (1) 事業目的

本事業は、ヤマハロボティクス株式会社が提案した「ポスト5G向けチップオンウェハダイレクト接合3D積層統合技術開発」が、国立研究開発法人新エネルギー・産業技術総合開発機構(NEDO)より公募された「ポスト5G情報通信システム基盤強化研究開発事業／先端半導体製造技術の開発」の助成事業として2022年8月に採択され、実施したものである。本事業は、国立研究開発法人産業技術総合研究所、学校法人東京理科大学を共同研究先として実施した。

近年、半導体の性能向上とコスト低減を支えてきた前工程の微細化は限界に近づいており、これに代わる手段として後工程による性能向上が期待されている。特に、3次元実装などのチップ積層化技術は急速に適用が拡大しており、NAND FlashやDRAMなどのメモリ用途に加え、近年ではロジックLSIなどへの適用も始まっている。これに伴い、接合ピッチのさらなる微細化や電極接続抵抗の低減による高速信号伝送性能の向上が求められており、高性能な実装技術の開発ニーズが高まっている。

半導体後工程におけるボンダ装置市場は、今後5年間で年平均成長率3.5%の拡大が予測されている。中でも、先端半導体向けのボンダ市場は、今後5年間にわたり年平均成長率18%で推移すると予測されており、その後も継続的な成長が期待されている。

従来、日本の装置メーカーが優位性を持っていた半導体後

工程設備産業において、近年は欧米やアジアの競合企業が合併等を通じてシェアを拡大している。このような状況下で、日本の技術力を結集し、世界の競合に対抗する高度な技術力の確立が求められている。

そこで当社は、先端装置市場における競争力を維持するため、次世代の実装工法および設備の開発によりグローバル市場をリードしていく。本事業では、ポスト5G社会において活用が期待される先端半導体向けの「チップオンウェハ(以下、COW: Chip on Wafer)」ダイレクト接合技術を早期に確立し、ポスト5G社会の実現に貢献するとともに、半導体後工程設備市場における日本のシェア拡大を図ることを目的とする。

#### (2) 事業目標

近年、先端デバイスでは性能向上のため、バンブピッチの微細化が50μmを下回り、貫通電極を有するチップ厚も50μm以下の極薄LSIチップを8段に積層した高速メモリが量産されている。半導体の微細化が限界に近づく中、今後の社会における“デジタルトランスフォーメーション”を加速するポスト5G向け先端デバイスの実現には、複数のLSIを高度なパッケージ技術で組み合わせ、複雑な機能を実現する必要がある。そのため、パッケージ配線の微細化および実装電極ピッチの狭ピッチ化が不可欠であり、高周波特性のさらなる改善や省電力化の実現も求められている。これらを達成するには、従来のはんだ実装接合に代わり、金属電極間を直接接合する“ダイレクト接合技術”的実用化が必要である(図1、図2、図3参照)<sup>[1][2]</sup>。

ダイレクト接合技術は、既に CMOS センサーなど一部の先端領域においてウェハ同士の貼り合わせによって実現されている。しかし、今後の3次元実装では、ウェハサイズやチップサイズの異なる LSI の積層接合が求められており、ウェハ同士の貼り合わせでは良品チップと不良品チップを接合してしまうなどのコスト面での課題があるため、COW によるダイレクト接合の実現が必要である。

COW によるダイレクト接合において最大の課題は、チップを個別に切断するダイシング工程で発生するパーティクルの除去である。

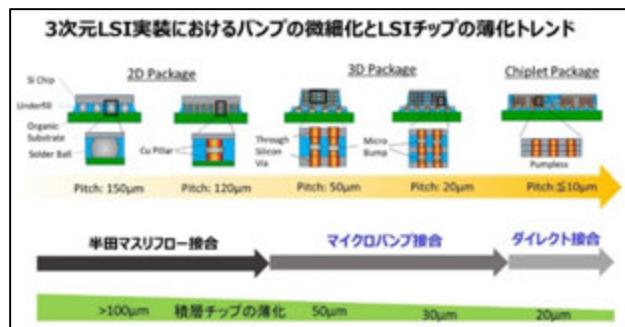


図1 3次元実装技術におけるバンプの微細化と  
LSIチップの薄化トレンド

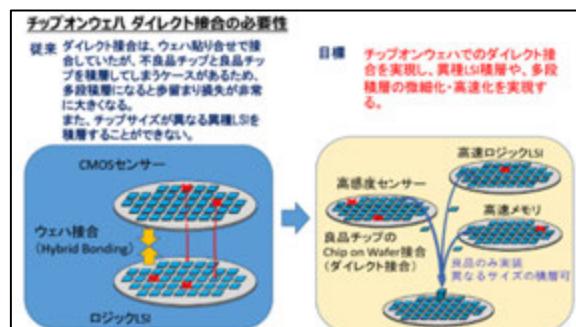


図2 ダイレクト接合技術の必要性

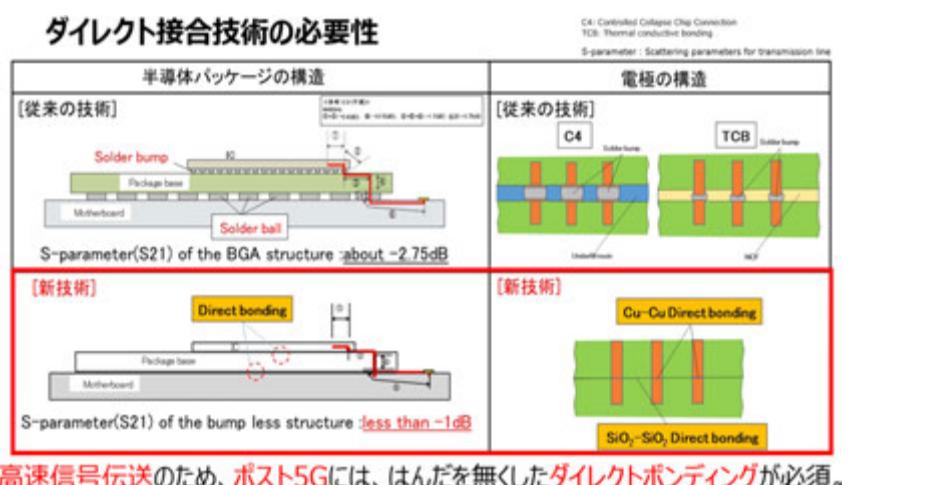


図3 チップオンウェハダイレクト接合技術の必要性

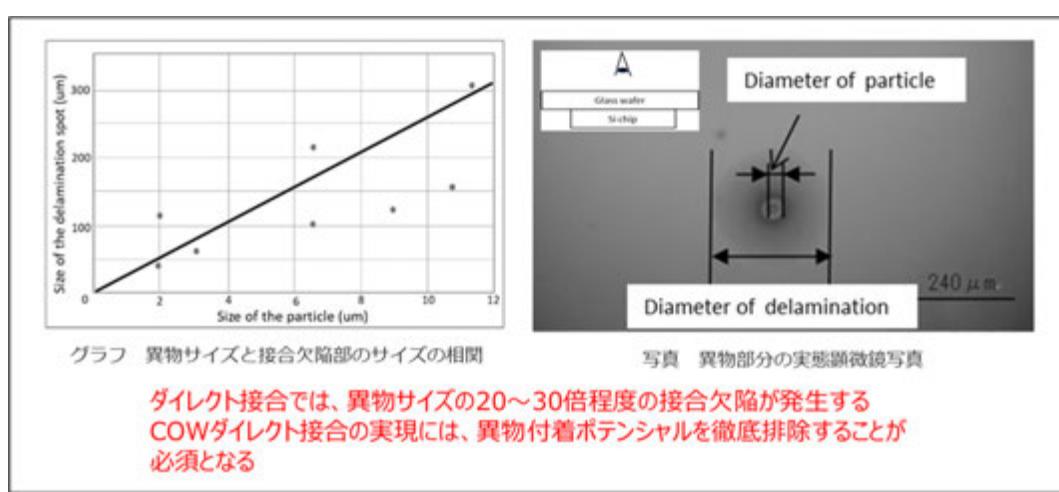


図4 ダイレクト接合への異物の影響

接合表面のパーティクルは接合界面に微小な剥がれを発生させ(図4参照)、電極の接続不良に直結する。このため、極めて高い清浄度が要求される。プラズマダイシングなど発塵の少ない工法の開発も進められているが、発塵を完全に防ぐことは困難であり、接合直前のパーティクル除去が不可欠である<sup>[3]</sup>。

さらに、チップの多段積層ニーズの高まりに伴い、積層対象となるチップの薄化も進行しており(図1参照)、厚さ50μmを下回る極薄チップを損傷なく清浄な状態で実装するためのハンドリング技術も重要な課題となっている。

本事業では、ダイレクト接合技術を用いて、10μm レベルの狭ピッチ電極を有する50μm 以下の極薄チップの3次元実装を量産レベルで実現することを目指し、超清浄なハンドリングおよび前処理、異物除去、半導体チップ上の微小異物検査技術を確立し、COW による3D 積層の統合プロセス技術の構築を目標とした。

高精度なダイレクト接合技術を COW で実現し、その工法を組み込んだ設備の販売を通じて、高性能半導体およびポスト 5G 技術の実現に貢献し、設備販売事業の発展を目指す。

## 1-2. 研究開発の概要

ポスト5G 社会における“デジタルトランスフォーメーション”を加速する先端デバイスの3次元実装を実現するには、複数の半導体チップをダイレクト接合技術により COW で実装する必要がある。しかし、従来技術では、ダイシング後の半導体チップ表面に異物や汚れが残留し、接合品質に悪影響を及ぼすほか、薄化されたチップに損傷を与えるなど、実用化に向けた深刻な障害となっていた。この課題を解決するため、本事業では以下の4つの要素技術を開発した。

- ①異物除去洗浄・前処理技術
- ②半導体チップ非接触ハンドリング技術
- ③異物検査技術
- ④接合技術

これらの研究開発を通じて、厚さ50μm 以下の極薄チップを高い生産性と高歩留で実装可能な、環境配慮型 COW ダイレクト接合技術を構築し、ポスト5G 社会の実現に向けた先端半導体の開発に貢献する。

## 2 COW ダイレクト接合プロセス開発のコンセプト

### 2-1. 異物に関する課題と解決のためのコンセプト

COW ダイレクト接合を実現するための重要な課題として、“異物”に関する技術的障壁と、その解決に向けたコンセプトに

ついて、2項目に分けて説明する。

半導体チップ上に残留する異物の1つ目は、ダイシング後のチップ上面に付着する異物である。

写真1は、ミラーウエハをダイシングし、ダイシングライン近傍を金属顕微鏡で拡大観察した例である。ウエハをダイシングすると、ダイシングラインに沿って異物が付着する。異物は大きく分けて2種類に分類され、1つは Si 屑、もう1つは有機物異物である。

Si 屑は不透明で、Si ウエハに由来するものであり、ウエハ用パーティクルカウンタで検出可能である。

一方、有機物異物は半透明で、ダイシングテープに由来する。この半透明の有機物異物は、ウエハ用パーティクルカウンタでは検出が困難である。

COW ダイレクト接合を実現するには、微小な半透明有機物異物を除去する洗浄・表面処理技術と、これらの異物を検査する技術が必要である。

そこで本事業では、水素水を用いた異物除去技術および異物検査技術の開発を実施した。

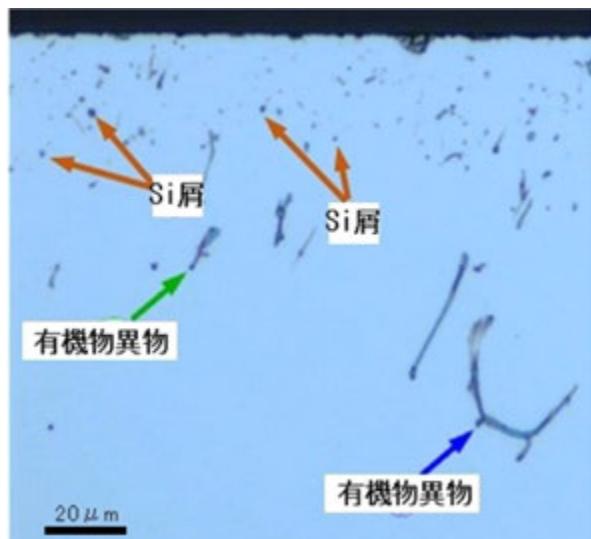
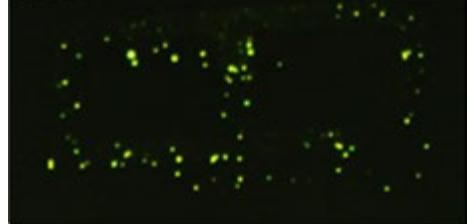
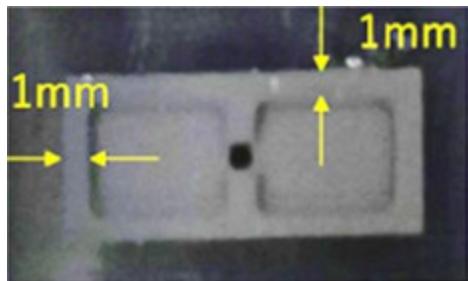


写真1 チップ上の異物観察例

半導体チップ上に残留する異物の二つ目は、チップ表面に治具が接触することにより付着する異物である。

写真2-1は、ポンダの搬送用コレットの外観である。通常、コレットと呼ばれる治具により、半導体チップを真空吸着して装置内を搬送する。

チップがコレットに接触すると、写真2-2のように、コレット形状に沿って異物がチップ表面に付着する。写真2-3は、COW 接合後の接合界面の SAT(Scanning Acoustic Tomography: 超音波映像装置)像である。この状態で COW 接合を行うと、コレット



トの接触痕に沿ってボイドが発生する。

そこで本事業では、この課題を解決するためのコンセプトとして、半導体チップの非接触ハンドリング技術の開発を実施した。

## 2-2. COW ダイレクト接合システムのコンセプト

図5は、プロセスの基本構想と NEDO 助成事業における各事業項目との関係を示す説明図である。

半導体前工程においてウエハが完成した後、真空プラズマ処理により表面を活性化する。その後、ウエハは半導体組立工程へ搬送され、ダイサーによってチップ個片に切断・分割される。個片化されたメモリチップは、COW ダイレクトボンディングシステムにより、異物除去およびプラズマ処理を施した後、COW 接合を行う。

本事業では、前述の4つの要素技術(①異物除去洗浄・前処理技術、②半導体チップ非接触ハンドリング技術、③異物検査技術、④接合技術)に基づき、以下の事業項目を実施した。

- ①接合表面の洗浄技術および前処理技術の開発
- ②半導体チップ非接触ハンドリング技術の開発
- ③①および②を組み合わせた統合プロセスの検討
- ④①～③の成果を定量的に検証するための、COW ダイレクト接合評価デバイスおよび評価技術の開発

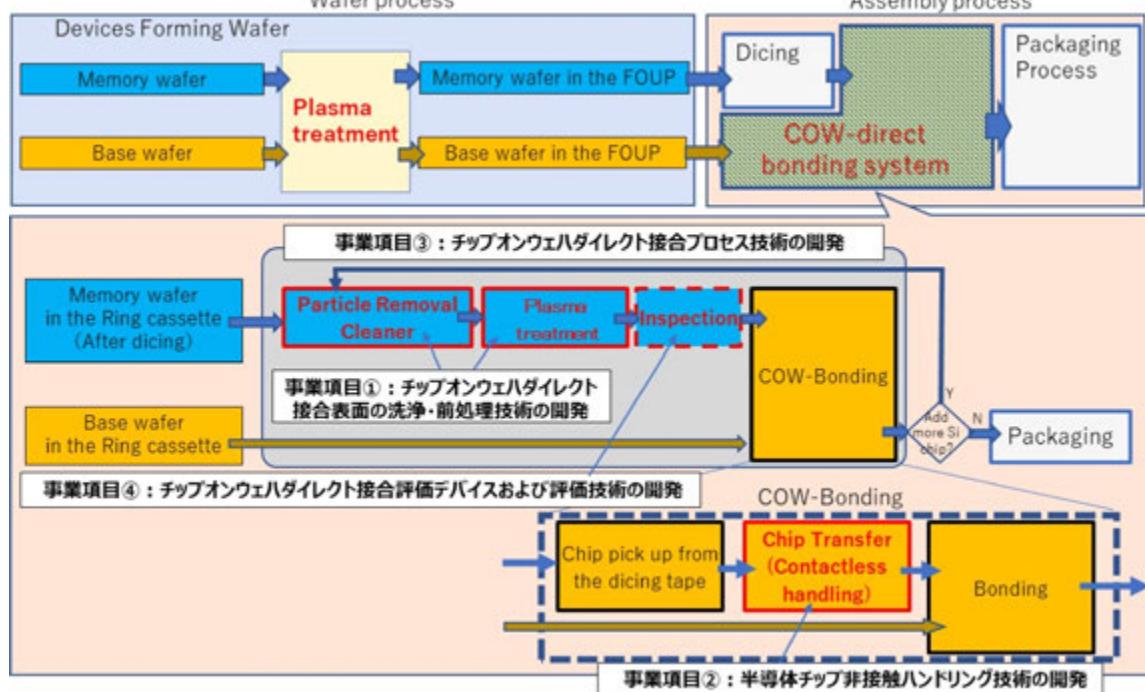


図5 本事業のプロセスの基本構想と事業項目の関係の説明図

### 3 要素技術開発

#### 3-1. COW ダイレクト接合表面の洗浄・前処理技術

##### 3-1-1. 異物除去洗浄技術開発

###### 3-1-1-1. 異物除去洗浄プロセスおよび装置の概要

図6は、開発した異物除去洗浄プロセスの説明図である。

ダイシング後の Si チップをダイシングテープに載せた状態で、異物除去洗浄を行う。

まず、水素水と Megasonic(超音波洗浄)を用いて Si 屑を除去する。水素水に Megasonic を印加することでマイクロバブルが発生し、そのキャビテーション効果により異物を除去する<sup>[4][5]</sup>。

次に、水素水 + Megasonic 洗浄とワイピング洗浄を併用し、樹脂異物を除去する。樹脂異物は粘着性があるため、ワイピング材で浮かせた後、水素水と Megasonic で除去する。

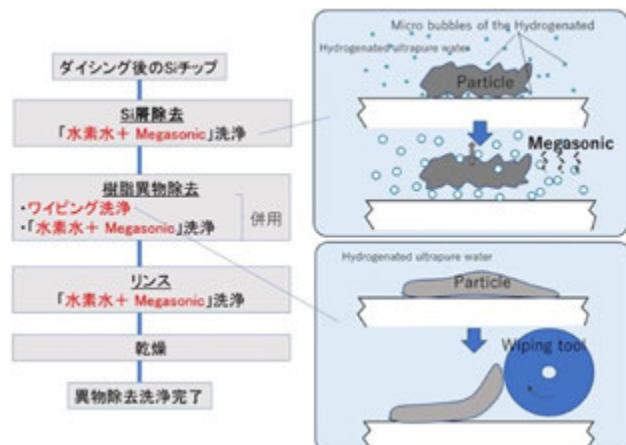


図6 開発した異物除去洗浄プロセスの説明図

続いて、水素水と Megasonic でリシスを行う。浮上した樹脂異物は粘着性があるため、再付着を防止する目的で水素水と Megasonic が有効に機能する。

最後に乾燥工程を経て、異物除去洗浄を完了する。

以下に、異物除去洗浄装置の構造および設計方針について説明する。

写真3は、異物除去洗浄装置の外観である。本装置は二階建ての構造であり、下段に異物除去洗浄装置本体、上段に大気圧プラズマ装置を配置している。将来的な量産対応を見据え、クリーンルーム内の設置面積を最小限に抑えることが重要である。このため、装置のフレーム構造を強化し、低背化を図ることで、二階建て構造を実現した。

また、装置上部にはクリーン化のための HEPA フィルタを搭載しており、定期交換に対応するため、一階の装置と二階の装置間には必要最低限の高さを確保している。

写真4は、異物除去洗浄装置の洗浄チャンバ付近の拡大写真である。本装置には、水素水に超音波を印加して洗浄対象物に吹き付ける“水素水 + Megasonic アーム”と、水素水を掛けながらワイピングを行う“水素水 + ワイピングアーム”が搭載されている。両アームはそれぞれ独立して動作可能である。

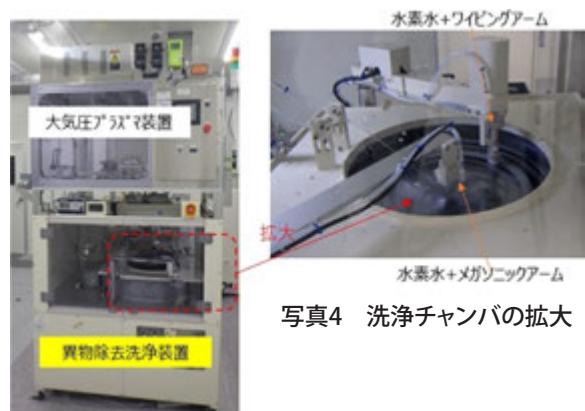


写真4 洗浄チャンバの拡大

#### 写真3 異物除去洗浄装置外観

写真5は、“水素水 + Megasonic アーム”的外観である。水素水はチューブを介して供給され、アーム先端の超音波ヘッドに導入される。ヘッド内で水素水が振動することで、水素のマイクロバブルが形成される。

アーム先端部は傾斜構造となっており、洗浄対象物に対して超音波ヘッドを最適な角度に設定することが可能である。

写真6は、“水素水 + ワイピングアーム”的洗浄時の外観である。アームの先端には、適度な弾性を有し、微小な異物をごく軽い圧力で除去する効果を持つワイピング材を取り付けている。

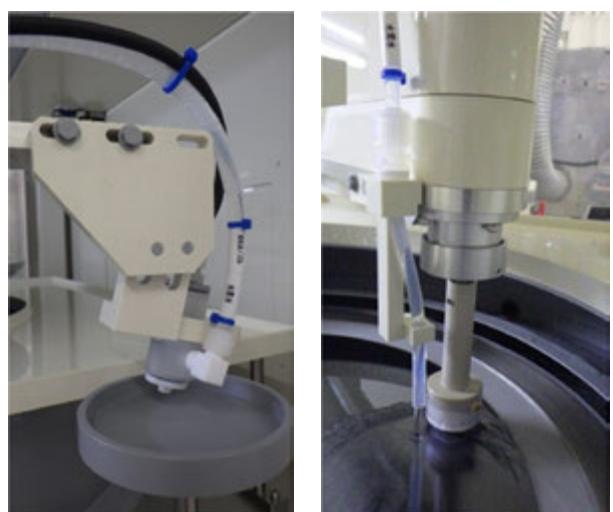


写真5 水素水 + Megasonic アーム



写真6 水素水 + ワイピングアーム

### 3-1-1-2. 洗浄効果の検証

2024年度の目標である、“1μm以上の異物が洗浄後のチップ上で10個以下”を証明するため、異物を試験サンプル上に故意に多数付着させ、洗浄効果の確認を金属顕微鏡により実施した。

写真7は、Si屑の除去例である。Si屑散布後には無数の異物が確認されたが、洗浄後にはすべて除去された。

写真8は、シリコーン樹脂異物の除去例である。

シリコーン樹脂異物散布後には多数の異物が確認されたが、洗浄後にはすべて除去されている。

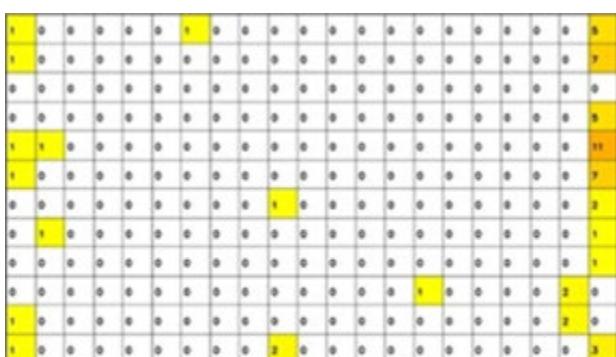
なお、写真7と写真8の拡大写真で示した通り、洗浄後のチップ端面でチッピングなどの欠陥は発生していない。

### 3-1-1-3. 乾燥残りの状況

次に、開発したQPM(Quantitative Phase Microscopy:定量位相顕微鏡)異物検査装置を用いて異物検査を実施した(QPMの詳細は3-3項にて説明する)。

図7は、QPMによる異物検査の一例である。6mm × 13mmのチップ表面を12 × 21のブロックに分割し、各ブロックに0.3μm以上の異物数を数値として表示している。

この検査により、金属顕微鏡では検出されなかった乾燥残りが新たに確認された。



0.3μm以上の異物数 : 53  
 1μm以上異物数 : 21

図7 QPMによる異物検査の一例(水素水洗浄後)

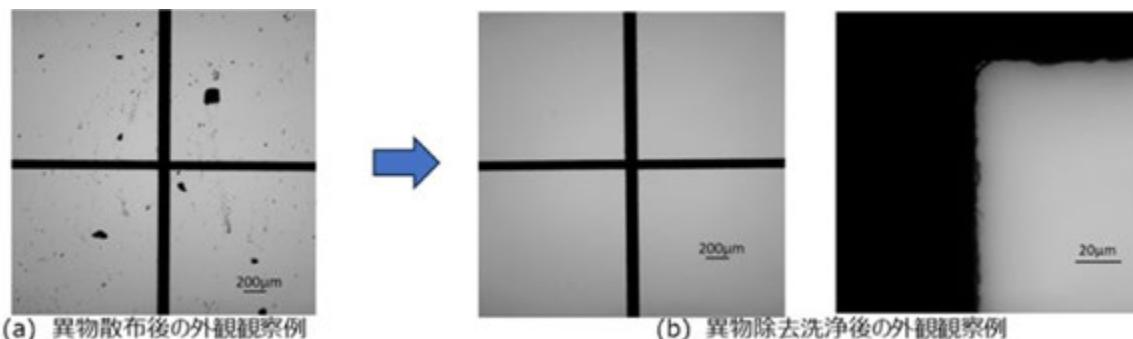


写真7 Si屑の除去例

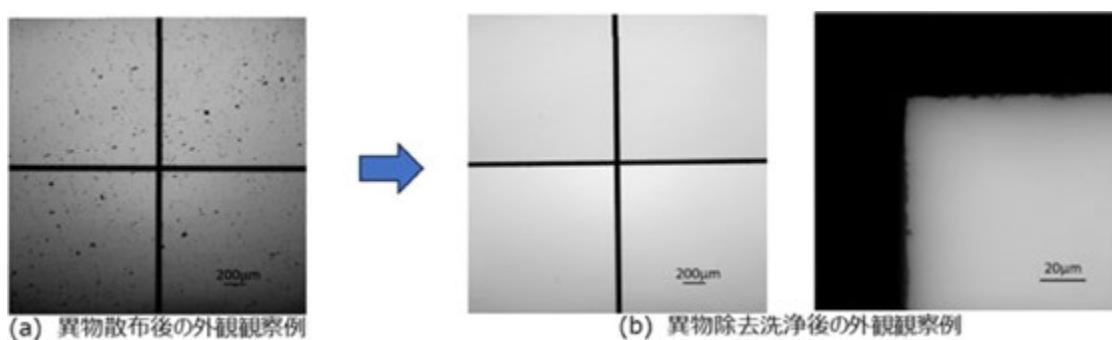


写真8 シリコーン樹脂異物の除去例

写真9-1は、ダイシング時に発生した乾燥残りの金属顕微鏡写真である。検出された異物は、微小な異物が線状に連なる乾燥残りが多く含まれていることが判明した。なお、この乾燥残りは、乾燥後に時間が経過すると写真9-2に示すとおり、水素水洗浄を用いても完全に除去することは困難である。

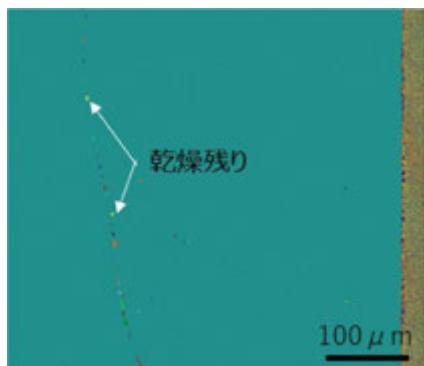


写真9-1 ダイシング後  
乾燥不良の金属顕微鏡写真(水素水洗浄前)

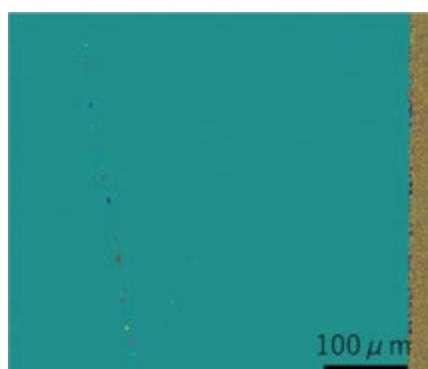


写真9-2 ダイシング後  
乾燥不良の金属顕微鏡写真(水素水洗浄後)

### 3-1-1-4. 乾燥残りの対策

図8に、乾燥残りの対策方法の説明図を示す。

従来の後工程では、ダイシング後のチップをそのまま接合に使用していた。

ダイシングは、ウエハからチップを切り出す個片化と洗浄および乾燥で構成される。本事業では当初、ダイシング後に水素水による異物除去洗浄を実施し、その後にプラズマ処理による表面活性化を行い、最後に接合することを想定していた。しかし、この方法では前述のとおり、ダイシング時に乾燥残りが発生する。

そこで、乾燥残りを防止する方法として、ダイシング後にウエハを速やかに水素水に浸漬し、乾燥させることなくダイシング装置から異物除去洗浄装置まで移動し、直ちに水素水による洗浄を行った。

水素水の浸漬は、樹脂製の容器に新鮮な水素水を満たし、その中にダイシング後のウエハを速やかに投入することで実施した。この処理により、ダイシング装置から洗浄装置への移動中の乾燥を防止することが可能となった。

なお、詳細は3-1-3項に記載するが、水素水によるCu電極の酸化抑制効果により、Cu電極の酸化を防止することが可能である。

表1は、対策プロセス適用前後のQPM異物検査結果である。プロセスの改善(水素水浸漬の適用)により、乾燥残りを含む異物を低減し、本事業の最終目標である10個/chip以下(異物サイズ0.3μm)を達成した。

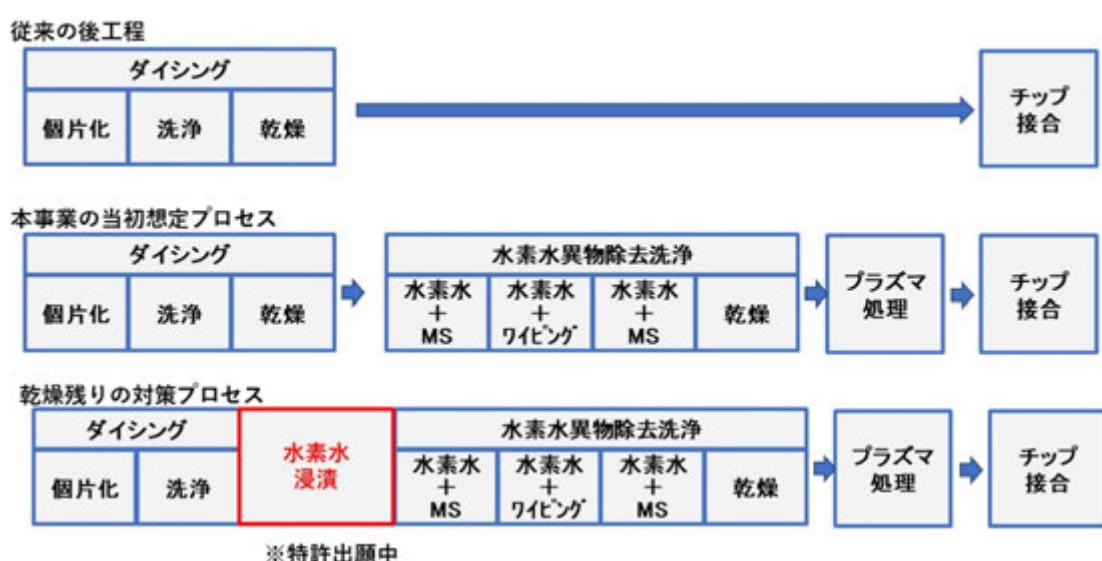


図8 乾燥残りの対策方法の説明図

表1 対策プロセス適用前後のQPM異物検査結果

(目標:10個/chip以下(異物サイズ:0.3μm)。チップサイズ:6mm×13mm)

異物除去洗浄プロセス	Edge部チップ	Middle部チップ	Center部チップ
改善前の異物除去洗浄プロセス	・0.3μm以上の異物数:37 [内訳] 1μm以上異物数:6 0.3μm以上1μm未満:31	・0.3μm以上の異物数:7 [内訳] 1μm以上異物数:0 0.3μm以上1μm未満:7	・0.3μm以上の異物数:0 [内訳] 1μm以上異物数:0 0.3μm以上1μm未満:0
改善後の異物除去洗浄プロセス	・0.3μm以上の異物数:10 [内訳] 1μm以上異物数:0 0.3μm以上1μm未満:10	・0.3μm以上の異物数:8 [内訳] 1μm以上異物数:4 0.3μm以上1μm未満:4	・0.3μm以上の異物数:1 [内訳] 1μm以上異物数:1 0.3μm以上1μm未満:0

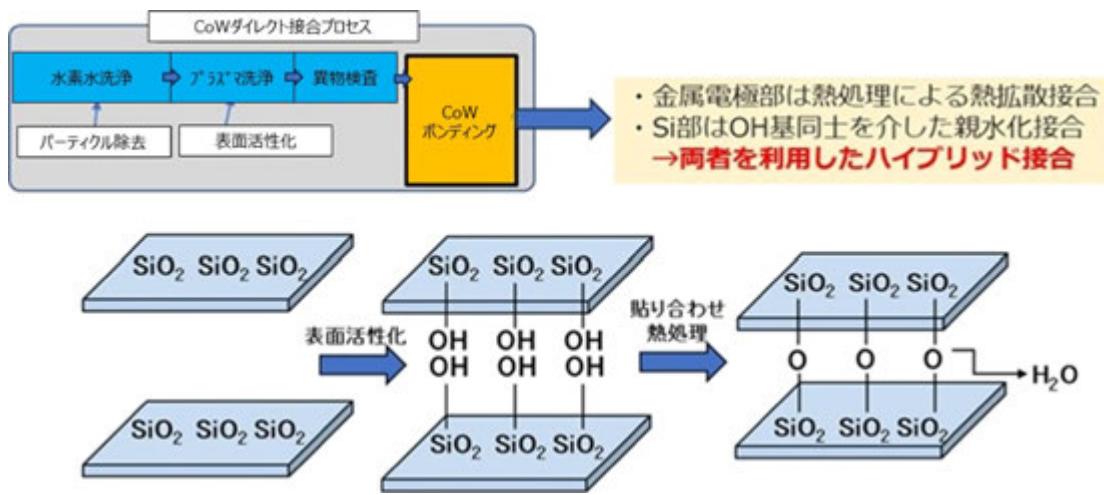
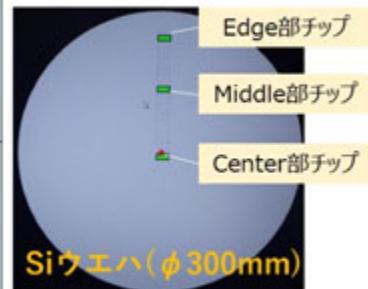


図9 ダイレクト接合の概要説明

### 3-1-2. プラズマ処理による表面活性化

本稿では、プラズマによる Si チップおよび Si ウエハの表面活性化処理の説明を行う。図9は、ダイレクト接合の概要説明である。また、参考として、写真10に真空プラズマ装置の外観を、図10にその構成概略図を示す。

ダイレクト接合は、Si チップ上の銅電極の接合に加え、 $\text{SiO}_2$  の接合も必要である。このため、ハイブリッド接合とも呼ばれる。

$\text{SiO}_2$ 同士の接合(親水化接合)は、水酸基(OH)を介して行われるため、 $\text{SiO}_2$ 表面は高い表面活性化エネルギーを有する状態であることが求められる。そこで、新たに導入した真空プラズマ装置による  $\text{SiO}_2$ 面の表面活性化効果を検証した。



写真10 真空プラズマ装置の外観

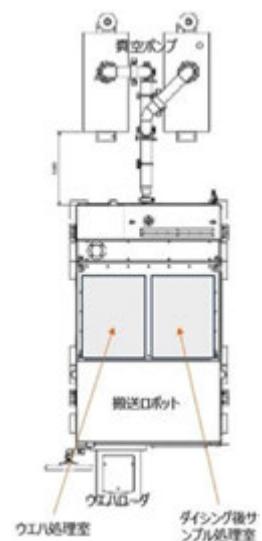


図10 真空プラズマ装置の構成

表2は、真空プラズマ処理の前後における親水性を純水の接触角で比較した結果である。

Si熱酸化膜およびCu薄膜のいずれも、真空プラズマ処理を施すことで純水の接触角が小さくなり、濡れ広がる状態となった。

以上の結果から、真空プラズマ処理により、SiO<sub>2</sub>面およびCu表面の活性化エネルギーが向上することが証明された。

表2 真空プラズマ処理前後の親水性比較

	純水接触角(度)	
	処理前	処理後
Si熱酸化膜	35.42	6.12
Cu薄膜	46.55	15.16

### 3-1-3. 水素水およびプラズマ処理がCu表面に与える影響の解析

#### 3-1-3-1. 大気中放置によるCu表面状態の変化

図11は、大気中放置によるCu表面状態の変化をXPS(X-ray Photoelectron Spectroscopy:X線光電子分光法)により分析した結果である。使用したサンプルは、ウェハ上にCuを成膜したものである。

前処理として酢酸溶液中に浸漬し、あらかじめ酸化物を除去した。

この前処理を実施する前は、金属Cuに加え、Cu<sub>2</sub>O、さらに酸化が進行したCuOおよびCu(OH)<sub>2</sub>(水酸化銅)、CuCO<sub>3</sub>(炭酸銅)が検出された。

前処理を施すことでCuO、Cu(OH)<sub>2</sub>およびCuCO<sub>3</sub>は除去され、金属CuとCu<sub>2</sub>Oのみが検出された。

この状態でクリーンルーム中に静置すると、2日目以降に表面酸化が進行し、金属CuおよびCu<sub>2</sub>Oが減少し、CuO、Cu(OH)<sub>2</sub>およびCuCO<sub>3</sub>が増加した。

なお、この酸化の進行は約14日で安定した。

#### 3-1-3-2. 水素水浸漬およびプラズマ処理の影響解析

図12は、水素水浸漬およびプラズマ処理の影響解析結果である。前処理後12日間クリーンルーム中に放置したサンプルを、非脱気水、脱気水および水素水に浸漬したところ、CuCO<sub>3</sub>、Cu(OH)<sub>2</sub>、およびCuOが減少し、ほとんどが金属CuとCu<sub>2</sub>Oになることが確認された。金属Cuの比率は、非脱気水<脱気水<水素水の順番であった。

また、同様のサンプルを大気圧プラズマに処理すると、Cuの酸化がさらに進行することが確認された。一方、真空プラズマ処理を施すことで、Cuの酸化物が減少することが確認された。これは、処理雰囲気中の酸素の残留量の違いによると考える。

以上の結果から、本事業では水素水による異物除去と真空プラズマによる前処理を採用することとした。

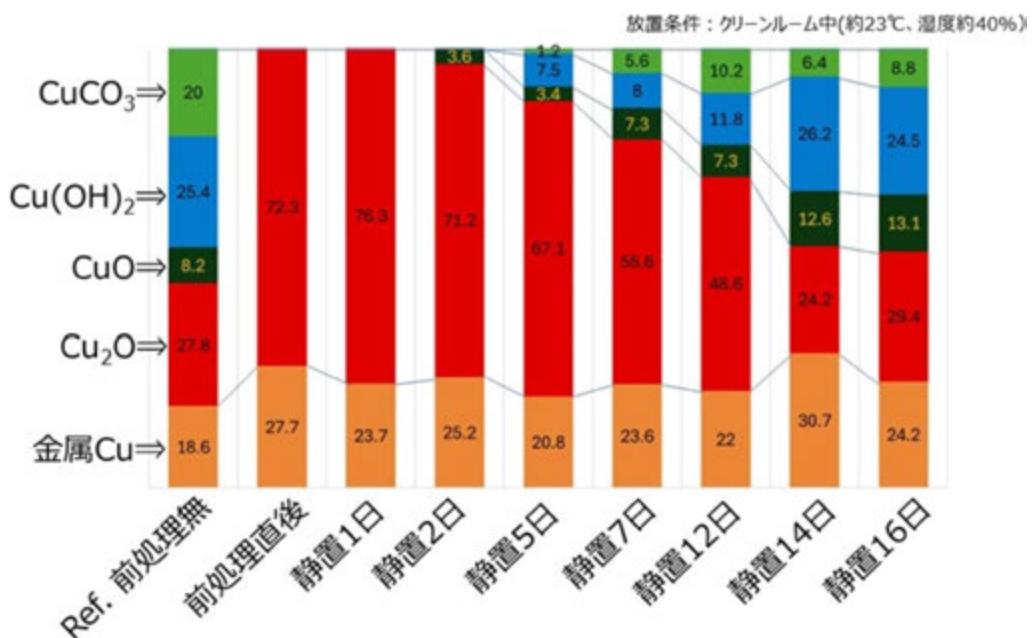


図11 大気中放置によるCu表面状態の変化をXPSにより調べた結果(グラフ中の数字は%)

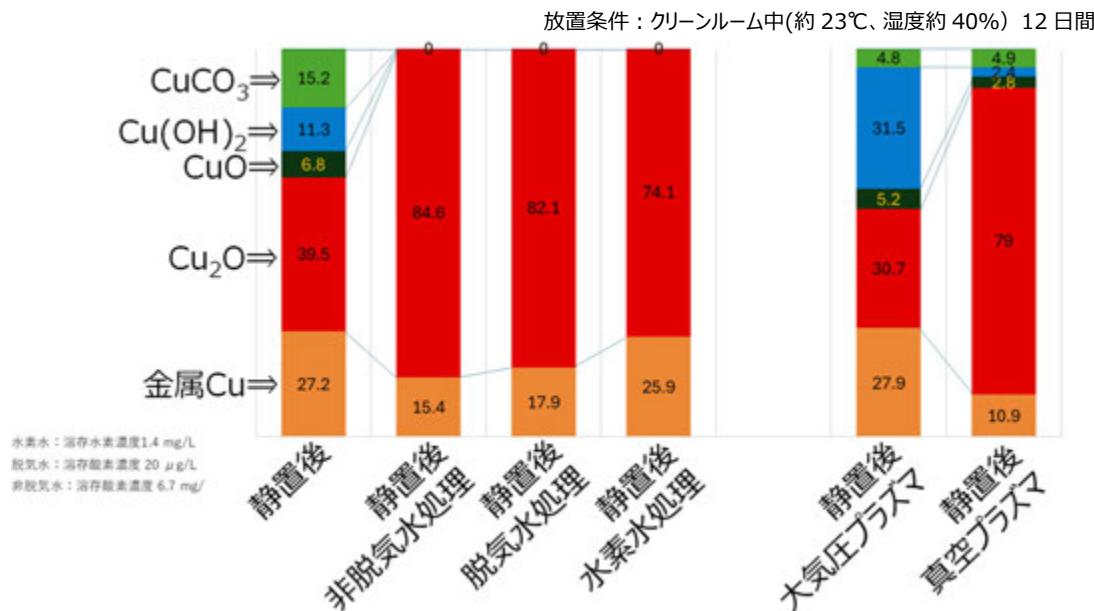


図12 水素水浸漬およびプラズマ処理の影響解析結果(グラフ中の数字は%)

### 3-2. 半導体チップ非接触ハンドリング技術の開発

#### 3-2-1. 超音波非接触ハンドリング技術の概要

図13は、本事業で開発した超音波非接触ハンドリング治具の説明図である<sup>[6][7]</sup>。振動板が振動することで、Siチップに浮上力が発生する。このとき、Siチップには水平方向の把持力が働き、この力により半導体チップのセルフアライメント(自動位置決め)が可能となる。このセルフアライメントにより、従来技術のペルヌーイチャックで必要であったチップ外形のガイドが不要になり、完全非接触のチップハンドリングが可能となった。

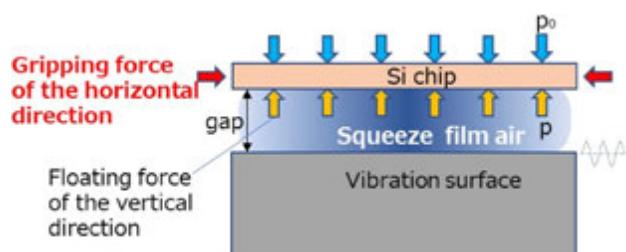


図13 超音波非接触ハンドリング治具の説明図

図14は、治具の小型化設計の説明図である。初期の治具は、ランジュバン型振動子を用いたもので、超音波発振子を含む高さが60mmあり、ボンダへの取付けが困難であった。そこで本事業では、ユニモルフ型の圧電素子を治具に一体化設計することで高さ15mmを実現した。

写真11は、開発した小型超音波非接触ハンドリング治具の外観である。主要部品は土台、振動板、圧電素子である。振動板は4端で土台に固定している。振動板上にはチップを吸引する

ための穴が形成されている。外観写真では見えないが、振動板の裏面には2個の圧電素子が接合されている。また、振動板裏面の中央部には、吸引チューブを取り付けている。

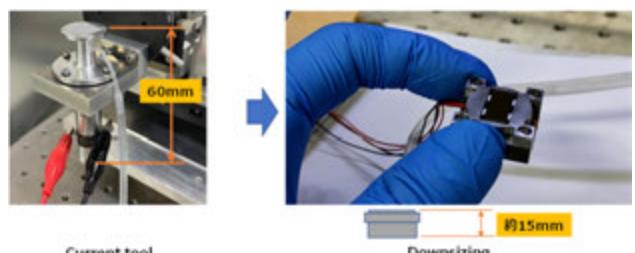


図14 治具の小型化の説明図



写真11 超音波非接触ハンドリング治具の外観

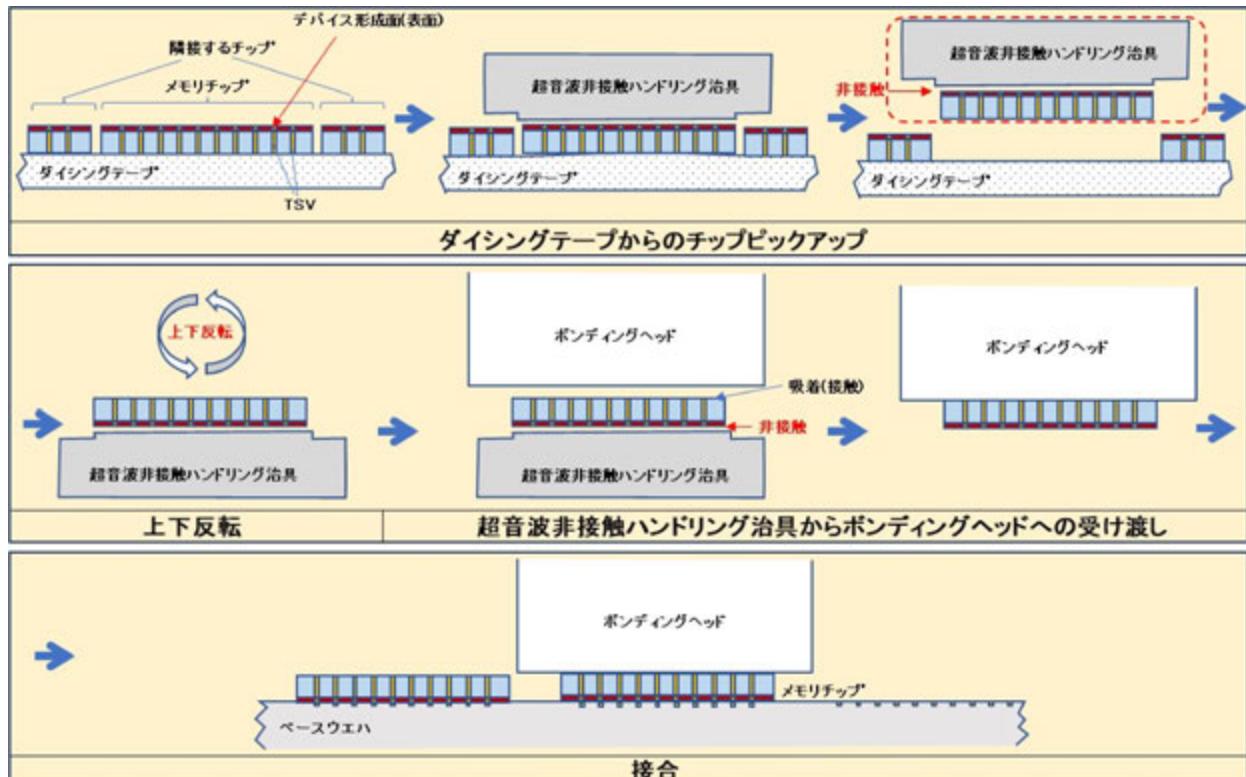


図15 チップ搬送の説明図

図15はチップ搬送の説明図である。

最初に非接触ハンドリング治具でダイシングテープからチップを非接触でピックアップする。その後、非接触ハンドリング治具に非接触でチップを持ったまま上下反転する。そして、非接触ハンドリング治具に非接触でチップを持ったままボンディングヘッドにチップを受け渡す。最後に、ボンディングヘッドでベースウェハにチップを接合する。

このように、ダイシングテープからチップをピックアップ後、接合するまでの間、チップのデバイス形成面は治具に一切接触しない。

### 3-2-2. チップの浮上状態確認

写真12は、Siチップを超音波非接触ハンドリング治具に把持させ、チップをセルフアライメントした時の写真である。

チップは浮上した状態で、自己整合的に治具の中央部に位置決めされる。



写真12 Siチップの把持状態

写真13は、Siチップ浮上観察の一例である。Siチップの厚さは50μmである。超音波出力と吸引を適切に調整することで、浮上したチップを安定して静止させることができる。

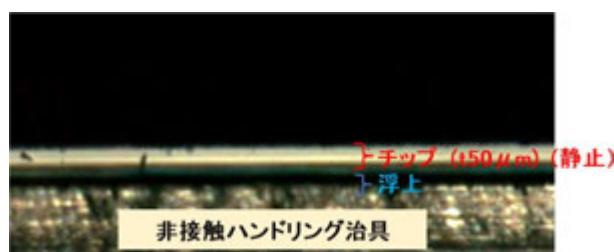


写真13 チップ浮上状態

### 3-2-3. 振動形態の最適化

超音波非接触ハンドリング治具の形状最適化は振動状態のシミュレーションによって実施した。図16は、振動板の溝有無による振動状態の比較シミュレーションである。具体的には、振動板のチップ把持領域(6mm × 13mm)の外側部分に溝を設けることで、超音波振動時にチップ把持領域が反ることなく上下動できるようにしている。

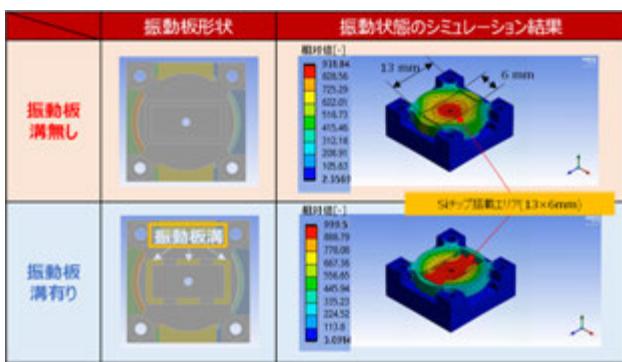


図16 振動板の溝有無による振動状態の比較シミュレーション

### 3-2-4. 接合部ボイド低減の効果確認

図17は、接触型の治具と本事業の非接触ハンドリング治具で、チップ接合後のボイドを SAT(Scanning Acoustic Tomograph: 超音波映像装置)観察で比較した評価結果である。

本事業の超音波非接触ハンドリング治具を用いることで接合部ボイドの低減を達成した。

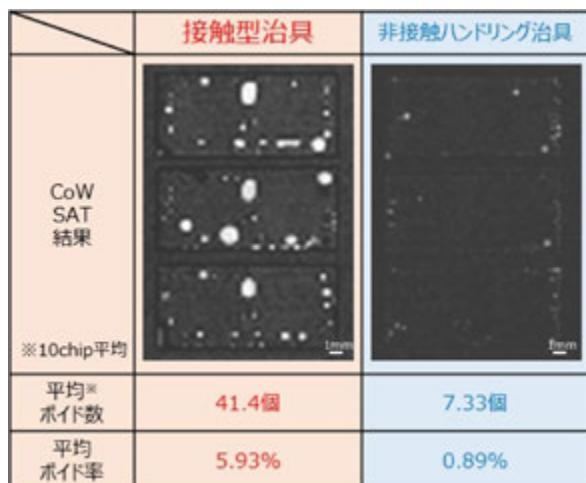


図17 接触ツール／非接触搬送ツールボイド比較

### 3-3. 異物検査技術の開発

#### 3-3-1. 半導体チップ表面異物検査の現状の課題と課題解決のためのコンセプト

半導体チップの表面異物検査における課題は、以下の2点である。

1. Si屑などの不透明異物や、樹脂などの透明異物の付着が想定される。

2. 接合後に異物の数十倍のサイズであるボイドが発生するため、サブミクロンサイズの微小異物も検知する必要がある。

これに対し、従来の半導体前工程で使用されている光散乱式パーティクル検査装置では、透明異物の検出が困難であるという課題があった。

そこで本事業では、これらの課題を解決するため、生物分野で活用されている QPM(Quantitative Phase Microscopy: 定量位相顕微鏡)の応用を進めた<sup>[8]</sup>。

図18は、QPMの観察原理を示すイメージ図である。QPMは、試料に単色光を照射し、反射光の位相差を高さ情報を変換して画像化する。

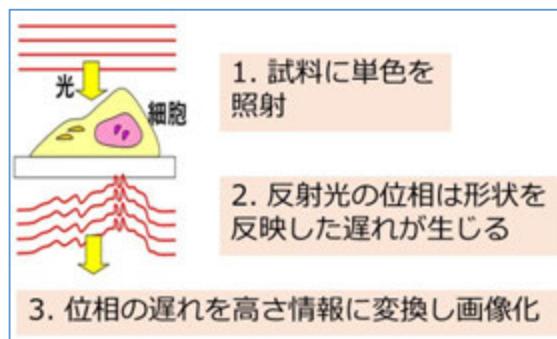


図18 QPM の観察原理

表3は、QPMと他の検査装置の適用可否を比較した一覧である。QPMは、垂直分解能、広範囲観察、透明異物検出などの点で他の検査装置に比べて優れていると考え、本事業での採用を決定した。

表3 QPM と他の検査装置の適用可否

	QPM	光散乱式パーティクル検査装置	レーザー顕微鏡	金属顕微鏡
水平分解能	可 サブミクロンレベル	可 サブミクロンレベル	可 サブミクロンレベル	不可 ミクロンレベル
垂直分解能	良 サブナノレベル	不可 ノン接触型は困難	可 ナノレベル	不可
広範囲観察	良 1shotで数30mm四方を撮像可能	可 Φ300mmウエルを検査可能	可 XY走査が必要であり、撮像に時間が必要	良 1shotで数30mm四方を撮像可能
不透明異物検出(Si屑など)	良	不可	良	不可
透明異物検出(樹脂など)	良	不可	良	不可

写真14は、QPM 検査装置の試作機の外観である。

検査装置は、 $\phi 400\text{mm}$ のダイシングリングのサンプルおよび $\phi 300\text{mm}$ ウェハの検査に対応している。

### 3-3-2. 検出能力限界の検証

図19は、QPM 異物検査装置試作機により標準パーティクル(大きな保証された標準粒子)を観察した結果である。

標準パーティクルは、ウェハ上に不透明な Si 粒子と、透明な  $\text{SiO}_2$  粒子が搭載されている。

観察の結果、QPM は不透明・透明を問わず、 $\phi 0.2\mu\text{m}$  以上の異物を検出可能であることが確認された。



写真14 試作機外観

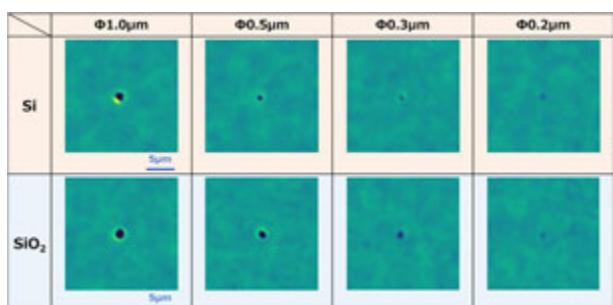


図19 QPM 異物検査装置試作機での標準パーティクル観察結果

### 3-4. 接合評価

#### 3-4-1. 巻き込みボイド対策

COW ダイレクト接合は、フリップチップボンダを用いてベースウェハ上にチップを搭載する方式である。このとき、ボンダのボンディングヘッド先端に取り付けたコレットでチップを把持し、チップをウェハに押し付ける。チップとベースウェハ間に空気が残留すると、巻き込みボイド(空隙)が発生する。

図20は、ボンディングコレットの形状と巻き込みボイドの発生状況を比較した一例である。

実験では、従来型の“平コレット”と、開発した“外周支持コレット”を比較した。

平コレットはチップサイズに合わせて製作し、表面は平坦で、真空吸引穴を9個形成している。

外周支持コレットもチップ外形に合わせて製作し、チップ外周部のみを支持する構造とし、中央部は凹形状としてチップと接触しないように設計した。

COW 接合後のボイド比較では、平コレットを用いた場合は、空気の巻き込みによる大きなボイドが多発した。一方、外周支持コレットでは、空気の巻き込みによる大きなボイドはほとんど見られなかった。

コレット形状が巻き込みボイドの発生に影響を与える要因は、チップを吸着した際のチップ反り形状に起因すると考えられ、現在も解析を継続している。



図20 ボンディングコレット形状と巻き込みボイドの比較

### 3-4-2. COW 電極の接合

開発した異物除去洗浄および前処理プロセスを適用したサンプルを用いて、COW 電極の接合を評価した。

写真15は、接合前処理として真空プラズマおよび水素水洗浄を施した電気評価用デバイスの COW 接合品における電極部の断面観察例である。Cu-Cu 電極の接合が確認された。

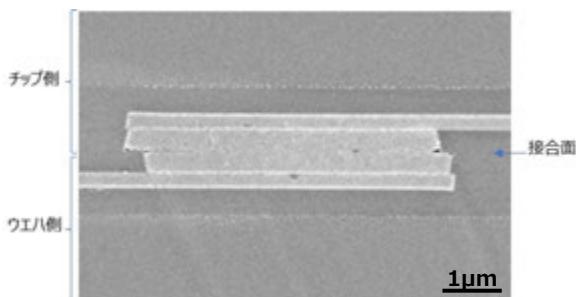


写真15 COW 接合品の電極部断面観察例

## 4 結言

ポスト5G 社会の先端半導体に用いられる COW ダイレクト接合を早期に確立することを目的として、要素技術の開発を実施した。その結果、以下の結論を得た。

- (1) 半導体チップへの異物付着防止を目的として、超音波非接触ハンドリング技術を開発した。ボンダに搭載可能な超小型ツールを設計・試作することで、厚さ50μm 以下のチップの完全非接触搬送を実現した。
- (2) 水素水を用いた異物除去洗浄装置および洗浄プロセスを開発した。ダイシングに起因する異物除去に加え、乾燥残りへの対策を講じることで、異物10個 / チップ以下(異物サイズ0.3μm)を達成した。
- (3) QPM(定量位相顕微鏡)を用いた異物検査装置を開発した。これにより、サブミクロンサイズの透明異物および不透明異物の検出が可能となり、異物除去洗浄プロセスの改善に貢献した。
- (4) ボンディングツールの改良および異物低減の取り組みにより、接合部のボイド低減を達成した。また、COW 実装後の Cu-Cu ダイレクト接合を実現した。

以上の開発成果により、COW ダイレクト接合技術の実用化に貢献した。

なお、本開発完了後には、半導体メーカーより試作評価の依頼を受けており、今後は実デバイスに近いサンプルを用いた評価を進める計画である。

### ■謝辞

本成果は、NEDO の助成事業“ポスト5G情報通信システム基盤強化研究開発事業”により得られたものである。

研究にあたっては、産業技術総合研究所の菊地克弥様、高橋健司様、島本晴夫様には、ダイレクト接合技術に関するご指導を賜った。また、東京理科大学の宮武教授には、超音波非接触ハンドリング技術に関するご指導をいただいた。

水素水を活用した異物除去洗浄プロセスの開発においては、栗田工業株式会社の井田純一様にご協力をいただいた。

本事業は、多くの方々のご支援とご協力のもとに実施・完遂することができた。ここに厚く御礼を申し上げる。

なお、ヤマハロボティクス株式会社は、本事業の共同研究先である産業技術総合研究所および東京理科大学と CEATEC AWARD 2025にてイノベーション部門賞を受賞致しました。

### ■参考文献

- [1] 須賀唯知：“ものづくりのための接合技術”，精密工学会誌, Vol. 79, No. 8, pp. 705-709, 2013
- [2] 仁村将次：“積層型機能性デバイスに向けたハイブリッド接合技術に関する研究”，早稲田大学, 博士論文, 2016
- [3] 李瑾, 菊地広, 永田憲雅, 井田純一, 宮崎洋一：“水素水とワイピングを用いた半導体チップの精密洗浄技術”，エレクトロニクス実装学会誌, 24巻, 3号, pp. 268-276, 2021
- [4] 大見忠弘：“半導体製造プロセスにおけるクリーン化技術”，表面技術, Vol. 50, No. 10, pp. 848-854, 1999
- [5] 森田博志：“機能水を用いたシリコン基板の新しい洗浄技術に関する研究”，東北大学, 博士論文, 2011
- [6] 菊沼隼, 宮武正明, 菊地広, 戸張優太：“半導体チップ搬送における超音波スクイーズ効果を利用した非接触ハンドリング技術開発”，エレクトロニクス実装学術講演大会講演論文集, 38巻, pp. 13A5-1, 2024
- [7] 代田雄高, 吉本成香, 宮武正明：“超音波スクイーズ効果を利用した半導体ウエハ用非接触チャックに関する研究”，精密工学会学術講演会講演論文集, 2006A巻, pp. 945-946, 2006
- [8] 戸張優太, 菊地広, 菊沼隼, 李瑾：“チップオンウェハダイレクト接合技術の開発”，第36回マイクロエレクトロニクス研究会 Proceeding, Nov. 2024

■著者



菊地 広  
Hiroshi Kikuchi  
ヤマハロボティクス株式会社  
技術研究本部  
先端技術開発部



菱沼 隼  
Hayato Hishinuma  
ヤマハロボティクス株式会社  
技術研究本部  
先端技術開発部



李 瑾  
Jin Li  
ヤマハロボティクス株式会社  
技術研究本部  
先端技術開発部



永田 憲雅  
Norimasa Nagata  
ヤマハロボティクス株式会社  
技術顧問



尾川 謙一  
Kenichi Ogawa  
ヤマハロボティクス株式会社  
技術研究本部  
先端技術開発部

# DiffMamba: Leveraging Mamba for Effective Fusion of Noise and Conditional Features in Diffusion Models for Skin Lesion Segmentation

Amit Shakya Shruti Phutke Chetan Gupta Rupesh Kumar Lalit Sharma  
Chetan Arora

This paper, titled “DiffMamba: Leveraging Mamba for Effective Fusion of Noise and Conditional Features in Diffusion Models for Skin Lesion Segmentation,” was presented at CVIP-2024 (9th International Conference on Computer Vision & Image Processing), held at IIITDM Kancheepuram, Chennai, India, from December 19–21, 2024. It received the Best Industry Paper Award.

Reprinted with permission. Copyright © 2024 CVIP and IIITDM.

Further use or distribution is not permitted without permission from CVIP and IIITDM.

## Abstract

Effective Skin Lesion Segmentation is crucial for dermatological care, it enables the early identification and accurate diagnosis of skin cancer. Denoising Diffusion Probabilistic Models (DDPMs) have recently become a major focus in computer vision. Its applications in image generation, such as Stable Diffusion, Latent Diffusion Models and Imagen, have showcased remarkable abilities in creating high-quality generative outputs. Recent research highlights that DDPMs also perform exceptionally well in medical image analysis, specifically in medical image segmentation tasks. Even though a U-Net backbone served as the foundation for these models initially, there is a promising opportunity to boost their performance by incorporating other mechanisms. Recent research includes a transformer-based framework for diffusion models, but the advancement comes with the challenge of inherent quadratic complexity. Research has shown that state space models (SSMs), like Mamba efficiently capture long-range dependencies while maintaining linear computational complexity. Due to these benefits, it outperforms many of the mainstream foundational architectures. However, we found that simply merging Mamba with diffusion results in suboptimal performance. To truly harness the power of these two advanced technologies for medical image segmentation, a more effective integration is required, we formulate a novel Mamba-Based Diffusion framework, called Diff-Mamba for skin lesion segmentation. We access its performance on the ISIC 2018 dataset for skin lesion segmentation, and our method outperforms existing state-of-the-art techniques. The code is available at: <https://github.com/amit-shakya-28/DiffMamba>

## 1 INTRODUCTION

**Background** Melanoma, the most fatal type of skin cancer, results from the abnormal growth of melanocyte cells due to the activation of mutation by unusual Deoxyribonucleic Acid (DNA) damages. Melanocyte cells create melanin, which is the substance responsible for skin colour<sup>[1]</sup>. Melanoma cases have increased sharply over the last 30 years, with around 10,000 deaths annually in the USA<sup>[35]</sup>. It is highly curable, if detected in early stage<sup>[19]</sup>. In dermatology, early melanoma diagnosis

is achieved by visual examination through methods like the ABCD criteria<sup>[27]</sup> (asymmetry, border irregularity, colour patterns, and diameter) and the seven-point checklist<sup>[3]</sup>. However, these methods are time consuming and face subjectivity issue. Automated medical image segmentation methods gained considerable interest recently for their potential to diagnose the diseases accurately. The effectiveness of these models in medical image segmentation stems from progress in deep learning technologies, ranging from widely used convolutional neural networks (CNN)<sup>[4][31][17]</sup> to the newer

vision transformer architectures (ViT)<sup>[5][39]</sup>.

**Denoising Diffusion Probabilistic Models** Recently, the DDPM has gained significant recognition as a potent class of generative models. The rising acknowledgment of DDPMs has sparked significant interest and research, fueled by their remarkable ability to produce high-quality and diverse samples, as demonstrated by models like DALLE2<sup>[29]</sup>, Imagen<sup>[34]</sup>, and Stable Diffusion<sup>[30]</sup>. Building on these advancements, researchers have introduced innovative techniques for medical image segmentation utilizing diffusion models. By leveraging DDPMs, many approaches have achieved cutting-edge results across various benchmarks. The outstanding results of these models arise from their built-in stochastic sampling process<sup>[28]</sup>. DDPMs can produce varied segmentation predictions through repeated runs, with the diversity of these outputs effectively reflecting the inherent uncertainty in medical images. This is particularly valuable for segments where organs or lesions often have uncleared or indistinct boundaries.

**Mamba Based Diffusion Framework** However, it is noteworthy that all these approaches are built upon traditional U-Net backbone. Compared to the growing trend of using state space models (Mamba), traditional U-Net models compromised the segmentation quality, due to which it generates a diverse but incorrect mask during ensemble. Ultimately, this can introduce persistent noise that permanently degrades performance. Building on this momentum, we want to combine the Mamba-based U-Net model such as U-Mamba<sup>[26]</sup> with the diffusion model. However, we observed that a straightforward implementation yielded inadequate performance. One of the main reasons behind it is that the features produced by the Mamba are not compatible with those from the diffusion backbone. The Mamba extracts detailed semantic information directly from the original image, while the diffusion backbone handles features from a noisy and corrupted mask, which complicates the process of combining these features. To mitigate these shortcomings, we formulate a novel Mamba-driven Diffusion Framework for skin lesion segmentation, called DiffMamba.

The main concept is to combine conditional embedding and diffusion embedding. To effectively link these two embeddings, we proposed a novel Mamba-based Fusion module for their integration. The feature fusion module merges noise and semantic features using a cross-mamba block (CroMB) and the merged features further in succeeding Mamba block (see Section 3.3). To align the noise and conditions at each step, CroMB utilizes cross-input features to the mamba block to enrich features, while integrated output of CroMB processed in further Mamba block are refined features through selective scanning and concatenation to produce the final fused output. The proposed work provides the following key aspects.

- We formulate a novel method that combine Mamba and diffusion model for skin lesion segmentation.
- We introduce an innovative Mamba-based fusion module that seamlessly merges conditional semantic features with diffusion noise. As far as the authors are aware, this is the first successful method to combine diffusion and condition embeddings in skin lesion segmentation.
- The proposed method is evaluated using the ISIC 2018 dataset for comparison.

The enhancement in Intersection over Union (IoU) and Dice score relative to current leading medical image segmentation methods demonstrates the efficacy of the proposed approach.

## 2 RELATED WORKS

### 2-1. Skin Lesion Image Segmentation

**Traditional Approaches** Before the advent of deep learning, image segmentation was predominantly driven by classical methods and machine learning techniques. These included approaches like adaptive thresholding<sup>[7]</sup>, support vector machines<sup>[48]</sup>, region growing<sup>[18]</sup>, unsupervised clustering<sup>[48]</sup>, and active contours<sup>[13]</sup>. These methods were heavily dependent on manually designed features, which were difficult to create and often lacked the adaptability and effectiveness required for handling

more complex datasets. As a result, they struggled with larger and more intricate data.

**Deep Learning Approaches** This section offers a summary of deep learning methods used for skin lesion segmentation, with a focus on the U-Net model introduced by Ronneberger et al.<sup>[32]</sup>, which efficiently leverages data augmentation to make the most of limited labeled samples in biomedical imaging. Bi et al.<sup>[6]</sup> created a multi-stage fully convolutional network for skin lesion segmentation, which includes stages for both coarse and fine boundary learning, along with a parallel integration method to enhance detection. Yuan et al.<sup>[45]</sup> proposed a fully automated skin lesion segmentation approach that employs a 19-layer DCNN and uses Jaccard Distance as the loss function. Despite extensive parameter optimization and testing on the ISBI 2016 and PH2 datasets, their approach did not perform as good as state-of-the-art methods, notably when handling low-contrast images. The 2020s saw a shift in computer vision with the rise of vision transformers, disrupting the dominance of CNNs and leading to innovations like Swin-UNet<sup>[15]</sup>, which integrates Swin Transformer<sup>[24]</sup> blocks into U-Net models. Hybrid architectures like UCTransNet<sup>[38]</sup> and MCTrans<sup>[46]</sup> combine CNNs and transformers, while all-transformer models such as SMESwin-UNet<sup>[47]</sup> utilize transformers throughout the entire U-Net structure. Despite these advancements, challenges remain due to the limited availability of annotated segmentation data compared to classification data, which affects the precision and reliability of segmentation algorithms.

## 2-2. Diffusion Models for Medical Image Segmentation

Building on recent advancements, researchers have introduced innovative medical image segmentation techniques that utilize diffusion models to address this complex problem. For example, EnsDiff<sup>[41]</sup> uses ground truths for training and treats input images as priors to create segmentation distributions, which aids in generating uncertainty maps and an implicit ensemble of segmentation maps. Kim et al.<sup>[20]</sup> introduce an innovative method for self-supervised vessel segmentation. On the other hand, MedSegDiff<sup>[42]</sup> utilizes diffusion probabilistic

models (DPM) for medical image segmentation, incorporating dynamic conditional encoding alongside the FF-Parser helps to mitigate the impact of high-frequency noise. The subsequent MedSegDiff-V2<sup>[43]</sup> improves upon this method by incorporating a conditional U-Net, which strengthens the interaction between semantic features and noise.

## 2-3. Mamba in Computer Vision

Previous methods in semantic segmentation either use CNNs<sup>[21]</sup>, which are scalable but constrained by small receptive fields and weight-sharing limitations, or Vision Transformers (ViTs)<sup>[36]</sup>, which provide better global context but suffer from quadratic complexity and efficiency issues. To overcome these challenges, Selective Structured State Space Models (Mamba)<sup>[9]</sup> have become increasingly popular for their ability to cover global receptive fields and use dynamic weights while maintaining linear complexity. Mamba has proven highly effective in long sequence modeling tasks, particularly in natural language processing<sup>[9]</sup>. Its capabilities are also being investigated in vision tasks like image classification<sup>[23]</sup>, medical image segmentation<sup>[26][33]</sup>, and 3D scene comprehension<sup>[22]</sup> introduced a residual state space block that combines channel attention with Mamba and a 2-D selective scanning technique for better image restoration. Inspired from its success, we introduced a fusion module in our proposal.

## 3 METHODOLOGY

---

This part provides a thorough explanation of our proposed Mamba-based Diffusion framework for semantic segmentation. We commence with an explanation of the foundational ideas of DDPMs and State Space Models. Then, we provide a summary of the architecture we propose, followed by a detailed examination of the fusion module.

### 3-1. Denoising Diffusion Probabilistic Models

We offer a concise summary of the diffusion models as described in<sup>[14]</sup>. These generative models are defined by a Markov chain and consist of a forward process, in

which the data undergoes gradual degradation by adding Gaussian noise, and a backward process, in which the degradation is progressively reversed by removing the noise. The forward process, denoted as  $q$ , is given by the following formulation:

$$q(x_{1:T}|x_0) = \prod_{t=1}^T q(x_t|x_{t-1}) \quad (1)$$

Here,  $T$  are number of steps  $x_1, x_2, \dots, x_T$  are the latent variables at each step, and  $x_0$  is the initial data sample. For every step in the forward pass, the Gaussian noise is given as:

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t \mathbf{I}_{n \times n}) \quad (2)$$

where  $\beta_t$  specifies the noise schedule with a constant and  $\mathbf{I}_{n \times n}$  is the identity matrix having size  $n \times n$ . The forward pass allows for sampling at any arbitrary timestamp  $t$ , is described in [14], which can be reparametrised to

$$x_t = \sqrt{\alpha_t}x_0 + \sqrt{1-\alpha_t}\epsilon, \quad \epsilon \sim \mathcal{N}(0, \mathbf{I}_{n \times n}) \quad (3)$$

where,

$$\begin{aligned} \alpha_t &= 1 - \beta_t \\ \bar{\alpha}_t &= \prod_{s=0}^t \alpha_s \end{aligned} \quad (4)$$

The reverse pass, parameterized with  $\theta$ , is defined as:

$$p_\theta(x_{0:T-1}|x_T) = \prod_{t=1}^T p_\theta(x_{t-1}|x_t). \quad (5)$$

Starting with  $p_\theta(x_T) = \mathcal{N}(x_T; 0, \mathbf{I}_{n \times n})$ , this process transforms the distribution from  $p_\theta(x_T)$  to  $p_\theta(x_0)$ . The reverse pass is carried out by applying Gaussian steps described with:

$$p_\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t)). \quad (6)$$

As shown in [14], we can then predict  $x_{t-1}$  from  $x_t$  with

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left( x_t - \frac{1-\alpha_t}{\sqrt{1-\alpha_t}} \epsilon_\theta(x_t, t) \right) + \sigma_t z, \quad (7)$$

where  $z \sim \mathcal{N}(0, \mathbf{I})$ , and  $\sigma_t$  represents the variance scheme to be learned, as proposed in [14]. As seen in Equation 7, the sampling process includes a random component  $z$ , resulting in stochastic behavior. Note that  $\epsilon_\theta$  refers to the U-Net model containing a Mamba-based fusion module we train, with input  $x_t = \sqrt{\alpha_t}x_0 + \sqrt{1-\alpha_t}\epsilon$ . The model  $\epsilon_\theta(x_t, t)$

that is subtracted from  $x_t$  during sampling, as described in Equation 7, needs to be learned by the model. This U-Net with Mamba is trained using the loss objectives specified in [14].

### 3-2. State Space Models

State Space Models (SSMs) [10][11] are a framework used for sequence-to-sequence tasks, characterized by their time-invariant dynamics, also referred to as linear time-invariant (LTI) properties. Because of their linear complexity, SSMs are ideal for capturing the dynamics of systems by mapping them to latent states, described as follows:

$$y(t) = Ch(t) + Dx(t), \quad \dot{h}(t) = Ah(t) + Bx(t) \quad (8)$$

Here,  $x(t) \in \mathbb{R}$  represents the input,  $h(t) \in \mathbb{R}^N$  is hidden state, and  $y(t) \in \mathbb{R}$  is the output.  $N$  denotes the state size, and  $\dot{h}(t)$  refers to the time derivative of  $h(t)$ . Additionally,  $A \in \mathbb{R}^{N \times N}$ ,  $B \in \mathbb{R}^{N \times 1}$ ,  $C \in \mathbb{R}^{1 \times N}$ , and  $D \in \mathbb{R}$  are the system matrices. Since the matrix  $D$  is considered a residual connection between the input and output, State Space Models (SSMs) are often represented by omitting  $D$ , leading to the following simplified forms:

$$\begin{aligned} \dot{h}(t) &= Ah(t) + Bx(t); & \text{state equation} \\ y(t) &= Ch(t); & \text{output equation} \end{aligned} \quad (9)$$

This representation captures the global feature dependency of SSMs, as the current output depends on all preceding states and the input. While the above equations assume continuous-time inputs, deep learning approaches treat the input as discrete in time. To handle discrete sequences such as text and images, SSMs use zero-order hold (ZOH) discretization [10]. This approach maps the continuous  $\{x_1, x_2, \dots, x_K\}$  (input sequence) to the discrete  $\{y_1, y_2, \dots, y_K\}$  (output sequence). To achieve this, a time scale parameter  $\Delta \in \mathbb{R}^D$  is introduced, transforming the matrices  $A$  to  $\bar{A}$  and  $B$  to  $\bar{B}$  as follows:

$$\begin{aligned} \bar{A} &= \exp(\Delta A), \\ \bar{B} &= (\Delta A)^{-1} (\exp(A) - I) \cdot \Delta B, \\ \bar{C} &= C, \\ y_k &= \bar{C}\bar{h}_k + \bar{D}\bar{x}_k, \\ h_k &= \bar{A}h_{k-1} + \bar{B}\bar{x}_k. \end{aligned} \quad (10)$$

In this context, the dimensions of all matrices remain consistent throughout each iteration of the process. In addition, as detailed in Mamba<sup>[9]</sup>, a first-order Taylor expansion is used to approximate the matrix  $B$ :

$$\bar{B} = (\exp(A) - I)A^{-1}B \approx \Delta B$$

where we simplify  $(\Delta A)(\Delta A)^{-1}\Delta B$  to  $\Delta B$ . This approximation facilitates the analysis and implementation of the model by reducing the computational complexity associated with matrix exponentiation. By simplifying  $(\exp(A) - I)A^{-1}B$  to  $\Delta B$ , we make the model more tractable and easier to work with in practical applications.

### 3-3. Proposed Architecture

Our approach utilizes the diffusion model<sup>[42]</sup>, comprising a forward diffusion process adding the Gaussian noise and the reverse process with neural network restoring the original data. This process is explained mathematically in Section 3.1. To ensure consistency with forward process, the noisy image is iteratively refined with the reverse process, progressively restoring it to achieve the final clear segmentation (see Figure 1). Similar to the standard diffusion probabilistic model (DPM) setup, we use an encoder-decoder network for training. For segmentation purpose, we enhance the model by conditioning the noise prediction function  $\epsilon_\theta$  on information from the original image.

$$\epsilon_\theta(x_t, I, t) = DC(MambaF(F_{I_t}, F_{x_t}), t) \quad (11)$$

Here,  $F_{I_t}$  and  $F_{x_t}$  (see Figure 1), signifies the feature embeddings of the original image (condition) and diffusion input at the current stage  $t$  respectively. These two embeddings are merge together using Mamba based fusion module (MambaF) and then fed into a UNet decoder (DC) for the reconstruction process. The complete process of DiffMamba is shown in Figure 1. To explain this, we consider the single step  $t$  in diffusion process. Initially, the noisy mask  $x_T$  is passed through a UNet, which is called the Diffusion Model. The model is directed by semantic features of the raw images using a separate UNet, referred to as the condition UNet. Following this, the semantic condition is integrated into

the encoder features of the Diffusion UNet, combining the semantic segmentation embeddings from the condition UNet. This integration is controlled by the Fusion Module, which refines the representation by connecting noise and semantic embeddings, utilizing the global and adaptive features of Mamba. The proposed approach used a standard noise prediction loss  $L_n$  for training, analogous to diffusion probabilistic models (DPMs)<sup>[14]</sup>.

**Features Fusion with Mamba** <sup>[37]</sup> proposed a feature fusion module that integrate multi-modal features using two blocks, Concat-Mamba Block (ConMB), and Cross-Mamba Block (CroMB). The CroMB enhances features through cross-multiplication and selective scanning, while ConMB combines outputs from CroMB using concatenation and selective scanning. Inspired from this approach<sup>[37]</sup>, we propose the Mamba Fusion block to merge the semantic (conditional) and noise features. The combined features are then forwarded to UNet decoder in the diffusion model for further processing. Let us suppose the conditional features and noisy features are represented by  $F_{I_t}^o$  and  $F_{x_t}^o$  respectively, the complete fusion process is formulated as:

$$\begin{aligned} F_{I_t}^o, F_{x_t}^o &= CroMB(F_{I_t}, F_{x_t}) \\ MambaF_{Output} &= Mamba(Conv < F_{I_t}^o, F_{x_t}^o >). \end{aligned} \quad (12)$$

where,  $< \cdot >$  is concatenation,  $Conv$  is  $1 \times 1$  Convolution,  $Mamba$  is a mamba block (see Mamba in Figure 1 for more details). Here, all the features remain in original dimension. The Mamba block basically works as a gating mechanism with the State space model. In Cross mamba block we utilize the cross features for generating the gating mechanism (see the inputs  $F_{I_t}$  and  $F_{x_t}$  provided differently to the two parallel mamba blocks in Figure 1). The Mamba block uses the 2D-selective scanning mechanism (SS2D)<sup>[12]</sup> as shown in Figure 1. Furthermore, refer to equation 11 to see how  $MambaF_{Output}$  features are used to condition the noise prediction function.

## 4

### EXPERIMENTS AND RESULT DISCUSSION

This section provides details about the dataset used in the

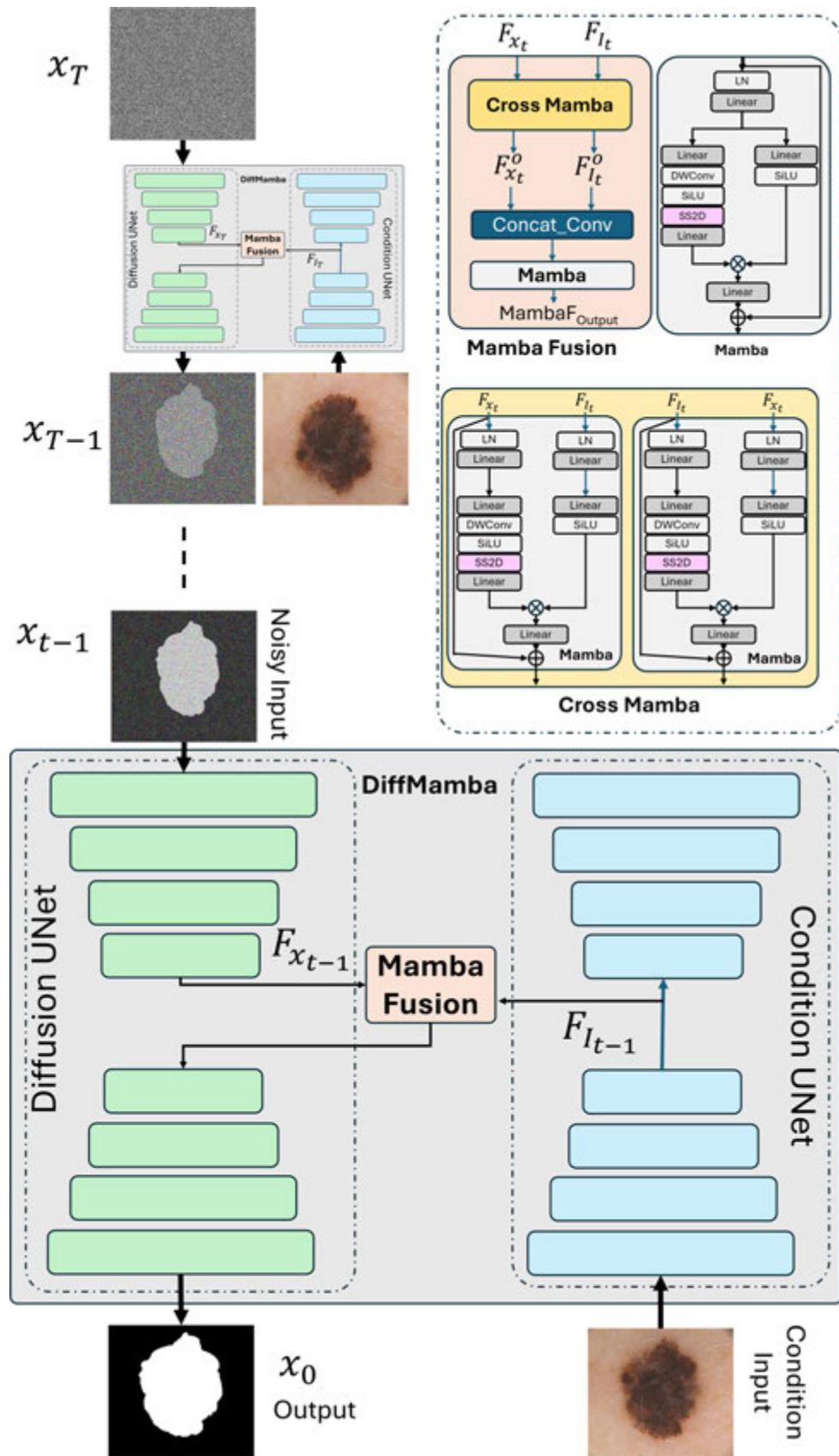


Fig. 1 Details of the Proposed Architecture for Skin Lesion Segmentation: We present a Mamba-based fusion method that integrates noise from the diffusion encoder with conditional features to enhance the effectiveness of image segmentation

experiments, outlines the implementation process, explains the evaluation metrics, and discusses the results.

#### 4-1. Dataset

The ISIC 2018 dataset<sup>[8]</sup>, from International Skin Imaging Collaboration (ISIC), offers a diverse and extensive collection of dermoscopy images. In which, Task 1 is dedicated to lesion segmentation having a total of 3,694 images, of these, 2,594 images are designated for training, consisting of 72% nevi, 20% melanomas, and 8% seborrheic keratoses. Additionally, 100 images are allocated for validation, and 1,000 for testing. The image resolutions range from 0.5 to 29 megapixels, with dimensions spanning from  $540 \times 576$  to  $4,499 \times 6,748$  pixels. Figure 2 illustrates sample images from the dataset.

#### 4-2. Training and Implementation Details

The proposed segmentation network was trained on images which are resized to a resolution of  $256 \times 256$ . The model was trained for 100k iterations, employing the AdamW<sup>[25]</sup> optimizer with a *batch size* = 8 and a *learning rate* = 0.0001. For inference, 100 diffusion steps were used, and the ensemble of 5 times model execution is considered, which is less than 25 runs performed in MedSegDiff<sup>[44]</sup>. The STAPLE algorithm<sup>[40]</sup> was applied to merge the samples. The experiments are implemented on PyTorch framework and executed on an NVIDIA A100 GPU.

#### 4-3. Evaluation Metrics

The proposed network's performance was measured using the IoU and Dice score to compare it with leading medical image segmentation methods.

**Dice Score** The Dice coefficient is measured using the recall and precision from prediction, evaluating the similarity of the prediction and the ground truth. It also accounts for false positives, which is generally useful in datasets with significant class imbalance, such as those found in medical image segmentation. Mathematically it is defined as:

$$\text{Dice} = \frac{2 \times \text{True Positive}}{2 \times \text{True Positive} + \text{False Positive} + \text{False Negative}} \quad (13)$$

**Intersection Over Union (IoU)** It evaluates the overlap of the prediction and the ground truth by evaluating the ratio of their common area to the total area covered by both. In mathematical terms, it is defined as:

$$\text{IoU} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive} + \text{False Negative}} \quad (14)$$

The difference between the two metrics is that the IoU penalizes under-and over-segmentation more than Dice score.

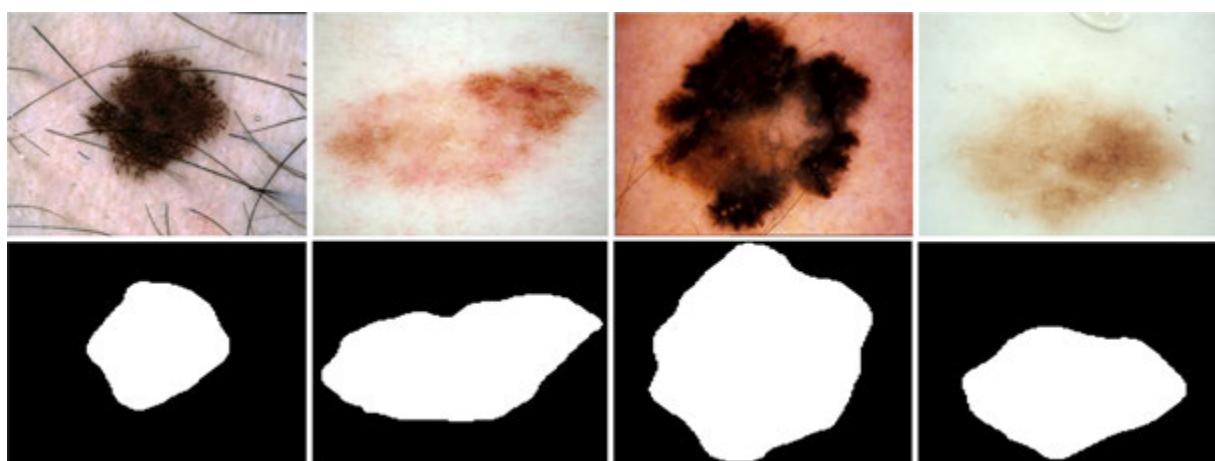


Fig. 2 Sample Images from ISIC 2018 dataset. First and second row represent the image and Ground Truth respectively

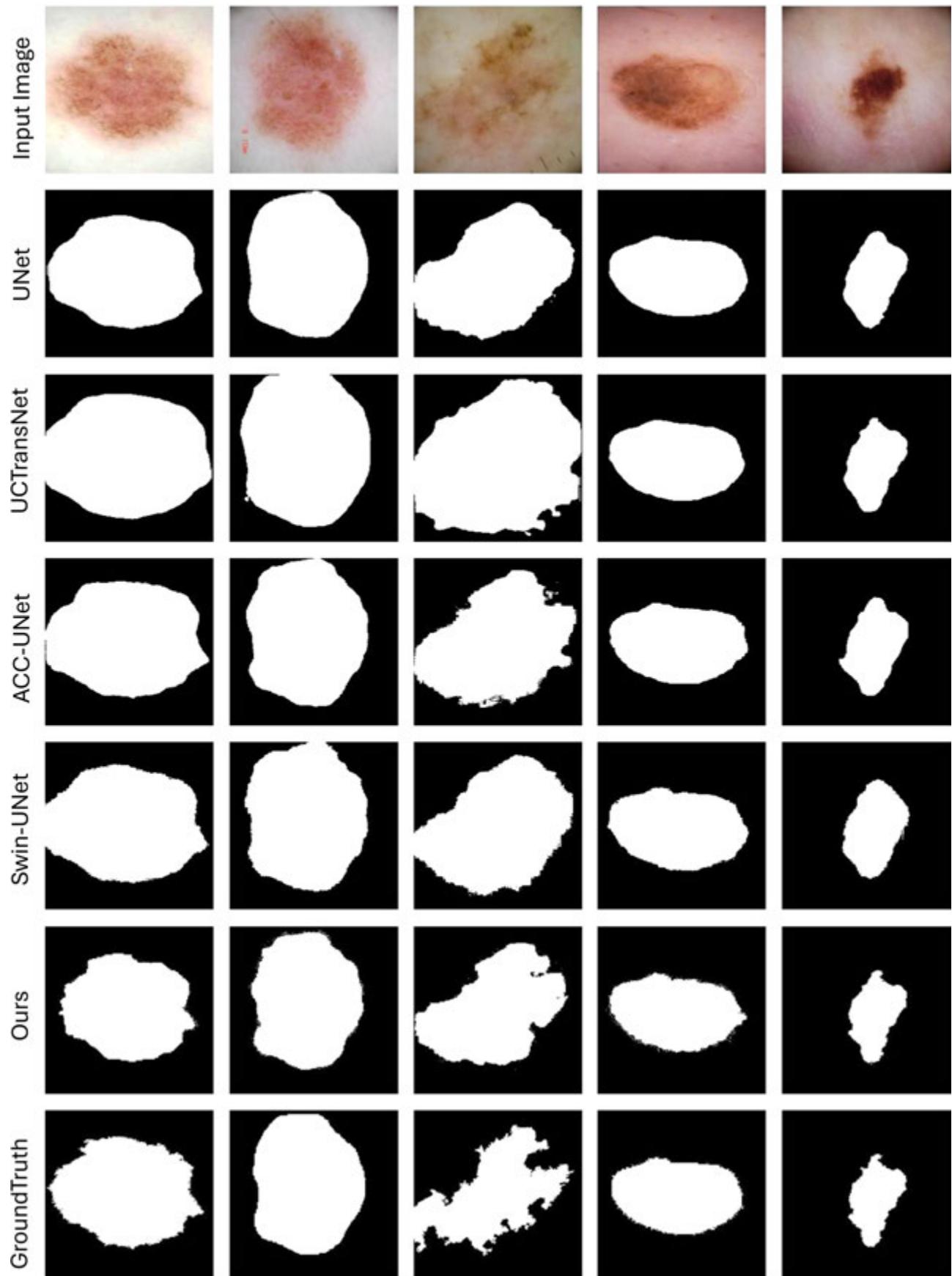


Fig. 3 Methods comparison based on qualitative results from the ISIC 2018 dataset

**Table 1** A comparison of the proposed method with existing state-of-the-art approaches for skin lesion segmentation. Note: The **best** is indicated with **bold**, and **second best** is indicated with underline

Dataset	ISIC 2018	
Metric	Dice score	IoU
UNet <sup>[32]</sup>	85.41	76.85
UCTransNet <sup>[38]</sup>	86.69	<u>78.35</u>
ACC-UNet <sup>[16]</sup>	86.57	78.81
Swin-UNet <sup>[15]</sup>	88.03	80.02
SegDiff <sup>[2]</sup>	87.30	79.42
EnsemDiff <sup>[41]</sup>	88.21	80.72
MedSegDiff <sup>[42]</sup>	91.30	<u>84.14</u>
Ours	92.72	<u>86.73</u>

#### 4-4. Result and Discussion

Table 1 provides an analysis of our proposed DiffMamba on the ISIC 2018 dataset. We consider IoU and Dice score (DSC) for evaluation. According to the results, our method outshines both CNN and Transformer based strategies, emphasizing its ability to accurately capture boundaries on the ISIC 2018 dataset. In particular, our proposed approach demonstrates better performance compared to methods that rely on transformers such as Swin-UNet, CNNs such as UNet, and hybrid models such as UCTransNet, ACC-UNet, and diffusion-based methods<sup>[38][16][15][42]</sup>. Additionally, DiffMamba outperforms the baseline model (Medsegdiff) with improvements of +1.42% in the DSC score and +2.59% in IoU on the ISIC 2018 dataset. Moreover, Figure 3 shows that our method excels in visually capturing detailed structures and defining boundaries more accurately than other models. This visual analysis highlights the superior performance of the Mamba-based fusion module, which effectively captures long-range dependencies while keeping computational complexity linear during the learning process.

## 5

## CONCLUSION

This study proposed the DiffMamba diffusion network for segmenting skin lesions. We improve the diffusion-based framework for medical image segmentation by integrating Mamba mechanism with the original UNet

backbone. We introduced a fusion module to align noise and semantic features. The comparative analysis is carried out on ISIC 2018 dataset which shows that our approach outperforms the existing SOTA methods. Considering its effectiveness on ISIC 2018 dataset, in future, the approach will be evaluated on the other medical image segmentation datasets to verify its generalizability. A Mamba-based diffusion model for skin lesion segmentation is being introduced for the first time as per our knowledge, we anticipate that DiffMamba will set a new standard for future research in this field.

## REFERENCES

- [1] American Cancer Society: Cancer Facts Figures 2016. American Cancer Society, Atlanta, GA, USA (2016)
- [2] Amit, T., Shaharbany, T., Nachmani, E., Wolf, L.: Segdiff: Image segmentation with diffusion probabilistic models. arXiv preprint arXiv:2112.00390 (2021)
- [3] Argenziano, G., Fabbrocini, G., Carli, P., Giorgi, V. D., Sammarco, E., Delfino, M.: Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions: Comparison of the abcd rule of dermatoscopy and a new 7-point checklist based on pattern analysis. Archives of Dermatology 134(12), 1563–1570 (1998)
- [4] Azad, R., Aghdam, E. K., Rauland, A., Jia, Y., Avval, A. H., Bozorgpour, A., Karim-ijafarbigloo, S., Cohen, J. P., Adeli, E., Merhof, D.: Medical image segmentation review: The success of u-net. arXiv preprint arXiv:2211.14830 (2022)
- [5] Azad, R., Kazerouni, A., Heidari, M., Aghdam, E. K., Molaei, A., Jia, Y., Jose, A., Roy, R., Merhof, D.: Advances in medical image analysis with vision transformers: A comprehensive review. arXiv preprint arXiv:2301.03505 (2023)
- [6] Bi, L., Kim, J., Ahn, E., Kumar, A., Fulham, M., Feng, D.: Dermoscopic image segmentation via multistage fully convolutional networks. IEEE Transactions on Biomedical Engineering 64(9), 2065–2074 (2017)
- [7] Celebi, M. E., Wen, Q., Iyatomi, H., Shimizu, K., Zhou, H., Schaefer, G.: A state-of-the-art survey on lesion border detection in dermoscopy images. Dermoscopy image analysis 10(1), 97–129 (2015)

- [8] Codella, N., Rotemberg, V., Tschandl, P., Celebi, M. E., Dusza, S., Gutman, D., Helba, B., Kalloo, A., Liopyris, K., Marchetti, M., et al.: Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). arXiv preprint arXiv:1902.03368 (2019)
- [9] Gu, A., Dao, T.: Mamba: Linear-time sequence modeling with selective state spaces. arXiv preprint arXiv:2312.00752 (2023)
- [10] Gu, A., Goel, K., Ré, C.: Efficiently modeling long sequences with structured state spaces. arXiv preprint arXiv:2111.00396 (2021)
- [11] Gu, A., Johnson, I., Goel, K., Saab, K., Dao, T., Rudra, A., Ré, C.: Combining recurrent, convolutional, and continuous-time models with linear state space layers. Advances in Neural Information Processing Systems 34, 572–585 (2021)
- [12] Guo, H., Li, J., Dai, T., Ouyang, Z., Ren, X., Xia, S. T.: Mambair: A simple baseline for image restoration with state-space model. In: European Conference on Computer Vision. pp. 222–241. Springer (2025)
- [13] Hemalatha, R., Thamizhvani, T., Dhivya, A. J. A., Joseph, J. E., Babu, B., Chan-drasekaran, R.: Active contour based segmentation techniques for medical image analysis. Medical and biological image analysis 4(17), 2 (2018)
- [14] Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. Advances in Neural Information Processing Systems 33, 6840–6851 (2020)
- [15] Hu, C., Wang, Y., Joy, C., Dongsheng, J., Xiaopeng, Z., Qi, T., Manning, W.: Swin-Unet: Unet-like pure transformer for medical image segmentation. In: Computer Vision-ECCV 2022 Workshops. pp. 205–218. Springer Nature Switzerland, Cham (2023)
- [16] Ibtehaz, N., Kihara, D.: Acc-unet: A completely convolutional unet model for the 2020s. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 692–702. Springer (2023)
- [17] Isensee, F., Jaeger, P. F., Kohl, S. A., Petersen, J., Maier-Hein, K. H.: nnu-net: A self-configuring method for deep learning-based biomedical image segmentation. Nature Methods 18(2), 203–211 (2021)
- [18] Iyatomi, H.: Computer-based diagnosis of pigmented skin lesions. book: Campolo D, ed. New developments in biomedical engineering. pp. 183–200 (2010)
- [19] Jerant, A. F., Johnson, J. T., Sheridan, C., Caffrey, T. J.: Early detection and treatment of skin cancer. American Family Physician 61(2), 357–386 (2000)
- [20] Kim, B., Oh, Y., Ye, J.: Diffusion adversarial representation learning for self-supervised vessel segmentation. In: The Eleventh International Conference on Learning Representations (2023)
- [21] LeCun, Y., Bengio, Y.: Convolutional Networks for Images, Speech, and Time Series. MIT Press, Cambridge, MA, USA (1998)
- [22] Liang, D., Zhou, X., Wang, X., Zhu, X., Xu, W., Zou, Z., Ye, X., Bai, X.: Point-mamba: A simple state space model for point cloud analysis. arXiv preprint (2024)
- [23] Liu, Y., Tian, Y., Zhao, Y., Yu, H., Xie, L., Wang, Y., Ye, Q., Liu, Y.: Vmamba: Visual state space model. arXiv preprint arXiv:2401.10166 (2024)
- [24] Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., Guo, B.: Swin transformer: Hierarchical vision transformer using shifted windows. In: 2021 IEEE/CVF International Conference on Computer Vision (ICCV). pp. 9992–10002. IEEE (October 2021)
- [25] Loshchilov, I., Hutter, F.: Fixing weight decay regularization in adam. arXiv preprint arXiv:1711.05101(5) (2017)
- [26] Ma, J., Li, F., Wang, B.: U-mamba: Enhancing long-range dependency for biomedical image segmentation. arXiv preprint arXiv:2401.04722 (2024)
- [27] Nachbar, F., Stolz, W., Merkle, T., Cognetta, A. B., Vogt, T., Landthaler, M., Bilek, P., B. -Falco, O., Plewig, G.: The abcd rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions. Journal of the American Academy of Dermatology 30(4), 551–559 (1994)
- [28] Rahman, A., Valanarasu, J. M. J., Hacihaliloglu, I., Patel, V. M.: Ambiguous medical image segmentation using diffusion models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 11536–11546 (2023)
- [29] Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., Chen, M.: Hierarchical text-conditional image generation with clip latents. arXiv preprint arXiv:2204.06125 (2022)
- [30] Rombach, R., Blattmann, A., Lorenz, D., Esser, P.,

- Ommer, B.: High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 10684–10695 (2022)
- [31] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 234–241. Springer (2015)
- [32] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, proceedings, part III 18. pp. 234–241. Springer (2015)
- [33] Ruan, J., Xiang, S.: Vm-unet: Vision mamba unet for medical image segmentation. arXiv preprint arXiv:2402.02491 (2024)
- [34] Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E., Ghasemipour, S., Ayan, B., Mahdavi, S., Lopes, R., et al.: Photorealistic text-to-image diffusion models with deep language understanding. arXiv preprint arXiv:2205.11487 (2022)
- [35] Skincancer.org: Melanoma - skincancer.org (2016), <http://www.skincancer.org/skin-cancer-information/>
- [36] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, Polosukhin, I.: Attention is all you need. In: Advances in Neural Information Processing Systems 30. Curran Associates, Inc. (2017)
- [37] Wan, Z., Wang, Y., Yong, S., Zhang, P., Stepputtis, S., Sycara, K., Xie, Y.: Sigma: Siamese mamba network for multi-modal semantic segmentation. arXiv preprint arXiv:2404.04256 (2024)
- [38] Wang, H., Cao, P., Wang, J., Zaiane, O.: Uctransnet: Rethinking the skip connections in u-net from a channel-wise perspective with transformer. Proceedings of the AAAI Conference on Artificial Intelligence 36(3), 2441–2449 (June 2022)
- [39] Wang, J., Wei, L., Wang, L., Zhou, Q., Zhu, L., Qin, J.: Boundary-aware transformers for skin lesion segmentation. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference. pp. 206–216. Springer (2021)
- [40] Warfield, S. K., Zou, K. H., Wells, W. M.: Simultaneous truth and performance level estimation (staple): an algorithm for the validation of image segmentation. IEEE Transactions on Medical Imaging 23(7), 903–921 (2004)
- [41] Wolleb, J., Sandkühler, R., Bieder, F., Valmaggia, P., Cattin, P.: Diffusion models for implicit image segmentation ensembles. In: International Conference on Medical Imaging with Deep Learning. pp. 1336–1348. PMLR (2022)
- [42] Wu, J., Fang, H., Zhang, Y., Yang, Y., Xu, Y.: Medsegdiff: Medical image segmentation with diffusion probabilistic model. arXiv preprint arXiv:2211.00611 (2022)
- [43] Wu, J., Fu, R., Fang, H., Zhang, Y., Xu, Y.: Medsegdiff-v2: Diffusion based medical image segmentation with transformer. arXiv preprint arXiv:2301.11798 (2023)
- [44] Wu, J., Fu, R., Fang, H., Zhang, Y., Yang, Y., Xiong, H., Liu, H., Xu, Y.: Medsegdiff: Medical image segmentation with diffusion probabilistic model. In: Medical Imaging with Deep Learning. pp. 1623–1639. PMLR (2024)
- [45] Yuan, Y., Chao, M., Lo, Y. C.: Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance. IEEE transactions on medical imaging 36(9), 1876–1886 (2017)
- [46] Yuanfeng, J., Zhang, R., Huijie, W., Zhen, L., Lingyun, W., Shaoting, Z., Ping, L.: Multi-compound transformer for accurate biomedical image segmentation. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2021. pp. 326–336. Springer International Publishing, Cham (2021)
- [47] Ziheng, W., Min, X., Fangyu, S., Ruinan, J. S. N., Ichen, Y., Ryoichi, N.: Smeswin unet: Merging cnn and transformer for medical image segmentation. In: Medical Image Computing and Computer Assisted Intervention (MICCAI) 2021 (2021)
- [48] Zortea, M., Skrøvseth, S. O., Schopf, T. R., Kirchesch, H. M., Godtliebsen, F.: Automatic segmentation of dermoscopic images by iterative classification. International journal of biomedical imaging 2011(1), 972648 (2011)

■著者



**Amit Shakya**  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India

**Shruti Phutke**  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India

**Chetan Gupta**  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India



**Rupesh Kumar**  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India

**Lalit Sharma**  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India

**Chetan Arora**  
Indian Institute of Technology  
Delhi

# Two-Stage, Global-Local Approach for Cell Nuclei Segmentation in Histopathology Images

Amit Shakya Revat Saharan Chetan Gupta Rupesh Kumar Lalit Sharma  
Srivatsava Naidu Subrahmanyam Murala Chetan Arora

This paper, titled “Two Stage, Global-Local Approach for Cell Nuclei Segmentation in Histopathology Image,” was published in the proceedings of CVIP-2024 and is available via IEEE Xplore (DOI: 10.1109/CVIP57892.2024.10899523).

Reprinted with permission. Copyright © 2024 IEEE.

Further use or distribution is not permitted without permission from IEEE.

## Abstract

Effective management of high-resolution, and spatially wide contextual cues is fundamental to the accurate semantic segmentation. Traditional approaches like multi-resolution feature maps, and skip-connection are effective but require changes in the backbone architecture, restricting utilization of newer models and architectures for the problem. In this work we propose an architecture-agnostic, two-stage, global-local frame-work, called GoLo, for the semantic segmentation, which can use arbitrary semantic segmentation models within its two stages. We focus on segmenting cell nuclei in histopathology image analysis, where accurate segmentation of cell nuclei boundaries is one of the key issues. The proposed framework consists of first stage with Global and second stage with Local learning approach. The first stage is proposed to process the image globally and provide the coarse nuclei segmentation map. In the second stage, to process the image locally, coarse segmentation map and input image is first converted into patches. These patches are then fed as input to the second stage to get the fine- grained segmentation map. Both stages are trained with a combination of dice and binary cross entropy loss. To show the effectiveness of our approach, we test 4 state-of-the-art segmentation architectures (ACC-UNet, UCTransnet, Swin-UNet, and Vanilla U-Net), on 4 different benchmark datasets (MoNuSeg, CPM-17, CoNSep, and TNBC). We evaluate performance of each technique before and after using our framework. We report an average improvement of 4.82% in mIoU, and 4.52% mDSC score, across techniques, and datasets.

## 1 INTRODUCTION

### A. Background

Cancer ranks as the second leading cause of death and disease burden worldwide, following cardiovascular disease<sup>[1]</sup>. It emphasizes the critical need for early detection to im- prove treatment outcomes. Changes in the shape and structure of cell nuclei are significant indicators in cancer diagnosis. Histopathological examination is the primary method in which pathologists diagnose cancer by visually inspecting tissue changes and cellular abnormalities under a microscope. This process is time-consuming and highly reliable on human expertise<sup>[2]</sup>. Studies have shown significant variability among pathologists in diagnosing biopsies, with an average disagreement rate of 24.7%<sup>[3]</sup>. This high

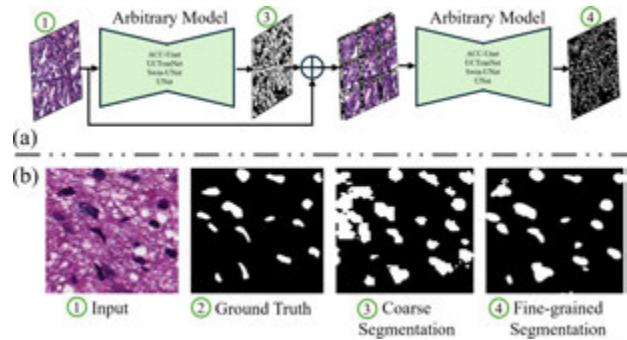


Fig. 1 The key idea of this work is to use a two-stage framework for accurate segmentation. In the first stage, we use the full image (which needs to be inputted at a lower resolution due to memory constraints) with a state-of-the-art model (we have experimented with ACC-UNet, UCTranNet, Swin-UNet, and UNet), which identifies all the nuclei but with inaccurate boundaries (image 3). In the second stage, we crop each individual nuclei based on the mask in the first stage (along with a small context around the boundary), and give full resolution version which leads to accurate boundaries (image 4)

variability highlights the potential for computer-assisted histopathology interpretation to improve diagnostic consistency and accuracy.

### **B. Cell Nuclei Segmentation**

One of the most important and fundamental task for computer-assisted histopathology interpretation is cell nuclei segmentation. It refers to the process of identifying and delineating the boundaries of cell nuclei within images obtained from microscopy. In traditional approaches to cell nuclei segmentation, scientists employed techniques such as Water-shed segmentation, K-means clustering, and fuzzy C-means to delineate cell nuclei within histopathology images<sup>[4]</sup>. These methods have limited performance for blurry images. So, scientists are looking for better ways to do this job automatically and accurately<sup>[5]</sup>. Deep learning has emerged as a superior solution for cell nuclei segmentation due to its ability to automatically learn and extract features from images, leading to more accurate and robust segmentation results<sup>[6]</sup>. Unlike traditional methods, deep learning models can segment nuclei of different cells with diverse image conditions<sup>[7][8]</sup>. This adaptability is crucial in biomedical imaging, where nuclei can vary greatly in size, shape, and intensity. Once trained, deep learning models can process large-scale image datasets quickly, making them suitable for high-throughput analysis. While these methods have shown some performance gains, they still suffer from issues like complex post processing and excessive redundant computations. While traditional methods have been effective to some extent, they often require manual intervention and parameter tuning, leading to subjective results and limited generalizability. U-Net<sup>[9]</sup>, an architecture built on Fully Convolutional Networks (FCN), is widely used and has achieved excellent results in medical image segmentation. Many single stage approaches have been proposed, an extensive review is presented in [10].

### **C. Two-Stage Approaches for Cell Nuclei Segmentation**

Among the numerous studies aimed at improving the original U-Net<sup>[11][13]</sup> there has been considerable focus on stacking or cascading multiple U-Nets. Sevastopolsky et

al.<sup>[13]</sup> combine two types of building blocks, U-Net and Res-UNet, to segment images of the optic disc and cup image segmentation. Christ et al.<sup>[12]</sup> use a sequential arrangement of two U-Nets to segment both the liver and lesions in CT images. In [14], an ensemble of FCN-based methods was introduced to distinguish overlapped segmented nuclei. Although this approach achieved better results than basic encoder-decoder models, it struggled with more difficult cases<sup>[15]</sup>. We observe that existing methods are often dependent on specific architectures. For example, Jiang et al.<sup>[16]</sup> proposed a two-step approach using a modified U-Net model: initially, a simplified U-Net variant generates a preliminary prediction, and then, in the second step, the architecture is enhanced by adding two decoding layers to refine the prediction. In contrast, our approach is architecture-agnostic. It allows the use of any arbitrary architecture in both the global and local stages, providing flexibility and refinement for cell nuclei segmentation without being tied to a specific model structure.

### **D. Our Proposal**

The key trade-off for an accurate segmentation is managing resolution and context. While per-pixel predictions can give a high-resolution mask, clearly delineating the boundaries, the context at individual pixel is often not enough to label. Giving context from a patch helps predict the label accurately but sacrifices the resolution. Our key observation is that the current two-stage state-of-the-art techniques which give binary segmented mask in the first stage do provide the necessary context, but by providing the full image in the second stage, doesn't allow the model to focus on high resolution which is crucial for accurate boundary segmentation. Hence, in this work, we propose a simple, yet highly effective technique which does coarse level segmentation in the first stage by using the full image. However, in the second stage, we crop the images according to the mask generated in the first stage, and then give high resolution images to the network in the second stage for highly accurate boundary prediction. Figure 1 visually describes our approach. Our main contributions are:

- We propose a novel two stage global-local learning approach for cell nuclei segmentation which effectively combines contextual cues with high resolution features for accurate boundary segmentation.
- The proposed pipeline is architecture agnostic, and can combine multiple latest evolving segmentation techniques in the two stages.
- Our proposed framework outperforms existing state-of-the-art techniques on four publicly available datasets.

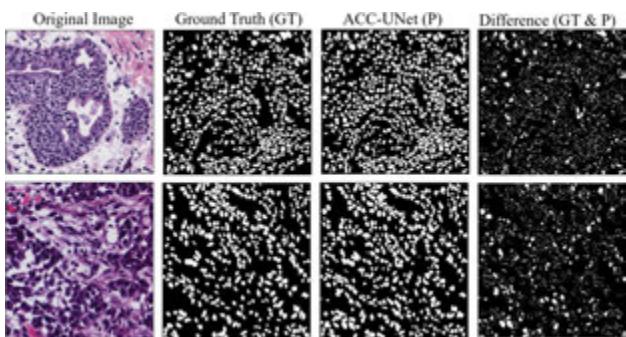


Fig. 2 Limitations in Nuclei Boundary Segmentation: Implications for Disease Diagnosis

## 2 PROPOSED APPROACH

Nuclei segmentation is important in medical imaging because it helps to clearly identify and outline the nuclei of cells in complicated tissue images. Several methods are used for nuclei segmentation<sup>[4]-[6][10]</sup> and while they effectively locate nuclei, they often struggle with precisely defining their boundaries (see Figure 2). This limitation raises concerns about their overall effectiveness. To address this critical aspect of biomedical image processing there is a dire need of refinement. In this work, we argue that providing localized features, along with the entire image can enhance the boundary delineation. The architecture of our proposed approach is illustrated in Figure 3. In this study, we propose a novel approach to nuclei segmentation using a two-stage learning framework called Global-Local (GoLo). The first stage is Global, and the second stage is Local. In a two-stage learning framework, we use the same segmentation model twice in sequence, with the second stage refining the results of the first one. The pipeline is designed such that the input image fed into the model and the resulting

segmented binary mask are of same original size.

In Stage-1 (Global), a U-Net based model is employed to globally process the histopathological images, aiming to predict the coarse binary segmentation map. To train this model, a combination of binary cross entropy and dice loss functions is utilized, facilitating the optimization process by penalizing deviations between the predicted and ground truth segmentation masks. However, this stage excels in accurately localizing nuclei, but have limitations with accurately delineating fine boundaries. To address this, the resulting segmentation mask from Stage-1 is resized to match the original image size ( $N \times N$ ) and then concatenate it with the input image. This process creates a four-channels image, incorporating both the original input image and the predicted segmentation mask of Stage-1. The four channel image consist of localization information into the input from Stage-1, enhancing the model's ability to capture spatial relationships between nuclei and their surroundings. Subsequently, this four-channel image is cropped into patches of size  $100 \times 100$ , serving as preprocessing for the Stage-2 model. In Stage-2 (Local), the same U-Net based model is employed, receiving the cropped patches of four-channel images as input. With the localization of cell nuclei already provided from first stage output as one of the channels in input, the primary focus of this model is to precisely delineate the boundaries of nuclei. Leveraging the rich spatial context encoded in the four channels, the model works to refine the segmentation by capturing fine details and accurately delineating boundaries. The predicted segmentation masks, resized back to  $100 \times 100$  patches, are then stitched together to reconstruct the full segmented image (see Figure 3). This stitching process ensures that the model effectively integrates information from multiple patches to produce a final binary segmented mask with smoother and more accurate boundaries. This refined segmentation is essential for enabling precise diagnosis and analysis in medical imaging applications. Final binary segmentation map is obtained by stitching together all the patches generated by the second stage. Formally, let's define some terms:

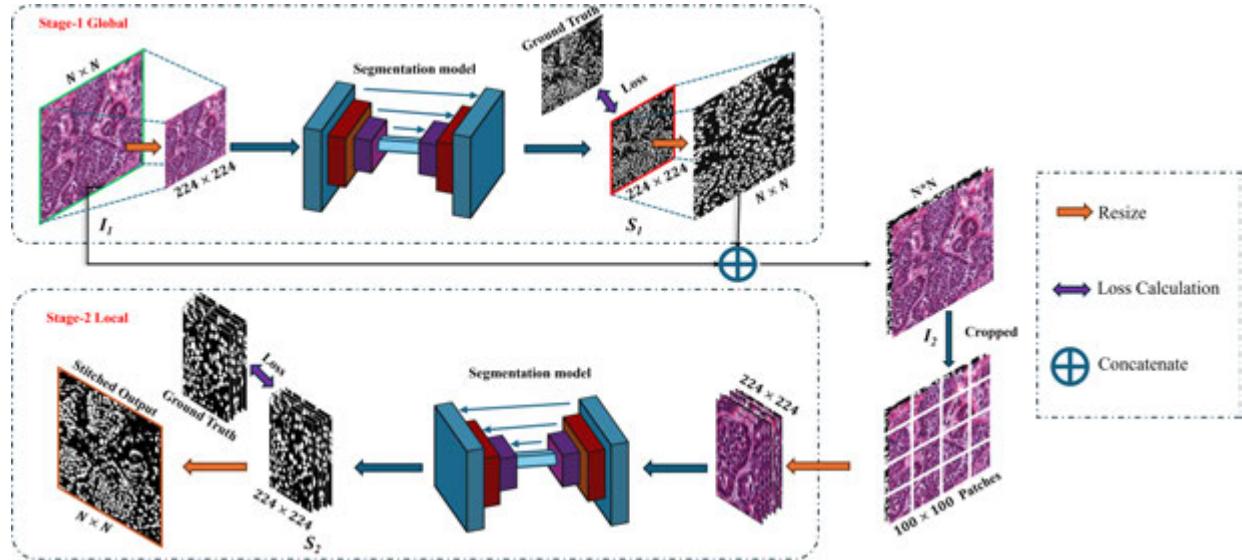


Fig. 3 Proposed Approach: Two Stages, Global- local features extractions (GoLo)

Table 1 COMPARATIVE QUANTITATIVE RESULTS FOR CELL NUCLEI SEGMENTATION

Methods	Approach	Datasets		MoNuSeg		CPM-17		CoNSeP		TNBC	
		mDSC	mIoU	mDSC	mIoU	mDSC	mIoU	mDSC	mIoU	mDSC	mIoU
UNet	Existing	69.81	55.12	83.55	72.00	74.21	59.18	70.77	54.35		
	Proposed	74.11	60.67	87.05	77.23	77.31	63.14	74.08	59.39		
Swin-UNet	Existing	65.58	49.81	79.08	69.60	61.00	41.23	66.75	50.35		
	Proposed	73.79	60.17	85.08	74.23	76.91	62.64	72.61	57.27		
UCTransnet	Existing	74.30	60.07	85.67	75.18	67.86	51.68	75.28	60.66		
	Proposed	77.38	63.80	86.85	76.94	74.58	59.68	74.67	59.85		
ACC-UNet	Existing	75.04	62.41	83.94	72.54	74.03	58.92	73.35	58.18		
	Proposed	77.93	65.93	86.87	76.97	75.69	61.01	77.74	57.51		

- $I_1$  represents an input image. In our case, it belongs to the RGB image domain, denoted as  $I_1 \in R^{w \times h \times 3}$ , where  $w$  and  $h$  represent the width and height of the image, respectively.
- $S_1$  denotes the binary segmentation map obtained from the first stage. It belongs to a set of possible segmentation maps  $S_1 \in \{0, 1\}^{224 \times 224 \times 1}$ . It is resized to  $w \times h \times 1$
- $I_2 \in R^{w \times h \times 4}$  which is the channel-wise concatenation of resized  $S_1$  and  $I_1$  both of size  $w \times h$ . It is then cropped into patches of  $100 \times 100$ .
- $S_2$  represents the binary segmentation map generated from the Stage 2. It belongs to a set of binary maps  $\in \{0, 1\}^{224 \times 224 \times 1}$  which again resized to  $100 \times 100$ . These patches are then stitched together to form the final binary segmentation map belongs to a set  $\{0, 1\}^{w \times h \times 1}$

In the two-stage learning framework: 1. First Stage (Global):

This stage takes the original image  $I_1$  as input to UNet based model<sup>[9][17]-[19]</sup>, generating a binary segmentation map  $S_1$ , which highlights the locations of the nuclei ignoring the fine boundaries. 2. Second Stage (Local): This stage uses the patches of  $I_2$  as input to U-Net based model, the output segmentation maps are stitched together to create the full binary segmentation map. The stitched segmented map has more accurate nuclei boundaries.

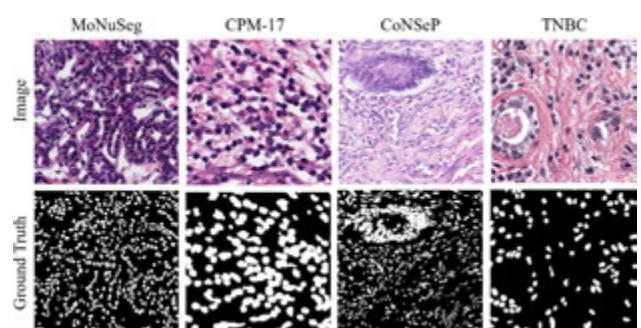


Fig. 4 Samples from different datasets

## A. Data Processing and Network Training

1) *Datasets Description:* To assess the effectiveness of the Global-Local training approach, we carried out experiments using four publicly available datasets: Multi-Organ Nucleus Segmentation<sup>[17]</sup>, CPM-17<sup>[20]</sup>, Colorectal Nucleus Segmentation and Phenotype<sup>[21]</sup>, and Triple Negative Breast Cancer Dataset<sup>[22]</sup> each targeting different types of cell nuclei. *Multi-organ Nuclei Segmentation (MoNuSeg)*<sup>[17]</sup> dataset contains 44 image tiles, each sized  $1,000 \times 1,000$  pixels. It contains Whole Slide Images (WSIs) of 7 organs (breast, kidney, colon, stomach, prostate, liver, and bladder) from various medical centres (i.e., various stains) of high-resolution WSI of H&E-stained slides from nine tissue types, digitised at  $40 \times$  magnification in eighteen different hospitals and obtained from National Cancer Institute's cancer Genome Atlas (TCGA). *CPM-17*<sup>[20]</sup> dataset consists of 32 images obtained from the University Hospitals Coventry and Warwickshire (UHCW). These images were captured using magnifications of  $40 \times$  and  $20 \times$ , with sizes ranging from  $500 \times 500$  to  $600 \times 600$  pixels. Within this dataset, a total of 7,570 nuclei have been meticulously annotated and labeled. It offers tissue images with annotated labels for nuclei segmentation and classification, sourced from patients diagnosed with head and neck squamous cell carcinoma (HNSCC), glioblastoma multiforme (GBM), non-small cell lung cancer (NSCLC), and lower-grade glioma (LGG) tumors. These nuclei span across four distinct cancer types, providing a diverse set of cellular structures for analysis and segmentation purposes. The *CoNSeP* dataset<sup>[21]</sup> contains 41 image tiles, each sized  $1000 \times 1000$  pixels, obtained from colorectal adenocarcinoma (CRA) WSIs. These images were scanned at  $40 \times$  objective magnification using an Omnyx VL120 scanner at the pathology department of University Hospitals Coventry and Warwickshire, UK. Naylor et al.<sup>[22]</sup> introduced a dataset called *Triple Negative Breast Cancer (TNBC)*. The dataset consists of 50 H&E-stained images, each with a resolution of  $512 \times 512$  pixels, and includes a total of 4022 annotated nuclei. These images are extracted from histopathological samples of 11 patients diagnosed with triple-negative breast cancer. Dataset includes various cell types such as normal epithelial and myoepithelial

breast cells, invasive carcinomatous cells, fibroblasts, endothelial cells, adipocytes, macrophages, and inflammatory cells, providing a comprehensive representation of the histological features observed in TNBC samples. To train the network on these datasets we utilized random flip and rotation augmentation.

2) *Training Loss:* The proposed approach is dedicated to semantic segmentation, and its loss function is constructed using both Binary Cross-Entropy (BCE) loss and Dice loss. These two are commonly used loss functions in biomedical image segmentation. BCE evaluates class predictions for each pixel individually, while Dice loss quantifies the overlap between two samples. To optimize performance and achieve early convergence in our task, we utilize a BCE-Dice loss function, which combines binary cross-entropy with Dice loss, resulting in improved performance and faster convergence. We formally define the loss function as Overall loss,  $L_O$

$$L_O = 0.5L_{BCE} + 0.5L_{DICE} \quad (1)$$

$$L_{BCE} = - \left[ \frac{1}{n} \sum_{i=1}^{N*K} X_i \log Y_i + \sum_{i=1}^{N*K} (1 - X_i) \log (1 - Y_i) \right] \quad (2)$$

$$L_{DICE} = 1 - \frac{2 \sum_{i=1}^{N*K} X_i Y_i + \epsilon}{\sum_{i=1}^{N*K} Y_i + \sum_{i=1}^{N*K} X_i + \epsilon} \quad (3)$$

Within equation 2 and 3,  $X_i$  represent the ground truth mask and  $Y_i$  represent the segmented mask.  $n$  denotes total number of nuclear pixels within the image and  $N*K$  denotes the set containing all pixels.

## 3

## RESULTS AND DISCUSSION

The proposed approach is quantitatively compared with existing approaches (UNet<sup>[9]</sup>, Swin-Unet<sup>[23]</sup>, UCTransnet<sup>[19]</sup>, and ACC-UNet<sup>[18]</sup>) in terms of mean Dice Similarity Coefficient (mDSC) and mean Intersection over Union (mIoU) on the MoNuSeg, CPM-17, CoNSeP, and TNBC datasets. Notably, our approach demonstrates superior mean DSC and mean (IoU) metrics, in comparison to

baseline methods indicating its effectiveness for nuclei segmentation tasks (see Table I). To ensure a fair comparison, we adhered to the same training and testing

set split criteria as employed for the baseline methods.

Furthermore, in Figure 5 we provide visual comparisons

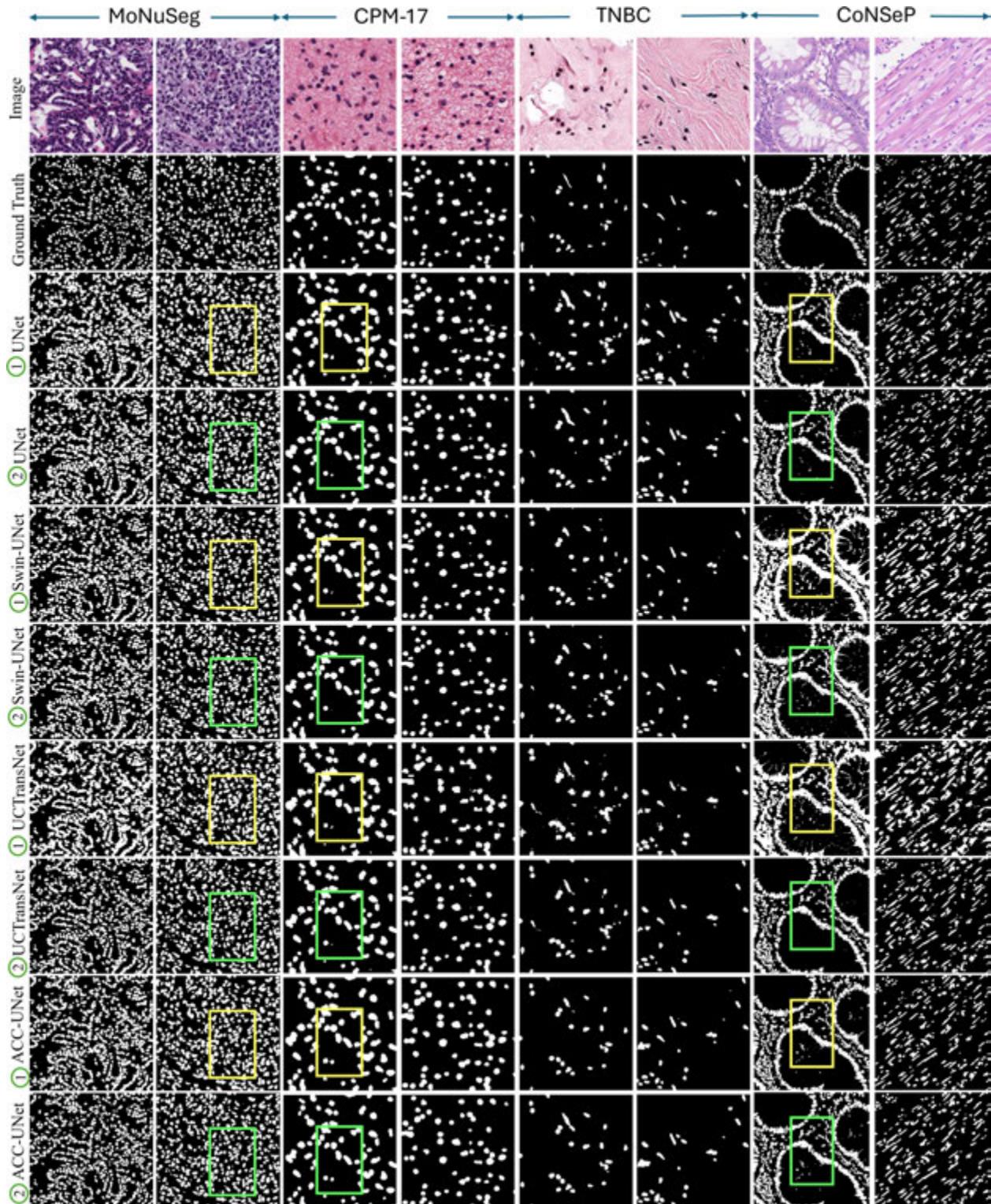


Fig. 5 Comparative results of nuclei segmentation on MoNuSeg, CPM-17, TNBC and CoNSeP dataset. Input image with their corresponding ground truth along with the segmentation results of existing methods (single-stage ①) and Proposed approach (GoLo ②)

between the segmentation results of our approach and the existing methods for MoNuSeg, CPM-17, CoNSEP, and TNBC datasets. For the MoNuSeg dataset, our approach refine the boundaries in comparison to existing methods (see box in Figure 5). Figure 5 show results on CoNSEP dataset, our prediction filtered out the apparent nodule region, which was predicted as a false positive by all the other models (see green box which signify the improvement in comparison to yellow box). In the CPM-17 and TNBC dataset, we not only successfully predicted the localization of nuclei but also identified fine boundaries that were mostly missed by the existing models. As clearly depicted in Figure 5 our Global-Local approach effectively delineates nuclei boundaries, outperforming the baseline models. Overall, our proposed model achieves state-of-the-art accuracy in nuclei segmentation by leveraging auxiliary information. This approach holds promise for direct deployment in cell pathology diagnosis systems, aiming to alleviate the workload of pathologists. It is important to note that our approach demonstrates a high level of generalization across varying tissue and cell types.

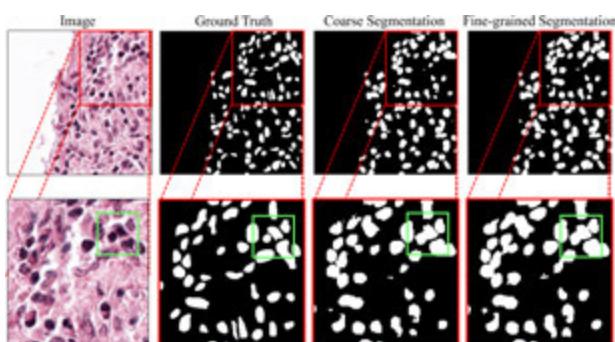


Fig. 6 Limitation of our proposed approach

#### A. Limitations of Proposed Approach

Though the proposed architecture offers significant benefits towards boundary delineation of nuclei, it performs poorly for touching nuclei with the unclear or blurred boundary (see the bounding box in the cropped portion of Figure 6). A traditional approach to address the touching cell nuclei problem involves applying a distance transform to the segmentation mask, followed by using a watershed algorithm. Although this method often proves ineffective for highly overlapping cases, this

observation will be considered as the future scope to further improve the segmentation accuracy of proposed approach.

## 4 CONCLUSION

In this study, we propose a novel two-stage learning approach called GoLo for cell nuclei segmentation in histopathology images, representing a significant advancement in addressing the nuclei boundary segmentation challenge. Breaking down the segmentation process into two distinct steps allows us to leverage global features for preliminary nucleus localization and boundary detection, followed by the refinement of boundaries using local features and additional information obtained in the first stage. In the end, our concept is architecture agnostic, universal which can be readily applicable to various histopathology image analysis applications.

## REFERENCES

- [1] H. Wang, M. Naghavi, C. Allen, R. M. Barber, Z. A. Bhutta, A. Carter, D. C. Casey, F. J. Charlson, A. Z. Chen, M. M. Coates et al., "Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the global burden of disease study 2015," *The Lancet*, vol. 388, pp. 1459-544, 2016.
- [2] W. He, T. Liu, Y. Han, W. Ming, J. Du, Y. Liu, Y. Yang et al., "A review: The detection of cancer cells in histopathology based on machine vision," *Computers in Biology and Medicine*, vol. 146, p. 105636, 2022.
- [3] W. Hsu, S. X. Han, C. W. Arnold, A. A. Bui, and D. R. Enzmann, "A data-driven approach for quality assessment of radiologic interpretations," *Journal of the American Medical Informatics Association*, vol. 23, no. e1, pp. e152-e156, 2016.
- [4] R. M. Thomas and J. John, "Detection and segmentation of mitotic cell nuclei in breast histopathology images," in *2017 International Conference on Networks & Advances in Computational Technologies (NetACT)*. IEEE, 2017, pp. 246-250.

- [5] A. Das, M. S. Nair, and S. D. Peter, "Computer-aided histopathological image analysis techniques for automated nuclear atypia scoring of breast cancer: a review," *Journal of digital imaging*, vol. 33, no. 5, pp. 1091-1121, 2020.
- [6] T. Wan, L. Zhao, H. Feng, D. Li, C. Tong, and Z. Qin, "Robust nuclei segmentation in histopathology using asppu-net and boundary refinement," *Neurocomputing*, vol. 408, pp. 144-156, 2020.
- [7] N. Kumar, R. Verma, S. Sharma, S. Bhargava, A. Vahadane, and A. Sethi, "A dataset and a technique for generalized nuclear segmentation for computational pathology," *IEEE Transactions on Medical Imaging*, vol. 36, no. 7, pp. 1550-1560, 2017.
- [8] P. Naylor, M. Laé, F. Reyal, and T. Walter, "Segmentation of nuclei in histopathology images by deep regression of the distance map," *IEEE Transactions on Medical Imaging*, vol. 38, no. 2, pp. 448-459, 2019.
- [9] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention-MICCAI 2015: 18th international conference*, vol. 18, 2015, pp. 234-241.
- [10] F. Xing and L. Yang, "Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: a comprehensive review," *IEEE Reviews in Biomedical Engineering*, vol. 9, pp. 234-263, 2016.
- [11] L. Bi and et al., "Dermoscopic image segmentation via multistage fully convolutional networks," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 9, pp. 2065-2074, 2017.
- [12] P. Christ and et al., "Automatic liver and lesion segmentation in ct using cascaded fully convolutional neural networks and 3d conditional random fields," in *MICCAI*. Springer, 2016, pp. 415-423.
- [13] A. Sevastopolsky and et al., "Stack-u-net: Refinement network for image segmentation on the example of optic disc and cup," *arXiv preprint arXiv:1804.11294*, 2018.
- [14] P. Naylor, M. Laé, F. Reyal, and T. Walter, "Nuclei segmentation in histopathology images using deep neural networks," in *14th IEEE International Symposium on Biomedical Imaging*, 2017, pp. 933-936.
- [15] ——, "Segmentation of nuclei in histopathology images by deep re-gression of the distance map," *IEEE Transactions on Medical Imaging*, vol. 38, no. 2, pp. 448-459, 2019.
- [16] Z. Jiang, C. Ding, M. Liu, and D. Tao, "Two-stage cascaded u-net: 1st place solution to brats challenge 2019 segmentation task," in *International MICCAI Brainlesion Workshop*, 2019, pp. 231-241.
- [17] N. Kumar, R. Verma, D. Anand, Y. Zhou, O. F. Onder, E. Tsougenis, H. Chen, P.-A. Heng, J. Li, Z. Hu et al., "A multi-organ nucleus segmentation challenge," *IEEE transactions on medical imaging*, vol. 39, no. 5, pp. 1380-1391, 2019.
- [18] N. Ibtehaz and D. Kihara, "Acc-unet: A completely convolutional unet model for the 2020s," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer Nature Switzerland, 2023, pp. 692-702.
- [19] H. Wang, P. Cao, J. Wang, and O. R. Zaiane, "Uctransnet: rethinking the skip connections in u-net from a channel-wise perspective with trans- former," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 36, no. 3, 2022, pp. 2441-2449.
- [20] Q. Vu, S. Graham, T. Kurc, M. To, M. Shaban, T. Qaiser, N. Koohbanani, S. Khurram, J. Kalpathy-Cramer, T. Zhao, R. Gupta, J. Kwak, N. Rajpoot, J. Saltz, and K. Farahani, "Methods for segmentation and classification of digital microscopy tissue images," *Frontiers in Bioengineering and Biotechnology*, vol. 7, p. 53, 2019. [Online]. Available: <https://doi.org/10.3389/fbioe.2019.00053>
- [21] S. Graham, Q. Vu, S. Raza, A. Azam, Y. Tsang, J. Kwak, and N. Rajpoot, "Hover-net: simultaneous segmentation and classification of nuclei in multi-tissue histology images," *Medical Image Analysis*, vol. 58, p. 101563, 2019. [Online]. Available: <https://doi.org/10.1016/j.media.2019.101563>
- [22] P. Naylor, M. Laé, F. Reyal, and T. Walter, "Segmentation of nuclei in histopathology images by deep regression of the distance map," *IEEE Transactions on Medical Imaging*, vol. 38, pp. 448-459, 2019. [Online]. Available: <https://doi.org/10.1109/TMI.2018.2865709>
- [23] H. Cao, Y. Wang, J. Chen, D. Jiang, X. Zhang, Q. Tian, and M. Wang, "Swin-unet: Unet-like pure transformer for medical image segmenta- tion," in *European conference on computer vision*. Springer, 2022, pp. 205-218.

■著者



**Amit Shakya**  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India  
Department of Biomedical  
Engineering, Indian Institute  
of Technology Ropar, India

**Revat Saharan**  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India

**Chetan Gupta**  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India



**Rupesh Kumar**  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India

**Lalit Sharma**  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India

**Srivatsava Naidu**  
Department of Biomedical  
Engineering, Indian Institute  
of Technology Ropar, India



**Subrahmanyam Murala**  
School of Computer Science  
and Statistics, Trinity College  
Dublin, Ireland

**Chetan Arora**  
Department of Computer  
Science and Engineering,  
Indian Institute of Technology  
Delhi, India

# Do not look so locally to fish skins: Improved YOLOv7 for fish disease detection with Transformers

Shruti Phutke Amit Shakya Chetan Gupta Rupesh Kumar Tsuyoshi Kuroda Lalit Sharma

This paper, titled “Do not look so locally to fish skins: Improved YOLOv7 for fish disease detection with Transformers,” was presented at CVIP-2024 (9th International Conference on Computer Vision & Image Processing), held at IIITDM Kancheepuram, Chennai, India, from December 19–21, 2024.

Reprinted with permission. Copyright © 2024 CVIP and IIITDM.

Further use or distribution is not permitted without permission from CVIP and IIITDM.

## Abstract

Aquaculture production significantly influences overall fish production, yet it is often adversely affected by various fish diseases. These diseases can be effectively identified by analyzing the condition of the fish's skin. Consequently, there is a growing demand for automated fish skin disease detection methods. By implementing such automated approaches, the efficiency and accuracy of disease detection can be enhanced, leading to better management of fish health and, ultimately, more sustainable aquaculture practices. In this work, we propose a novel Transformer based modified YOLO approach for detection of five different fish skin diseases. We propose a Transformer feature extraction module (TFEM) to effectively capture the long-range dependencies from input image. The proposed TFEM is incorporated in the YOLOv7 back-bone for efficient feature learning. We assessed the performance of our proposed TFEM by comparing it with various YOLOvX approaches to confirm its effectiveness. Both qualitative and quantitative results demonstrate that our method is highly capable of accurately detecting five distinct fish diseases. The source code is available at: [https://github.com/shrutiphutke/Fish\\_disease\\_detection\\_YOLO\\_transformer](https://github.com/shrutiphutke/Fish_disease_detection_YOLO_transformer)

## 1 INTRODUCTION

With the world's population steadily increasing, the demand on traditional food sources is mounting, highlighting the need for sustainable alternatives. Aquaculture emerges as a key solution, offering a dependable source of high-protein food while reducing pressure on marine and terrestrial

ecosystems<sup>[1]</sup>. This not only bolsters food security but also enhances the stability of global food systems. In aquatic food resources, fish food is considered as key resource worldwide in turn becoming the most important factor of the economy. The affected fish due to different diseases causes the limited availability of good quality aquatic food and thus impacts the financial outcomes.

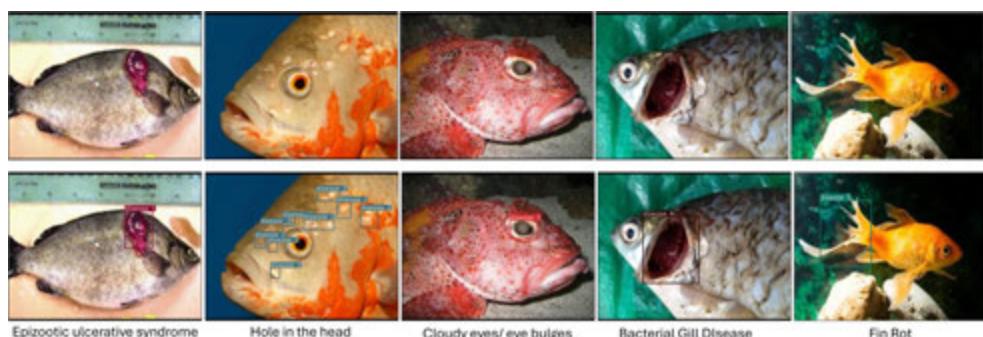


Fig. 1 Sample images of different fish skin diseases

There are various fish diseases caused due to poor water quality, bacterial infection, anchor worms, improper-diet, over-crowding of fish in the tank, etc. The diseased fish generally shows different symptoms such as behavioural or skin color/texture changes. The existing methodology for fish disease detection use expert base system<sup>[29]</sup>, microscopic images<sup>[16]</sup>, fluorescent images<sup>[26]</sup>, ultra sound images<sup>[14]</sup>, body surface images<sup>[25]</sup>, etc. Identifying the type of diseases on the basis of skin abnormalities using body surface images provides the high detection efficiency as compared to disease detection based on the behaviour. In this regard, researchers have proposed different approaches for fish disease detection<sup>[11]</sup>.

In addition to traditional hand-crafted methods, various deep-learning approaches have been proposed for the detection<sup>[27][28]</sup> and classification<sup>[12][4]</sup> of fish diseases. These methods encompass a range of techniques, including those that focus on identifying abnormalities in fish behavior and those that analyze skin abnormalities for disease detection and classification. By leveraging deep learning, these approaches aim to enhance the accuracy and efficiency of diagnosing fish diseases, offering more sophisticated and automated solutions compared to conventional methods.

Researchers have primarily focused on the classification of fish diseases rather than on the detection and localization of these diseases<sup>[13]</sup>. Within this focus, there are numerous approaches developed for identifying various fish diseases<sup>[12][4][20]</sup>. However, there are relatively few methods dedicated specifically to the localization of diseases<sup>[25]</sup>. This emphasis on classification highlights the importance of accurately identifying the type of disease affecting fish, while the challenge of pinpointing the exact location of the disease remains less explored.

Existing approaches for detecting fish skin diseases have employed various YOLOvX architectures on fish disease datasets to deliver detection results. These models have been applied directly to the datasets, yielding insights into the presence and extent of skin diseases in fish. In this study, we introduce a modified YOLOv7 approach,

distinguishing it from existing methods, with the transformer based backbone for efficient feature extraction in turn helping for efficient fish disease detection. We considered five different fish skin diseases such as epizootic ulcerative syndrome, holes on the head, cloudy eyes/eye bulges, bacterial gill disease, fin rot, etc. The sample images for each type of disease are shown in Figure 1. The contributions of our work are summarized as:

- We propose a novel modified Yolov7 approach for fish skin disease detection.
- A transformer feature extraction module is proposed in YOLO backbone to extract the efficient features for fish disease detection.
- The extensive quantitative and qualitative comparison is performed to verify the effectiveness of proposed approach for fish disease detection.

## **2 RELATED WORKS**

In this section, we discuss the different approaches proposed for fish skin disease classification and detection.

### **2-1. Fish Skin Disease Classification**

Segmenting fish skin based on texture involves using various clustering techniques to differentiate and categorize different regions of the skin. This process is followed by extracting features from the segmented images, which captures important information about the texture and appearance of the skin. When these extracted features are analyzed using machine learning methods like Support Vector Machines (SVM), it has been shown to enhance the accuracy and effectiveness of fish disease detection<sup>[12][4]</sup>. This approach leverages the ability of clustering to preprocess and organize data, and the power of SVM to classify and identify potential diseases based on the features extracted from the segmented images. This combination has demonstrated success in improving the detection of fish diseases, providing a robust framework for analyzing and managing fish health. In this regard, Sikder et al.<sup>[20]</sup> used a Fuzzy C-means approach for clustering followed by Gray level

co-occurrence metrics (GLCM) based features for classifying six different fish skin diseases. A similar GLCM feature based approach is proposed in [2]. Mia et al.<sup>[13]</sup> used a K-means clustering as a pre-processing step and extracted different features such as mean, standard deviation, etc. to classify the three different fish diseases using various machine learning approaches.

The convolutional neural networks (CNNs), with their end-to-end effective feature extraction ability have been widely utilized for fish disease detection. Gupta et al.<sup>[7]</sup> proposed a 15 layer CNN approach for wound and lice detection in fish skin. Chen et al.<sup>[5]</sup> experimented the fish diseased detection using different CNN approaches such as ResNet, DenseNet, etc. Wang et al.<sup>[25]</sup> analyzed different color spaces of the input image such as RGB, YCbCr, XYZ and proposed a AlexNet and ResNet based approach for classifying three different fish diseases. Further, Azhar et al.<sup>[3]</sup> used a GoogleNet like architecture to classify the white spot fish disease. A similar, CNN based approach is proposed in [8] to classify white spot, red spot and healthy fish. a fine-tuning based approach is proposed in [15] for fish disease classification.

## 2-2. Fish Skin Disease Detection

Building on the success of classifying fish diseases using various machine learning and convolutional neural network (CNN)-based approaches, researchers have turned their attention to the localization of fish skin diseases. This shift has led to the development of several new techniques aimed at pinpointing the exact locations of diseases on fish skin. Localization involves not only identifying the presence of a disease but also determining its specific area or region, which is crucial for accurate diagnosis. These emerging approaches combine advanced image analysis and deep learning methods to enhance the precision of disease detection, providing more detailed and actionable insights into the distribution and severity of skin conditions. As a result, these localization techniques represent a significant advancement in the field, improving the management of fish diseases. Yasruddin et al.<sup>[27]</sup> pioneered the use of Faster Region-based Convolutional Neural Networks (Faster R-CNN) for

localizing fish diseases. Yu et al.<sup>[28]</sup> proposed a Modified YOLOv4 based approach by considering different backbones such as MobileNet, CSPDarkNet, etc. Wang et al.<sup>[25]</sup> proposed an improved YOLOv5 network consisting of channel and spatial attention mechanism for underwater fish disease detection. Given its effectiveness in capturing global dependencies, researchers have begun applying it to fish disease detection. In [30], authors proposed an improved YOLOv5 based approach by replacing on convolution module in backbone with MobileNetVit (a single transformer block) for fish disease detection based on abnormal behaviour. As seen from the literature, there are very few approaches for fish skin disease detection. Also, considering the existing YOLO based approaches providing sufficient ability for fish skin disease detection, there is a dire need of efficient feature extraction module in the backbone of YOLOv7 for fish skin disease detection. In regards to this, we proposed an efficient transformer based feature extraction module for fish skin disease detection.

## 3

## PROPOSED METHOD

From the existing literature, it is evident that YOLO-based approaches have been employed for fish skin disease detection. However, these methods often utilize the YOLO models directly without accounting for the specific details or characteristics of the input images. As seen from the Figure 1, we can see that different type of skin disease has unique properties such as for Epizootic ulcerative syndrome (EUS) the fish may have wound covering only a smaller region or while body, the hole in head disease has many locations present on the head of fish, etc. The existing approaches may fail at detecting these disease efficiently due to their localized visual ability due to the usage of convolution operation for feature extraction. Also, this oversight may impact the accuracy of disease detection and localization, suggesting a need for more tailored approaches that consider these nuances. Taking into account the ability of transformers to capture global dependencies effectively<sup>[23][17]</sup>, we propose the transformer based feature extraction module in the backbone of YOLOv7<sup>[24]</sup>.

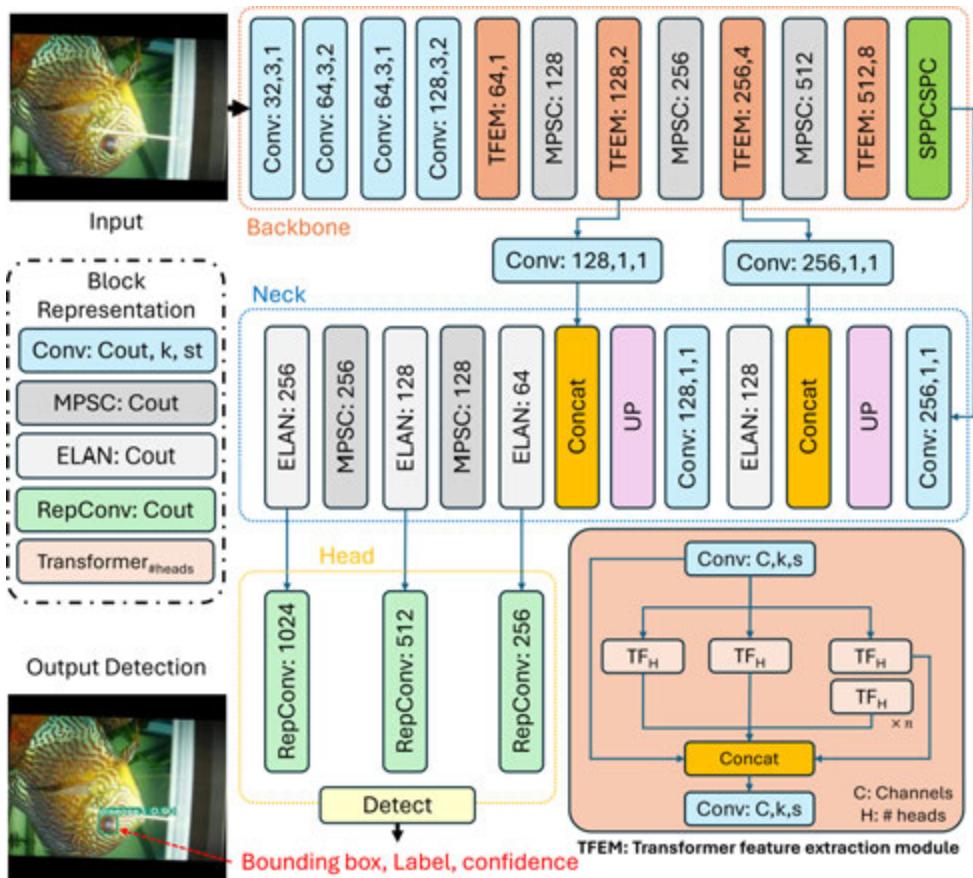


Fig. 2 Proposed architecture for fish disease detection. This is a YOLOv7 based architecture consisting of Backbone, Neck and Head. We propose a Transformer based feature extraction module (TFEM) for efficient feature extraction while capturing long-range dependencies. Further, the architecture consists of different blocks such as: MPSC- Maxpooling-strided convolution, ELAN- Efficient layer aggregation network, UP- upsampling layer, RepConv- Represented convolution, etc.

The proposed YOLOv7 based approach consists of different feature extraction parts such as a backbone, a neck and the head (refer Figure 2). The backbone part mainly contributes in extracting the efficient features from the input image. This plays an important role in capturing the textures, edges, structures, and the complex details from the image. The neck part is mainly responsible to refine and process the features from backbone. With the efficiently extracted features from backbone, the neck part contributed to better detection and localization of diseased region. Further, the head part converts the features into the bounding box co-ordinates, disease class and the confidence score of prediction. By considering the necessity to provide an effective feature extraction module in the backbone, we provided the Transformer feature extraction module (TFEM) in backbone of disease detection network.

### 3-1. Overview

The input image with size  $M \times N \times 3$  is fed as input to the backbone of the network. This image is processed through three consecutive convolution layers to transform it from spatial domain to feature domain. Further, extracted features are down-sampled using max-pooling and stride convolution (MPSC) to a size  $\frac{M}{2} \times \frac{N}{2} \times 3C$  and forwarded to the proposed TFEM. The TFEM processes

The features and extract the features by capturing the long-range dependencies with respect to the input. These features are again down-sampled using MPSC to a size  $\frac{M}{4} \times \frac{N}{4} \times 2C$ . The successive four TFEM blocks with number of heads  $H \in (1, 4)$  are used providing the feature map of size  $\frac{M}{16} \times \frac{N}{16} \times 5C$ . These features are then fed to spatial pyramid pooling cross stage partial connection (SPPCSPC) block which deals with the multi-scale feature processing (refer backbone in Figure 2).

The processed features from the backbone are then fed to neck part of the architecture. The neck part has up-sampling followed by efficient layer aggregation network (ELAN) block. The ELAN block is responsible for efficiently merging the features from the backbone to learn the high-contextual information. The features from last two TFEM module are merged and processed in the neck using the ELAN block producing the feature maps of size  $\frac{M}{4} \times \frac{N}{4} \times 2C$ . After this, the features are successively down-sampled using MPSC and processed in ELAN block in order to forward them to the detection head. The features of size,  $\frac{M}{4} \times \frac{N}{4} \times 2C$ ,  $\frac{M}{8} \times \frac{N}{8} \times 3C$ ,  $\frac{M}{8} \times \frac{N}{8} \times 4C$  are utilized to forward to the detection head.

The detection head consists of represented convolution (RepConv) layers for generating the final detection outcomes in terms of bounding box coordinates, disease labels/type, and confidence score. These predicted bounding box coordinates are then appended on the input to localize the disease predicted by the detection head with respective confidence score (see Output Detection in Figure 2).

### 3-2. Transformer Feature Extraction Module

The architecture of proposed transformer feature

extraction module is shown in Figure 2 (refer TFEM). After processing the features from a convolution layer, they are fed to three different transformer blocks in parallel. Each of the transformer block consists of a multi-head attention followed by a feed forward network. The multi-head attention deals with capturing the long-range dependencies. The process of multi-head attention operation is given as:

$$\text{Attention} = \sigma \left( \frac{QK^T}{\sqrt{d}} \right) V \quad (1)$$

where,  $\sigma$  is Sigmoid activation function,  $d$  is a scaling parameter,  $Q$  is query,  $K$  is key, and  $V$

$$Q, K, V = \phi(\text{inputfeatures}) \quad (2)$$

where,  $\phi$  is a  $3 \times 3$  depth-wise separable convolution. Further, the feed forward network in transformer follows a gated attention mechanism as follows:

$$Y_{\text{out}} = \phi(\text{inputfeatures}) \cdot \text{Gelu}(\phi(\text{inputfeatures})) \quad (3)$$

where,  $\text{Gelu}$  is Gelu activation function. This TFEM process the input features efficiently by the use of stacked Transformer blocks which successively captures more receptive features with respect to the input. The details of the Transformer block, MPSC, SPPCSPC, ELAN, and RepConv are provided in Figure 3.

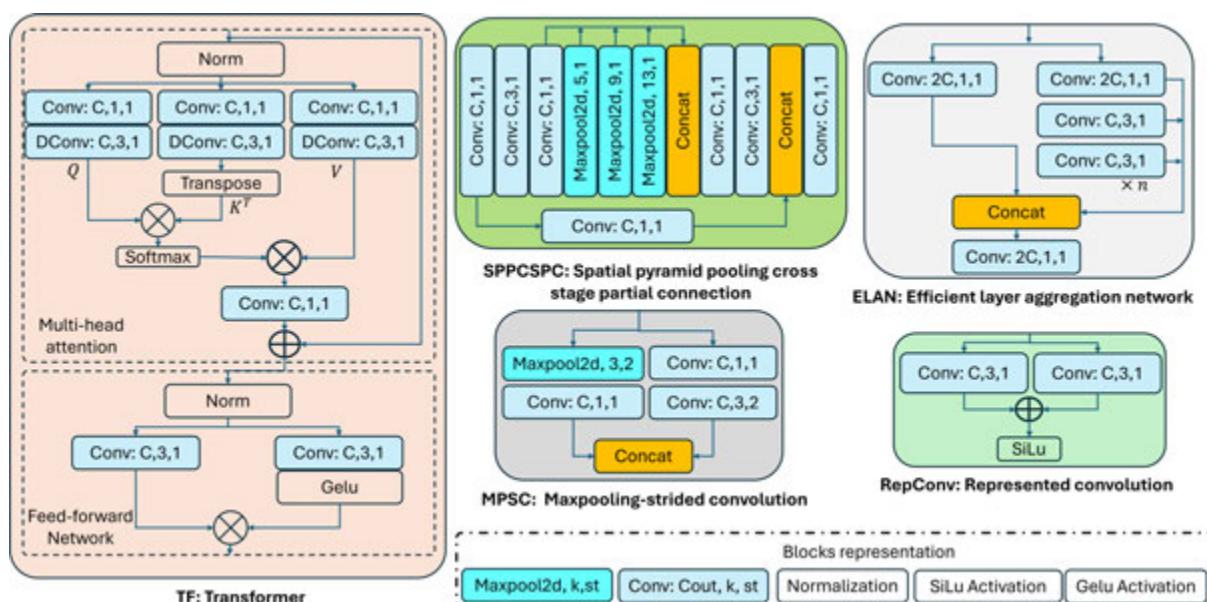


Fig. 3 Details of Transformer (TF), Max-pooling and stride convolution (MPSC), Spatial pyramid pooling cross stage partial connection (SPPCSPC), Efficient layer aggregation network (ELAN), and Represented convolution (RepConv) used in the proposed architecture in Figure 2 for fish skin disease detection

## 4 EXPERIMENTS AND RESULT DISCUSSION

This section details about the dataset utilized, evaluation measures, training and implementation, results and discussion, and limitations of proposed approach with future scope.

### 4-1. Dataset

In this work, we utilized a fish disease detection dataset collected from Roboflow. This dataset consists of 4419 images of diseased fishes which are then split into 3858, 381 and 180 for training, validation and testing respectively. The dataset has different augmentations such as horizontal and vertical flip, random rotations, randomly cropping of 0% to 3% of original images, and random shear, etc. This dataset considers five different fish diseases like: Epizootic ulcerative syndrome (EUS), hole/swelling in the head, cloudy eyes or eye bulges, bacterial gill disease and fin rot. The causes of different diseases are described as follows<sup>[19]</sup>:

**Epizootic ulcerative syndrome:** are sores that arise from inflammation of the external tissues. They can be caused by a range of factors, including physical injury and bacterial infections. Other contributing factors may include parasites, poor water quality, high ammonia levels, and low pH.

**Hole/swelling in the head:** The protozoan parasite Hexamita infects the intestines of fish that are already stressed due to overcrowding, poor water quality, or shipping. Parasite infections can advance from the intestines to affect the entire system. When parasites move to the sensory pores on the fish's head, they cause surrounding tissues to deteriorate and result in the development of deep lesions.

**Cloudy eyes/eye bulges:** It is caused by fluid accumulation behind the eye, but it is frequently the result of an opportunistic bacterial infection triggered by stress or poor tank conditions.

**bacterial gill disease** is generally caused due to ammonia poisoning in the new fish tank. Can be detected by red or inflamed areas around the gills.

**Fin rot** disease is generally caused due to poor water quality. This can be detected with the change in color of fins or tail, deterioration of fins and/or tail, frayed edges, inflammation at the base, etc.

### 4-2. Evaluation Measures

To assess the effectiveness of our proposed approach over existing state-of-the-art approaches, we employed several evaluation metrics: Recall, Precision, and mean Average Precision (mAP). Recall measures the ability of the approach to identify all relevant instances, Precision assesses the accuracy of the identified instances, and mean Average Precision (mAP) provides a comprehensive evaluation by averaging precision across different recall levels and multiple classes. Using these metrics allows us to thoroughly evaluate the performance and robustness of our method. In a multi-class imbalanced classification problem, precision is computed by dividing the total number of true positives across all classes by the combined total of true positives and false positives for all classes.

Table 1 Quantitative comparison of the proposed work and existing approaches

Method	Recall%	Precision%	mAP50%	mAP50:95%
YOLOv3 <sup>[18]</sup>	90.7	85.9	91.1	69.9
YOLOv5 <sup>[9]</sup>	90.3	<b>91.7</b>	92.9	70.3
YOLOv7 <sup>[24]</sup>	92.8	86.6	93.2	70.2
Ours	<b>94.2</b>	86.4	<b>93.4</b>	<b>71.4</b>

$$Precision = \frac{\sum_{c \in C} TruePositives_c}{\sum_{c \in C} (TruePositives_c + FalsePositives_c)} \quad (4)$$

In a multi-class imbalanced classification problem, recall is calculated by dividing the total number of true positives across all classes by the sum of true positives and false negatives for all classes.

$$Recall = \frac{\sum_{c \in C} TruePositives_c}{\sum_{c \in C} (TruePositives_c + FalseNegatives_c)} \quad (5)$$

$$mAP = \frac{1}{C} \sum_{c \in C} AP_c \quad (6)$$

where,  $AP$  is average precision,  $C$  total number of diseases considered (in this work  $C = 5$ ). The  $AP$  is given as:

$$AP = \sum_{k=0}^{th-1} [Recall_k - Recall_{k+1}] \times Precision_k \quad (7)$$

where,  $th$  is confidence score (here we considered case 1:  $th=0.5$  and case 2:  $th=0.5 - 0.9$ ),  $Recall_{th}=0$ ,  $Precision_{th}=1$ .

### 4-3. Training and Implementation

All the baselines and the proposed approach are trained

and tested on the images with  $640 \times 640$  resolution with *batch size* = 16 for 200 epochs. The networks are trained end-to-end from scratch and the parameters are optimized using Adam optimizer<sup>[10]</sup> with an initial learning rate of 0.001. A localization loss, confidence loss and classification loss functions are used to train the networks<sup>[22]</sup>. The Exponential moving average (EMA) approach<sup>[21]</sup> with a decay of 0.999 is used to train the networks. The training is carried out on NVIDIA A100 GPU. For fair comparison, we trained all the existing approaches and proposed approach with the same training/testing setup and dataset splits.

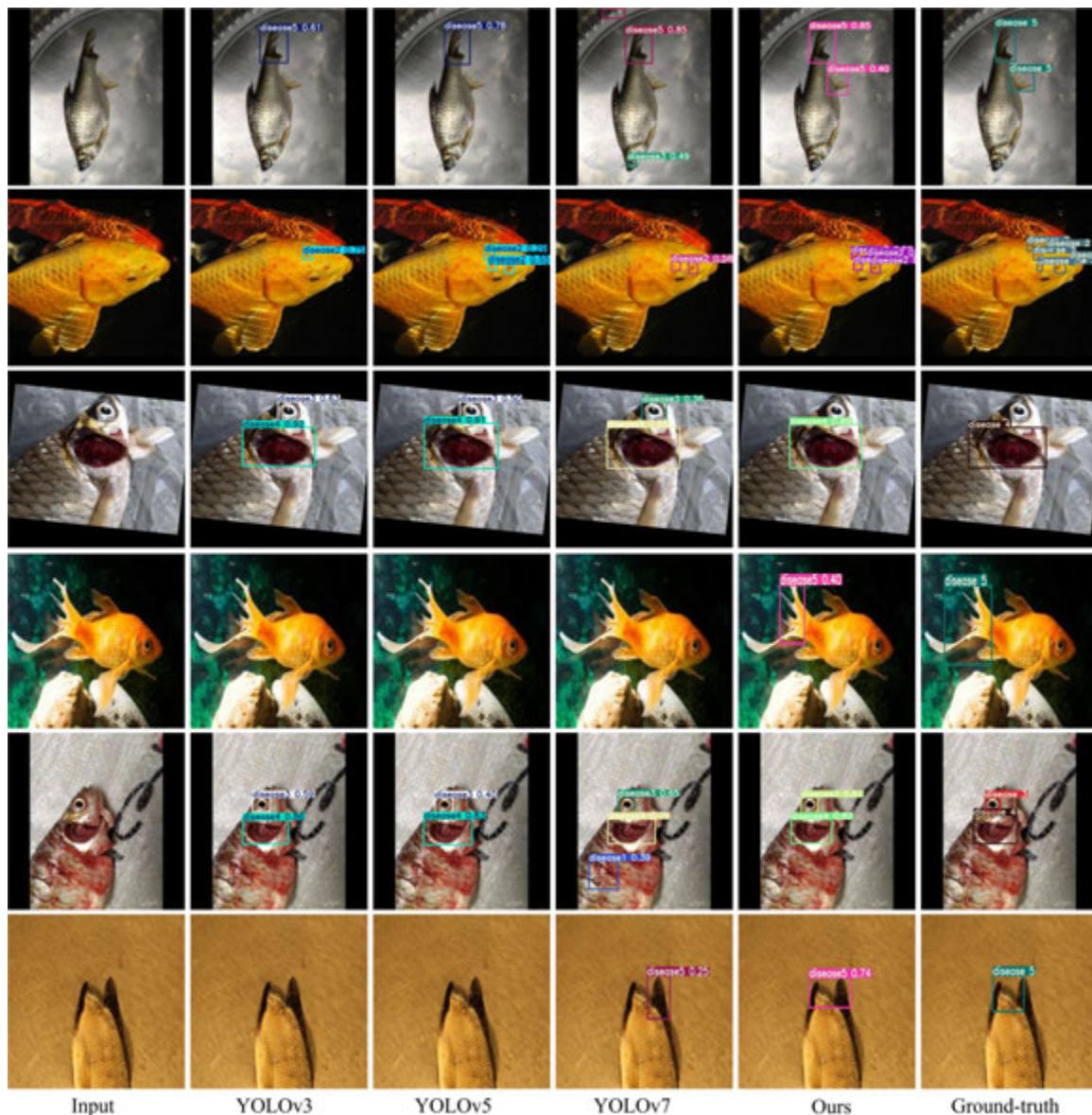


Fig. 4 Qualitative comparison of the proposed method (Ours) with existing state-of-the-art approaches for fish disease detection

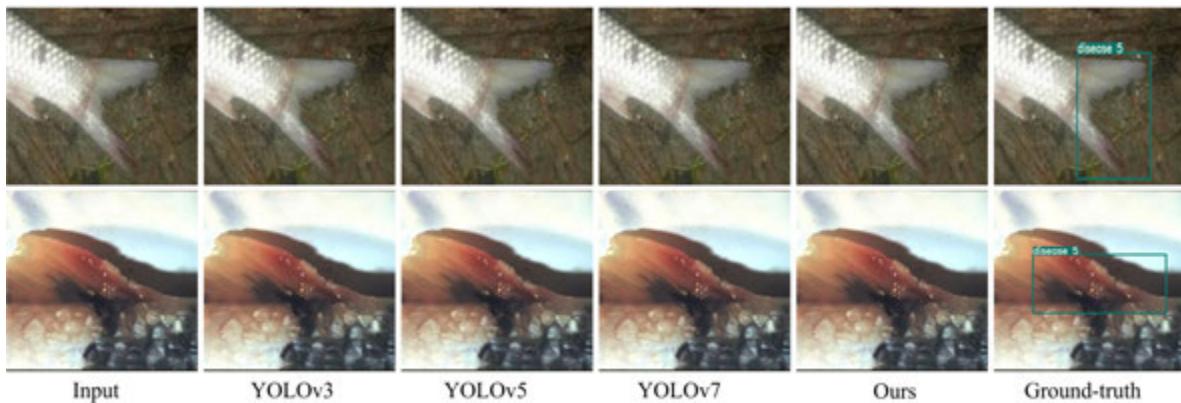


Fig. 5 Limitations of the proposed and existing approaches for fish disease detection

#### 4-4. Results and Discussion

**Quantitative Comparison** Table 1 shows the comparison of the proposed approach with existing state-of-the-art approaches for detection of fish disease. From this comparison, we can see that the proposed approach proves its efficiency for three evaluation measures out of four considered. The proposed approach has an increment of 1.4% in Recall which shows the ability of proposed approach for correct positive predictions by minimizing the false negatives. Similarly, there is an improvement of 0.2% and 1.1% in *mAP* at 50% and 50:90% threshold respectively.

**Qualitative Comparison** The visual comparison of proposed approach for fish disease detection with existing state-of-the-art approaches is provided in Figure 4. The results shows that the proposed approach is efficient in detecting the correct disease with a high recall (refer row 1,2, and 4 of Figure 4). Also, the proposed approach performs better in minimizing any false positive detection as compared specifically with YOLOv7 (refer row 1,3, and 5 of Figure 4). Further, as seen from the last row of Figure 4, our approach efficiently detects the fish disease with high confidence (0.74) whereas some of existing approaches (YOLOv3, YOLOv5) fail at detection and the other (YOLOv7) has incorrect localization of disease (refer last row of Figure 4).

**Limitations and future scope** In this section, we discuss the limitations of the proposed and existing approaches for detection of fish diseases. All the approaches

considered along with the proposed approach fail at detecting the fish disease in case of camouflage condition (refer row 1 of Figure 5). Further, as shown in row 2 of Figure 5, the detection approaches fail at detecting the fish disease when the image consists of localized part of fish body. In order to overcome these limitations, in future we can extend this work with more efficient feature learning module such as Mamba models<sup>[6]</sup>. Also, the work will be extended with the efficient performance on different scenarios such as poor lightning, water turbidity, etc.

## 5

## CONCLUSION

In this work, we propose a modified YOLOv7 architecture consisting of transformer feature efficient module (TFEM) in the backbone part for fish disease detection. The proposed TFEM based backbone allows the network to learn long-range dependencies with respect to input image in turn helping the effective disease detection capability. The numeric and visual comparison of the proposed approach is carried out with existing YOLOvX based approaches proving its applicability for fish skin disease detection. Further, the limitations of all the approaches with future scope for reliable fish skin disease detection are discussed.

## REFERENCES

- [1] <https://www.fao.org/4/x5568E/x5568e09.htm>
- [2] Ahmed, M. S., Aurpa, T. T., Azad, M. A. K.: Fish disease

- detection using image based machine learning technique in aquaculture. *Journal of King Saud University- Computer and Information Sciences* 34(8), 5170-5182 (2022)
- [3] Azhar, A. S. B. M., Harun, N. H. B., Hassan, M. G. B., Yusoff, N. B., Pauzi, S. N. B. M., Yusuf, N. N., Chu, K. B.: Early screening protozoan white spot fish disease using convolutional neural network. *Journal of Advanced Research in Applied Sciences and Engineering Technology* 37(1), 49-55 (2024)
- [4] Chakravorty, H., Paul, R., Das, P.: Image processing technique to detect fish disease. *International Journal of Computer Science and Security (IJCSS)* 9(2), 121-131 (2015)
- [5] Chen, J. C., Chen, T. L., Wang, H. L., Chang, P. C.: Underwater abnormal classification system based on deep learning: A case study on aquaculture fish farm in taiwan. *Aquacultural Engineering* 99, 102290 (2022)
- [6] Gu, A., Dao, T.: Mamba: Linear-time sequence modeling with selective state spaces. *arXiv preprint arXiv:2312.00752* (2023)
- [7] Gupta, A., Bringsdal, E., Knausgård, K. M., Goodwin, M.: Accurate wound and lice detection in atlantic salmon fish using a convolutional neural network. *Fishes* 7(6), 345 (2022)
- [8] Hasan, N., Ibrahim, S., Aqilah Azlan, A.: Fish diseases detection using convolutional neural network (cnn). *International Journal of Nonlinear Analysis and Applications* 13(1), 1977-1984 (2022)
- [9] Jocher, G., Stoken, A., Borovec, J., Changyu, L., Hogan, A., Diaconu, L., Ingham, F., Poznanski, J., Fang, J., Yu, L., et al.: ultralytics/yolov5: v3. 1-bug fixes and performance improvements. *Zenodo* (2020)
- [10] Kingma, D. P.: Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014)
- [11] Liu, C., Wang, Z., Li, Y., Zhang, Z., Li, J., Xu, C., Du, R., Li, D., Duan, Q.: Research progress of computer vision technology in abnormal fish detection. *Aquacultural Engineering* p. 102350 (2023)
- [12] Malik, S., Kumar, T., Sahoo, A.: Image processing techniques for identification of fish disease. In: 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP). pp. 55-59. IEEE (2017)
- [13] Mia, M. J., Mahmud, R. B., Sadad, M. S., Al Asad, H., Hossain, R.: An in-depth automated approach for fish disease recognition. *Journal of King Saud University- Computer and Information Sciences* 34(9), 7174-7183 (2022)
- [14] Navot, N., Sinyakov, M. S., Avtalion, R. R.: Application of ultrasound in vaccination against goldfish ulcer disease: A pilot study. *Vaccine* 29(7), 1382-1389 (2011)
- [15] Okawa, R., Iwasaki, N., Okamoto, K., Marsh, D.: Detection of abnormal fish by image recognition using fine-tuning. *Artificial Life and Robotics* 28(1), 175-180 (2023)
- [16] Park, J. S., Oh, M. J., Han, S.: Fish disease diagnosis system based on image processing of pathogens' microscopic images. In: 2007 Frontiers in the Convergence of Bioscience and Information Technologies. pp. 878-883. IEEE (2007)
- [17] Patil, P. W., Gupta, S., Rana, S., Venkatesh, S., Murala, S.: Multi-weather image restoration via domain translation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 21696-21705 (2023)
- [18] Redmon, J.: Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767* (2018)
- [19] Sharma, M., Shrivastav, A., Sahni, Y., Pandey, G.: Overviews of the treatment and control of common fish diseases (2012)
- [20] Sikder, J., Sarek, K. I., Das, U. K.: Fish disease detection system: a case study of freshwater fishes of bangladesh. *Int. J. Adv. Comput. Sci. Appl. (IJACSA)* 12(6), 867-871 (2021)
- [21] Tarvainen, A., Valpola, H.: Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. *Advances in neural information processing systems* 30 (2017)
- [22] Terven, J., Córdova-Esparza, D. M., Romero-González, J. A.: A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. *Machine Learning and Knowledge Extraction* 5(4), 1680-1716 (2023)
- [23] Vaswani, A.: Attention is all you need. *Advances in Neural Information Processing Systems* (2017)
- [24] Wang, C. Y., Bochkovskiy, A., Liao, H. Y. M.: Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In: Proceedings of the IEEE/

CVF conference on computer vision and pattern recognition.  
pp. 7464-7475 (2023)

[25] Wang, Z., Liu, H., Zhang, G., Yang, X., Wen, L., Zhao, W.: Diseased fish detection in the underwater environment using an improved yolov5 network for intensive aquaculture. Fishes 8(3), 169 (2023)

[26] Yang, X., Nian, R., Lin, H., Duan, C., Sui, J., Cao, L.: Detection of anisakid larvae in cod fillets by uv fluorescent imaging based on principal component analysis and gray value analysis. Journal of food protection 76(7), 1288-1292 (2013)

[27] Yasruddin, M. L., Ismail, M. A. H., Husin, Z., Tan, W. K.: Feasibility study of fish disease detection using computer vision and deep convolutional neural network (dcnn) algorithm. In: 2022 IEEE 18th International Colloquium on Signal Processing & Applications (CSPA). pp. 272-276. IEEE (2022)

[28] Yu, G., Zhang, J., Chen, A., Wan, R.: Detection and identification of fish skin health status referring to four common diseases based on improved yolov4 model. Fishes 8(4), 186 (2023)

[29] Zhang, X., Fu, Z., Wang, R.: Development of the es-fdd: an expert system for fish disease diagnosis. In: Oceans' 04 MTS/IEEE Techno-Ocean'04 (IEEE Cat. No. 04CH37600). vol. 1, pp. 482-487. IEEE (2004)

[30] Zhang, Z., Lu, X., Cao, S.: An efficient detection model based on improved yolov5s for abnormal surface features of fish. Math Biosci Eng 21(2), 1765-1790 (2024)

■著者



Shruti Phutke  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India



Amit Shakya  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India



Chetan Gupta  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India



Rupesh Kumar  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India



Tsuyoshi Kuroda  
Technical Research and  
Development Center Yamaha  
Motor Co., Ltd.



Lalit Sharma  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India

# OSCMamba: Omni-directional Selective Scan Convolution Mamba for Medical Image Classification

Shruti Phutke Amit Shakya Chetan Gupta Rupesh Kumar Lalit Sharma

This paper, titled "OSCMamba: Omni-directional Selective Scan Convolution Mamba for Medical Image Classification," was presented at CVIP-2024 (9th International Conference on Computer Vision & Image Processing), held at IIITDM Kancheepuram, Chennai, India, from December 19–21, 2024.

Reprinted with permission. Copyright © 2024 CVIP and IIITDM.

Further use or distribution is not permitted without permission from CVIP and IIITDM.

## Abstract

The advancement of various learning approaches has a great impact in computer vision applications specifically for medical image analysis. Being the most important task, classification accuracy of medical images has been successively improved using different methods such as Convolutional neural networks (CNNs), Transformers, etc. These models have some limitations such as the CNNs perform poorly when the feature extraction considering long-range dependency is concerned. Whereas Transformers perform well while dealing with the long-range dependencies for feature extraction, leading to the quadratic complexity. The evolution of state space models (SSMs) deals with the limitations of both the CNNs and Transformers. This has the advent of capturing the long-range dependencies and the linear complexity. Further, the scanning mechanism in the SSM provides the advantage of focusing on the required features while ignoring the rest. The existing Mamba based approach for medical image classification considers only horizontal and vertical feature scanning ignoring the diagonal information. While the omni-directional selective scan considers all of them. With this motivation, we propose omni directional selective scan-based convolution mamba (OSCMamba) approach for medical image classification. The OSCMamba approach is applied on different medical image modalities for image classification. The detailed experimental analysis with AUC and ACC on six different datasets proves the efficiency of the proposed OSCMamba based medical image classification approach. The source code is available at: <https://github.com/shrutiphetke/OSCMamba>

## 1

## INTRODUCTION

Classifying the medical images into various categories depending on the clinical information present in an image has a vital role in the field of medical image analysis. Different medical imaging modalities such as X-Rays, Ultrasound, Optical coherence tomography (OCT), Histopathology, etc. represent the information in unique ways. Traditional diagnostic methods are often subjective, and experiments suggest that the rate of disagreement among pathologists for diagnosis is about 24%<sup>[24]</sup>, highlighting the need for more consistent approaches. Computer-aided diagnosis (CAD) has emerged as a major research focus, aiming to assist radiologists by providing

a second opinion that enhances accuracy, consistency, and reduces image analysis time. Recent contributions in the field of computer vision and deep learning led to precise and quick CAD<sup>[12][39]</sup> of medical images.

Extensive research on Convolution neural networks (CNNs)<sup>[54][5][15]</sup> and Transformers<sup>[36][46][41]</sup> has demonstrated significant performance gain in CAD. Detecting abnormalities in the medical images relies on the global visualization of distinct features from each organ. CNNs excel at extracting local features, but they have difficulty capturing long-range dependencies needed for global feature learning. The self-attention mechanism of Transformers deals with capturing the long-range

dependencies in turn providing the global feature extractions. Despite of providing high classification accuracy, Transformers comes with high computational complexity due to its quadratic self-attention mechanism.

Due to the limitations of CNNs and Transformers, state space models (SSMs)<sup>[29]</sup> have attracted research interest for their capability to capture long-range dependencies using a linear state space framework. The conventional SSMs<sup>[19][20]</sup> have limited performance due to their limited ability of focusing or ignoring specific input features. To overcome this limitation, the advanced SSMs like Mamba<sup>[18]</sup> allows to attentively focus on relevant information via selective scanning mechanism and the hardware-aware algorithm allows linear data processing. The Mamba models are consistently used in various applications such as Language Modelling<sup>[38]</sup>, Speech separation<sup>[27]</sup>, Classification<sup>[58]</sup>, etc. Further, the Mamba have also shown great achievement in different computer vision applications like image restoration<sup>[57]</sup>, image super-resolution<sup>[7]</sup>, image deblurring<sup>[16]</sup>, etc.

In medical image classification, efficiently identifying and considering relevant features is crucial for achieving high classification accuracy. The existing Mamba based approach for medical image classification<sup>[53]</sup> effectively over- comes the limitations of CNNs and Transformers while ignoring the efficient feature scanning approach. With this motivation, in this work, we proposed a omni-directional selective scanning convolution block (OSCB) based approach for medical image classification. The contributions of the proposed work are:

- We propose a novel Mamba based Medical Image classification approach by exploiting the efficient feature scanning mechanism.
- The omni-directional selective scanning-based convolution block is proposed for efficient feature extraction.
- The extensive comparison of proposed method is carried out on the datasets with six different medical imaging modalities.

The consistent improvement of area under curve (AUC) and accuracy (ACC) as compared to existing state-of-the-art medical image classification methods shows the effectiveness of the proposed OSCB approach.

## 2

## RELATED WORKS

This section gives a brief overview of the related works in the field of medical image classification and the Mamba approaches for computer vision applications.

### 2-1. Medical Image Classification

Classifying medical images is an essential task for diagnosing a patient's health condition. The use of computer-aided diagnosis for medical images helps doctors analyze them more efficiently and effectively. Earlier researchers used machine learning methods for medical image classification. In this, the features of the input image are extracted with conventional computer vision approach (hand- crafted features) which are then used to train the classifier such as Support vector machine (SVM)<sup>[49]</sup>, K-nearest neighbour (KNN)<sup>[35]</sup>, etc. The widely used SVM classifier is applied across various medical imaging modalities, including retinopathy<sup>[40]</sup>, functional magnetic resonance imaging (fMRI)<sup>[49]</sup>, ultrasound<sup>[44]</sup>, and others. The KNN classifier is proposed for computed tomography (CT) image classification in<sup>[35]</sup>. Similarly, Iwahori *et al.*<sup>[25]</sup> proposed K-means clustering approach for endoscopic image classification. The manual feature extraction followed by classification approach is time-consuming and may miss important features in the images, potentially resulting in less accurate or less reliable classification outcomes. In comparison, convolutional neural networks (CNNs) automatically learn and extract hierarchical features from raw image data, allowing them to identify intricate patterns and details more efficiently. This capability often leads to improved classification accuracy and better generalization<sup>[33]</sup>.

**CNN in Medical Image Classification** In<sup>[42]</sup>, authors proposed a CNN based approach for diagnosing the diabetic retinopathy (DR) from fundus images. Similarly, Zhang *et al.*<sup>[54]</sup> proposed a CNN based approach and

compared existing CNN approaches for Pneumonia detection from chest X-Ray. Further, in<sup>[5][15]</sup> the authors proposed a CNN based approach with minimal number of hidden layers for skin cancer classification. In<sup>[11]</sup>, authors proposed the integration of CNN models with illumination normalization techniques to achieve higher classification accuracy. Balasubramaniam *et al.*<sup>[3]</sup> proposed a modified corrected ReLu activation based LeNet approach for breast cancer diagnosis. Several transfer learning-based methods have been proposed for medical image classification to overcome the challenge of limited datasets<sup>[45]</sup>. In<sup>[43][26]</sup> authors proposed a transfer learning on VGG, AlexNet, DenseNet201, ResNet18, SqueezeNet, etc. networks for pneumonia detection. Further, some works<sup>[22]</sup>, <sup>[8]</sup>, and <sup>[30]</sup> utilized the transfer learning approach for Ultrasound and retinopathy image classification respectively. Though CNN approaches achieved improved classification accuracy as compared to conventional machine learning approaches, they have limited performance due to their localized feature processing. This leads to limited ability of capturing the long-range dependency with respect to the input. The Transformers perform well in capturing the long-range dependencies which further help in efficient classification of medical images.

**Transformers in Medical Image Classification** Medical imaging modalities have organ specific representation that needs to be processed with highly efficient feature representation<sup>[46]</sup>. The well known fact of the Transformers of being able to efficiently capture the input dependent feature representation makes them a suitable choice over CNNs for medical image classification. Matsoukas *et al.*<sup>[37]</sup> provided the analysis of utilizing the Transformers over CNN for medical image classification task. Multiple approaches are proposed for retina disease classification<sup>[48]</sup>, tumor classification<sup>[10]</sup>, etc. Sun *et al.*<sup>[48]</sup> proposed a encoder with pixel relation and decoder with lesion-aware transformer for diabetic retinopathy grading. In<sup>[10]</sup>, the authors proposed a hybrid CNN and Transformer based approach for multi-modal medical image classification. Gheflati *et al.*<sup>[17]</sup> proposed a similar hybrid approach with pre-trained model for breast cancer

detection using ultrasound images. These approaches utilized the existing vision transformer ViT<sup>[13]</sup> model as a plug-and-play module for classification task. Further, Omid *et al.*<sup>[36]</sup> proposed a hybrid CNN Transformer approach with modified Transformer layer using an efficient convolution operation for medical image classification. These Transformer based approaches provide efficient classification accuracy but with increased computational complexity.

## 2-2. Mamba in Computer Vision

Considering the fact that CNNs capture only the local relationship of the features whilst the Transformers have quadratic complexity, researchers come up with the selective state space models called as Mamba<sup>[18]</sup>. The Mamba models are efficient at capturing the long-range dependencies unlike CNNs and have a linear computational complexity unlike Transformers. With the success of Mamba in Natural language processing tasks<sup>[38][27][50][58]</sup>, it further achieved re-markable performance in vision applications<sup>[57][21][7][16]</sup>. Zheng *et al.*<sup>[57]</sup> proposed a U-shape Mamba approach for single image dehazing. In<sup>[21]</sup> a residual state space block is proposed with channel attention and Mamba with 2-D selective scanning approach for image restoration. Chen *et al.*<sup>[6]</sup> proposed a Mamba-in-mamba with 2-D selective scanning for small target detection. Further, Yue *et al.*<sup>[53]</sup> proposed a SS-Conv-SSM approach consisting of a 2-D selective scanning-based Mamba in parallel with the convolution path for medical image classification. Shi *et al.*<sup>[47]</sup> proposed a omni-selective scan by processing six directional information to overcome the unidirectional scanning limitation of 2-D selective scan for image restoration. Further, Zhao *et al.*<sup>[56]</sup> proposed an omni-directional selective scan-based approach for remote sensing image dense prediction.

Following the success of Mamba models in computer vision applications and with the utilization of efficient selective scanning mechanism, we propose a omni-directional selective scan-based convolution layer for medical image classification. The details of the proposed approach and experiments on existing dataset are provided in successive Sections.

## 3 METHODOLOGY

In this section, first we introduce the preliminaries of Mamba approach and then give details about the proposed omni-directional selective scan-based convolution block based approach for medical image classification.

### 3-1. State Space Models

The state space models (SSMs) are used to make the predictions of next state ( $y(t)$ ) depending upon the input ( $x(t)$ ) provided in current state ( $h(t)$ ). The SSMs assume the inputs to be continuous in time and can be represented by two ordinary differential equations (ODEs) as state equation and the output equation.

$$\begin{aligned} h'(t) &= Ah(t) + Bx(t); & \text{state equation} \\ y(t) &= Ch(t) + Dx(t); & \text{output equation} \end{aligned} \quad (1)$$

where,  $A$  is the state transition matrix describing how the state  $h(t)$  changes with influence of the input  $x(t)$  through input projection matrix  $B$ .  $C$  indicates translation of the state towards output with respect to input  $x(t)$  through the feed-forward matrix  $D$ . As the matrix  $D$  is considered as a skip connection between input and output, the SSMs are represented by neglecting the matrix  $D$  and represented as:

$$\begin{aligned} h'(t) &= Ah(t) + Bx(t); & \text{state equation} \\ y(t) &= Ch(t); & \text{output equation} \end{aligned} \quad (2)$$

This represents the global feature dependency of SSMs since the current output is dependent on all the preceding states and the input. Unlike above equations where the input is considered as continuous in time, the deep learning approaches assume the input to be discrete in time. The S4<sup>[20]</sup> and Mamba<sup>[18]</sup> convert these equations from continuous ODEs to discrete time representation by utilizing zero-order hold approach. In order to achieve this, a time scale parameter  $\Delta$  is introduced and the matrix  $A$  and  $B$  are transformed into  $\bar{A}$  and  $\bar{B}$  respectively as follows:

$$\begin{aligned} \bar{A} &= \exp(\Delta A) \\ \bar{B} &= (\Delta A)^{-1}(\exp(\Delta A) - I) \cdot \Delta B \end{aligned} \quad (3)$$

This discrete representation now allows to transform the discrete input  $x_k$  to the discrete output  $y_k$  with  $k$  as discrete time step by using following representation:

$$\begin{aligned} h_k &= \bar{A}h_{k-1} + \bar{A}x_k; & \text{state equation} \\ y_k &= Ch_k; & \text{output equation} \end{aligned} \quad (4)$$

The SSMs can be represented with Convolution kernel ( $\bar{K}$ ) as:

$$\begin{aligned} \bar{K} &= (C\bar{B}, C\bar{A}\bar{B}, \dots, C\bar{A}^{L-1}\bar{B}) \\ y &= x * \bar{K} \end{aligned} \quad (5)$$

where,  $x$  is input and  $y$  is output,  $L$  is length of input  $x$ . The matrix  $\bar{A}$  is build with HiPPO to memorize all the hidden states in<sup>[20]</sup> called as structured state space for sequences (S4). The SSMs are efficient for modelling the input sequences but fail at filtering the irrelevant information and the ease of parallel scanning. To overcome this limitation a Selective State Space Models<sup>[58]</sup> propose a selective scan approach. Also to solve the issue of GPU utilization, hardware-aware state expansion approach is enabled in selective scan mechanism.

### 3-2. Proposed Architecture

In this section, we first give the pipeline of the proposed classification architecture. Further, the details of the omni-directional selective scanning convolution block followed by scanning mechanism are provided.

**Architecture Overview** In the proposed architecture, the input image  $I \in \mathbb{R}^{H \times W \times C_{in}}$  ( $C_{in} = 3$  for RGB image and  $C_{in} = 1$  for gray-scale image) is firstly, converted into  $4 \times 4$  sized non-overlapping patches in Patch embedding layer to process them in Mamba block. These patches are then fed to the first omni-directional selective scan convolution block (OSCB) followed by the patch merging layer in turn generating the feature map of size  $h \times w \times C$ , where  $h = \frac{H}{4}$ ,  $w = \frac{W}{4}$  and  $C = 96$  (see Figure 1). We call the each OSCB in Figure 1 as an encoder layer and there are four such layers with  $n_l \in [2, 2, 4, 8]$  repeated OSCB where  $l \in (1, 4)$  is number of layers. Further, a global average pooling followed by a fully connected layer is used to predict the output class from the input image.

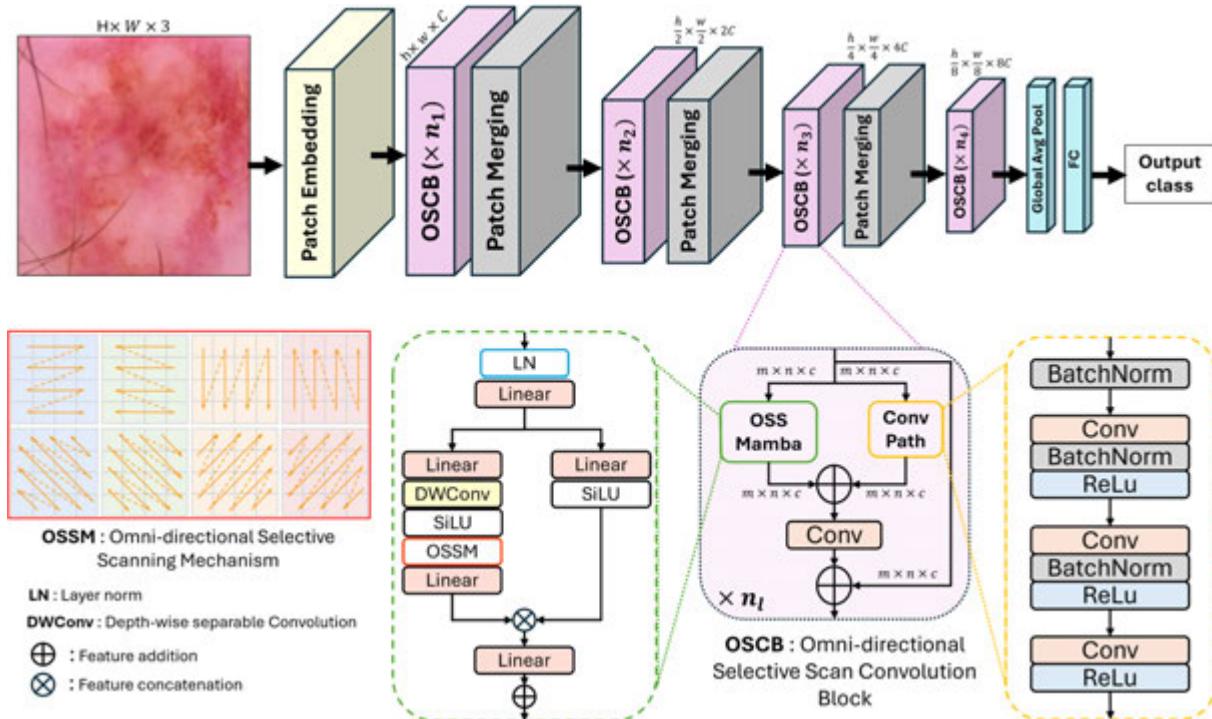


Fig. 1 Proposed Omni-directional Selective Scan Convolution mamba (OSCMamba) architecture for medical image classification

### Omni-directional Selective Scanning Convolution Block

Figure 1 shows the proposed omni-directional selective scan convolution block (OSCB). Unlike<sup>[53]</sup>, we process the incoming features separately in omni-directional selective scan path and the convolution path. This helps the network to learn the global and local feature dependencies separately. These global and local features are then added together followed by a convolution layer. Further, a residual connection is provided to preserve the feature information (see OSCB in Figure 1). The *convolution path* (see Conv Path in OSCB of Figure 1) in OSCB processes the image with Convolution → Batch Normalization → ReLu to capture the local feature dependency. Further, in the *omni-directional selective scanning (OSS) Mamba* path, the incoming features are normalized and projected linearly thorough linear projection layer.

The OSS Mamba block is integration of the Gated MLP where the input features are first split into equal parts along channel dimension<sup>[55]</sup>. One path processes the features with OSS mechanism (OSSM) and the other projects the linearly convoluted features via an activation function (see the left and right split in *OSS Mamba* block

of OSCB in Figure 1). In the first branch, the features are linearly projected, and a depth-wise separable convolution is applied on the features this helps in reducing the number of parameters followed by an activation function. These processed features are then forwarded to the OSSM block where these features are processed in omni-directional selective scanning mechanism in order to capture the global dependency from all the directions. These global features are then linearly projected and multiplied with the activated features from the other path (see *OSS Mamba* block in Figure 1).

**Omni-directional Selective Scan Mechanism** The existing selective scanning approaches such as Bi-directional, Cross-Scan, Continuous 2D, etc. consider the feature representation only in vertical and horizontal direction<sup>[55]</sup>. The Bi-directional scanning approach considers only horizontal direction for the forward and reverse scanning (similar to first two columns of the first row of Omni-directional Selective Scanning mechanism in Figure 1). Whereas the cross-scan and continuous 2D scan considers the horizontal and vertical scanning mechanism

for forward and reverse approach (similar to the first row of Omni-directional Selective Scanning mechanism in Figure 1). Unlike these scanning approaches, the omnidirectional selective scanning mechanism (OSSM) considers the horizontal, vertical, diagonal, and off-diagonal scanning for both forward and backward scan (see OSSM in Figure 1). These scanned features in eight directions are then processed in state space model (S6) separately<sup>[56]</sup>. Further the processed directional features are combined achieving the global information. This approach allows the efficient spatial directional feature learning.

## 4

## EXPERIMENTS AND RESULT DISCUSSION

In this section, we provide the details of the dataset utilized for experimental analysis, the implementation details, evaluation metrics, and discussion on result analysis.

### 4-1. Dataset

For experimental analysis, we considered the MedMNIST<sup>[52]</sup> dataset which consists of set of multiple medical image modality datasets. It covers different medical imaging modalities such as Ultrasound, X-Ray, Optical coherence tomography (OCT), Dermatology, Microscope, fundus camera, etc. Due to wide variety of dataset with different classification categories such as multi-class, multi-label, binary class, etc. MedMNIST is considered as the benchmark dataset for medical image

classification. Table 1 shows the training, validation and testing splits of six different datasets used for experimental analysis. The sample images from each dataset are provided in Figure 2 and details of each dataset are provided below:

**BreastMNIST** This dataset consists of 780 Breast Ultrasound images from three different categories: malignant, benign and normal with the original resolution of  $500 \times 500$ <sup>[2]</sup>. In<sup>[52]</sup>, this dataset is again converted into binary classification task with positive class consisting of benign and normal images and negative class consisting of malignant images.

**RetinaMNIST** This dataset has 1600 retina fundus images with  $3 \times 1736 \times 1824$  pixel resolution from DeepDRiD dataset<sup>[34]</sup>. It has 5 different severity ratings of diabetic retinopathy. The total training dataset<sup>[34]</sup> is divided into training and validation set with a 9:1 ratio. The test set is the actual validation set from<sup>[34]</sup>.

**PneumoniaMNIST** This is binary classification task dataset consisting of 5856 gray-scale pediatric chest X-Ray images with pixel resolution ranging in  $(384 - 2916) \times (127 - 2713)$ <sup>[32][31]</sup>. The total dataset is divided into training and validation set with a 9 : 1 ratio and the test set is the actual validation set from<sup>[32][31]</sup>.

**DermaMNIST** This dataset is a large collection of common pigmented skin lesions image from multi-source dermascope<sup>[51][9]</sup>. It has seven different categories of skin

**Table 1 Overview of different datasets in MedMNIST<sup>[52]</sup>. BC: Binary-Class, OR: Ordinal Regression, MC: Multi-Class**

Dataset	Data Modality	Task (#Classes/Labels)	# Train/Validation/Test
BreastMNIST	Breast Ultrasound	BC (2)	546/78/156
RetinaMNIST	Fundus Camera	OR (5)	1,080/120/400
PneumoniaMNIST	Chest X-Ray	BC (2)	4,708/524/624
DermaMNIST	Dermatoscope	MC (7)	7,007/1,003/2,005
BloodMNIST	Blood Cell Microscope	MC (8)	11,959/1,712/3,421
OCTMNIST	Retinal OCT	MC (4)	97,477/10,832/1,000

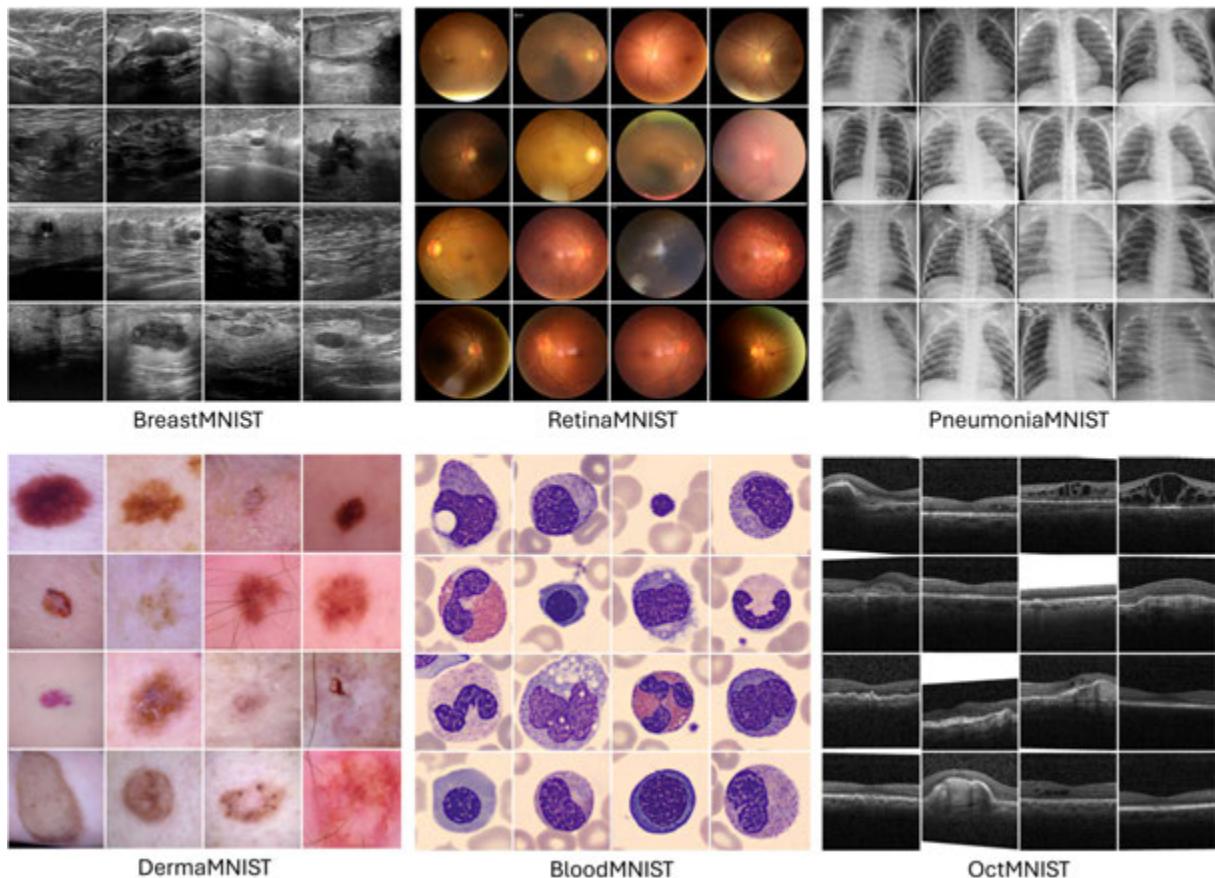


Fig. 2 Sample images from different medical image classification MedMNIST dataset

lesions such as actinic keratoses and intra-epithelial carcinoma, dermatofibroma, enign keratosis-like lesions, melanoma, basal cell carcinoma, melanocytic nevi, and vascular lesions. DermaMNIST consists of 10015 dermatoscopic images with  $3 \times 600 \times 450$ -pixel resolution having a 7:1:2 split ratio for training, validation and test splits.

**BloodMNIST** The BloodMNIST dataset consists of cellular images with 8 different categories such as basophil, erythroblast, eosinophil, immature granulocytes (metamyelocytes, myelocytes, and promyelocytes), lymphocytes, monocyte, neutrophil, and platelet<sup>[1]</sup>. This dataset is collected from the individuals that are free from any pharmacologic treatment, without any infection and hematologic or oncologic disease during the blood sample collection. It consists of 17,092 images with a training, validation and test split of 7:1:2. The images have  $3 \times 360 \times 363$  pixels resolution.

**OctMNIST** This dataset is derived from<sup>[31]</sup> consisting of 109309 optical coherence tomography (OCT) images for retinal disease analysis. It has 4 different categories like choroidal neovascularization, drusen, normal, diabetic macular edema. The image resolution varies in the range of  $(384-1536) \times (277-512)$ . The training and validation split for OctMNIST is taken from original training dataset<sup>[31]</sup> with a 9:1 split ratio and the testing split is a validation set from original dataset.

#### 4-2. Implementation Details and Evaluation Metrics

The pre-processing on the images from all the dataset is carried out similar to MedMNIST<sup>[52]</sup>. To train the proposed classification network, we resized the images into  $224 \times 224$  resolution and set the batch size = 64. The network is trained for 200 epoch using the early stopping criteria evaluated on validation accuracy with a patience of 50. The network parameters are optimized using Cross-Entropy loss and the Stochastic Gradient Descent (SGD) optimizer, with a learning rate set to 0.001. The training

is carried out on NVIDIA A100 GPU. For evaluation of the performance of proposed network with state-of-the-art medical image classification methods, we considered the Area under the receiver operating characteristic (ROC) curve (AUC) and Accuracy (ACC) as evaluation metric similar to<sup>[52]</sup>.

#### 4-3. Result Analysis

The comparison of the proposed approach is done with

existing state of the art medical image classification methods in terms of AUC and ACC as provided in Table 2, 3. We evaluated the proposed approach on six different medical imaging modalities dataset. As seen from the Table 2, 3 we achieve significant improvement in ACC on five among six medical image modality datasets. Where as there is significant improvement in AUC on four medical image modality datasets. Though accuracy (ACC) is more sensitive to class inconsistency

**Table 2 Evaluation of the proposed method and existing state-of-the-art approaches for medical image classification. Note: **Bold** and underline shows the best and second best values respectively**

Dataset	OCTMNIST		DermaMNIST		RetinaMNIST	
Metric	ACC	AUC	ACC	AUC	ACC	AUC
ResNet18 <sup>[23]</sup>	0.943	0.743	0.917	0.735	0.717	0.524
ResNet18 <sup>[23]</sup>	0.958	0.763	0.920	0.754	0.710	0.493
ResNet50 <sup>[23]</sup>	0.952	0.762	0.913	0.735	0.726	0.528
ResNet50 <sup>[23]</sup>	0.958	0.776	0.912	0.731	0.716	0.511
auto-sklearn <sup>[14]</sup>	0.887	0.601	0.902	0.719	0.690	0.515
AutoKeras <sup>[28]</sup>	0.955	0.763	0.915	0.749	0.719	0.503
Google AutoML <sup>[4]</sup>	0.963	0.771	0.914	0.768	0.750	0.531
MedVit-T <sup>[36]</sup>	0.961	0.767	0.914	0.768	0.752	0.534
MedVit-S <sup>[36]</sup>	0.960	0.782	<u>0.937</u>	<u>0.780</u>	<u>0.773</u>	<u>0.561</u>
MedVit-L <sup>[36]</sup>	0.945	0.761	0.920	0.773	<u>0.754</u>	0.552
MedMamba <sup>[53]</sup>	<u>0.993</u>	<u>0.914</u>	0.907	0.758	–	–
Ours	<u>0.995</u>	<u>0.927</u>	0.948	0.794	0.741	<u>0.573</u>

**Table 3 Evaluation of the proposed method and existing state-of-the-art approaches for medical image classification**

Dataset	PneumoniaMNIST		BloodMNIST		BreastMNIST	
Metric	AUC	ACC	AUC	ACC	AUC	ACC
ResNet18 <sup>[23]</sup>	0.944	0.854	0.998	0.958	0.901	0.863
ResNet18 <sup>[23]</sup>	0.956	0.864	0.998	0.963	0.891	0.833
ResNet50 <sup>[23]</sup>	0.948	0.854	0.997	0.956	0.857	0.812
ResNet50 <sup>[23]</sup>	0.962	0.884	0.997	0.950	0.866	0.842
auto-sklearn <sup>[14]</sup>	0.942	0.855	0.984	0.878	0.836	0.803
AutoKeras <sup>[28]</sup>	0.947	0.878	0.998	0.961	0.871	0.831
Google AutoML <sup>[4]</sup>	0.991	0.946	0.998	0.966	0.919	0.861
MedVit-T <sup>[36]</sup>	0.993	0.949	0.996	0.950	0.934	0.896
MedVit-S <sup>[36]</sup>	0.995	<u>0.961</u>	<u>0.997</u>	0.951	<b>0.938</b>	<b>0.897</b>
MedVit-L <sup>[36]</sup>	0.991	0.921	0.996	0.954	<u>0.929</u>	0.883
MedMamba <sup>[53]</sup>	<u>0.965</u>	0.912	<b>0.999</b>	<u>0.984</u>	0.879	0.872
Ours	<b>0.988</b>	<b>0.962</b>	<b>0.999</b>	<b>0.985</b>	0.899	<u>0.885</u>

than AUC, we achieve remarkable improvement in ACC as compared to state-of-the-art medical image classification methods. It is observed that, on BreastMNIST and RetinaMNIST datasets MedVit<sup>[36]</sup> achieves notable improvement. The reason behind this improvement is due to the augmentation utilized in MedVit<sup>[36]</sup>, since both the BreastMNIST and RetinaMNIST have very few images in training set. Whereas in our proposed approach we have only considered horizontal flip augmentation unlike MedVit<sup>[36]</sup>. Apart from all the existing approaches, only considering a recent Mamba based classification approach<sup>[53]</sup> for comparison, we achieve the remarkable improvement on overall considered datasets in terms of AUC and ACC.

## 5 CONCLUSION

This work presents a novel Omni-directional selective scan convolution layer based medical image classification approach. The proposed approach is evaluated on six different medical imaging modalities such as Ultrasound, X-Ray, optical coherence tomography, histopathology, etc. The effectiveness of the proposed approach is verified by the comparison with the state-of-the-art medical image classification methods in terms of AUC and ACC. The extensive result analysis of the proposed approach verifies its effectiveness for the task of medical image classification.

## REFERENCES

- [1] Acevedo, A., Merino González, A., Alférez Baquero, E. S., Molina Borrás, Á., Boldú Nebot, L., Rodellar Benedé, J.: A dataset of microscopic peripheral blood cell images for development of automatic recognition systems. *Data in brief* 30(article 105474) (2020)
- [2] Al-Dhabyani, W., Gomaa, M., Khaled, H., Fahmy, A.: Dataset of breast ultrasound images. *Data in brief* 28, 104863 (2020)
- [3] Balasubramaniam, S., Velmurugan, Y., Jaganathan, D., Dhanasekaran, S.: A modified lenet cnn for breast cancer diagnosis in ultrasound images. *Diagnostics* 13(17), 2746 (2023)
- [4] Bisong, E., et al.: Building machine learning and deep learning models on Google cloud platform. Springer (2019)
- [5] Chaturvedi, S. S., Tembhere, J. V., Diwan, T.: A multi-class skin cancer classification using deep convolutional neural networks. *Multimedia Tools and Applications* 79(39), 28477–28498 (2020)
- [6] Chen, T., Tan, Z., Gong, T., Chu, Q., Wu, Y., Liu, B., Ye, J., Yu, N.: Mim-istd: Mamba-in-mamba for efficient infrared small target detection. *arXiv preprint arXiv:2403.02148* (2024)
- [7] Cheng, C., Wang, H., Sun, H.: Activating wider areas in image super-resolution. *arXiv preprint arXiv:2403.08330* (2024)
- [8] Cheng, P. M., Malhi, H. S.: Transfer learning with convolutional neural networks for classification of abdominal ultrasound images. *Journal of digital imaging* 30, 234–243 (2017)
- [9] Codella, N., Rotemberg, V., Tschandl, P., Celebi, M. E., Dusza, S., Gutman, D., Helba, B., Kalloo, A., Liopyris, K., Marchetti, M., et al.: Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). *arXiv preprint arXiv:1902.03368* (2019)
- [10] Dai, Y., Gao, Y., Liu, F.: Transmed: Transformers advance multi-modal medical image classification. *Diagnostics* 11(8), 1384 (2021)
- [11] Dash, S., Parida, P., Mohanty, J. R.: Illumination robust deep convolutional neural network for medical image classification. *Soft Computing* pp. 1–13 (2023)
- [12] Doi, K., MacMahon, H., Katsuragawa, S., Nishikawa, R. M., Jiang, Y.: Computer-aided diagnosis in radiology: potential and pitfalls. *European journal of Radiology* 31(2), 97–109 (1999)
- [13] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al.: An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929* (2020)
- [14] Feurer, M., Klein, A., Eggensperger, K., Springenberg, J., Blum, M., Hutter, F.: Efficient and robust automated machine learning. *Advances in neural information processing systems* 28 (2015)

- [15] Fu'adah, Y. N., Pratiwi, N. C., Pramudito, M. A., Ibrahim, N.: Convolutional neural network (cnn) for automatic skin cancer classification system. In: IOP conference series: materials science and engineering. vol. 982, p. 012005. IOP Publishing (2020)
- [16] Gao, H., Ma, B., Zhang, Y., Yang, J., Yang, J., Dang, D.: Learning enriched features via selective state spaces model for efficient image deblurring. In: ACM Multimedia (2024)
- [17] Gheftati, B., Rivaz, H.: Vision transformers for classification of breast ultrasound images. In: 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). pp. 480–483. IEEE (2022)
- [18] Gu, A., Dao, T.: Mamba: Linear-time sequence modeling with selective state spaces. arXiv preprint arXiv:2312.00752 (2023)
- [19] Gu, A., Dao, T., Ermon, S., Rudra, A., Ré, C.: Hippo: Recurrent memory with optimal polynomial projections. Advances in neural information processing systems 33, 1474–1487 (2020)
- [20] Gu, A., Goel, K., Ré, C.: Efficiently modeling long sequences with structured state spaces. arXiv preprint arXiv:2111.00396 (2021)
- [21] Guo, H., Li, J., Dai, T., Ouyang, Z., Ren, X., Xia, S. T.: Mambair: A simple base- line for image restoration with state-space model. arXiv preprint arXiv:2402.15648 (2024)
- [22] Gupta, S., Agrawal, S., Singh, S. K., Kumar, S.: A novel transfer learning-based model for ultrasound breast cancer image classification. In: Computational Vision and Bio-Inspired Computing: Proceedings of ICCVBIC 2022, pp. 511–523. Springer (2023)
- [23] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
- [24] Hsu, W., Han, S. X., Arnold, C. W., Bui, A. A., Enzmann, D. R.: A data-driven approach for quality assessment of radiologic interpretations. Journal of the American Medical Informatics Association 23(e1), e152–e156 (2016)
- [25] Iwahori, Y., Hattori, A., Adachi, Y., Bhuyan, M. K., Woodham, R. J., Kasugai, K.: Automatic detection of polyp using hessian filter and hog features. Procedia computer science 60, 730–739 (2015)
- [26] Jain, R., Nagrath, P., Kataria, G., Kaushik, V. S., Hemanth, D. J.: Pneumonia de-tection in chest x-ray images using convolutional neural networks and transfer learning. Measurement 165, 108046 (2020)
- [27] Jiang, X., Han, C., Mesgarani, N.: Dual-path mamba: Short and long-term bidirectional selective structured state space models for speech separation. arXiv preprint arXiv:2403.18257 (2024)
- [28] Jin, H., Song, Q., Hu, X.: Auto-keras: An efficient neural architecture search system. In: Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining. pp. 1946–1956 (2019)
- [29] Kalman, R.E.: A new approach to linear filtering and prediction problems (1960)
- [30] Kandel, I., Castelli, M.: Transfer learning with convolutional neural networks for diabetic retinopathy image classification. a review. Applied Sciences 10(6), 2021 (2020)
- [31] Kermany, D., Zhang, K., Goldbaum, M.: Large dataset of labeled optical coherence tomography (oct) and chest x-ray images. Mendeley Data 3(10.17632) (2018)
- [32] Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C., Liang, H., Baxter, S. L., McKeown, A., Yang, G., Wu, X., Yan, F., et al.: Identifying medical diagnoses and treatable diseases by image-based deep learning. cell 172(5), 1122–1131 (2018)
- [33] Li, Q., Cai, W., Wang, X., Zhou, Y., Feng, D. D., Chen, M.: Medical image clas-sification with convolutional neural network. In: 2014 13th International Conference on Control Automation Robotics Vision (ICARCV). pp. 844–848 (2014). <https://doi.org/10.1109/ICARCV.2014.7064414>
- [34] Liu, R., Wang, X., Wu, Q., Dai, L., Fang, X., Yan, T., Son, J., Tang, S., Li, J., Gao, Z., et al.: Deepdrid: Diabetic retinopathy—grading and image quality estimation challenge. Patterns 3(6) (2022)
- [35] Manju, B., Meenakshy, K., Gopikakumari, R.: Prostate disease diagnosis from ct images using ga optimized smrt based texture features. Procedia Computer Science 46,

- 1692–1699 (2015)
- [36] Manzari, O. N., Ahmadabadi, H., Kashiani, H., Shokouhi, S. B., Ayatollahi, A.: Medvit: a robust vision transformer for generalized medical image classification. *Computers in Biology and Medicine* 157, 106791 (2023)
- [37] Matsoukas, C., Haslum, J. F., Söderberg, M., Smith, K.: Is it time to replace cnns with transformers for medical images? *arXiv preprint arXiv:2108.09038* (2021)
- [38] Mehta, H., Gupta, A., Cutkosky, A., Neyshabur, B.: Long range language modeling via gated state spaces. *arXiv preprint arXiv:2206.13947* (2022)
- [39] Melendez, J., Van Ginneken, B., Maduskar, P., Philipsen, R. H., Reither, K., Breuninger, M., Adetifa, I. M., Maane, R., Ayles, H., Sánchez, C. I.: A novel multiple-instance learning-based approach to computer-aided detection of tuberculosis on chest x-rays. *IEEE transactions on medical imaging* 34(1), 179–192 (2014)
- [40] Niemeijer, M., Abramoff, M. D., van Ginneken, B.: Image structure clustering for image quality verification of color retina images in diabetic retinopathy screening. *Medical image analysis* 10(6), 888–898 (2006)
- [41] Patil, P. W., Gupta, S., Rana, S., Venkatesh, S., Murala, S.: Multi-weather image restoration via domain translation. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. pp. 21696–21705 (2023)
- [42] Pratt, H., Coenen, F., Broadbent, D. M., Harding, S. P., Zheng, Y.: Convolutional neural networks for diabetic retinopathy. *Procedia computer science* 90, 200–205 (2016)
- [43] Rahman, T., Chowdhury, M. E., Khandakar, A., Islam, K. R., Islam, K. F., Mahbub, Z. B., Kadir, M. A., Kashem, S.: Transfer learning with deep convolutional neural network (cnn) for pneumonia detection using chest x-ray. *Applied Sciences* 10(9), 3233 (2020)
- [44] Rani, A., Mittal, D., et al.: Detection and classification of focal liver lesions using support vector machine classifiers. *Journal of Biomedical Engineering and Medical Imaging* 3(1), 21 (2016)
- [45] Salehi, A. W., Khan, S., Gupta, G., Alabdullah, B. I., Almjally, A., Alsolai, H., Siddiqui, T., Mellit, A.: A study of cnn and transfer learning in medical imaging: Advantages, challenges, future scope. *Sustainability* 15(7), 5930 (2023)
- [46] Shamshad, F., Khan, S., Zamir, S. W., Khan, M. H., Hayat, M., Khan, F. S., Fu, H.: Transformers in medical imaging: A survey. *Medical Image Analysis* 88, 102802 (2023)
- [47] Shi, Y., Xia, B., Jin, X., Wang, X., Zhao, T., Xia, X., Xiao, X., Yang, W.: Vmambair: Visual state space model for image restoration. *arXiv preprint arXiv:2403.11423* (2024)
- [48] Sun, R., Li, Y., Zhang, T., Mao, Z., Wu, F., Zhang, Y.: Lesion-aware transformers for diabetic retinopathy grading. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 10938–10947 (2021)
- [49] Tan, L., Chen, Y., Maloney, T. C., Caré, M. M., Holland, S. K., Lu, L. J.: Combined analysis of smri and fmri imaging data provides accurate disease markers for hearing impairment. *NeuroImage: Clinical* 3, 416–428 (2013)
- [50] Tang, S., Dunnmon, J. A., Liangqiong, Q., Saab, K. K., Baykaner, T., Lee-Messer, C., Rubin, D. L.: Modeling multivariate biosignals with graph neural networks and structured state space models. In: *Conference on Health, Inference, and Learning*. pp. 50–71. PMLR (2023)
- [51] Tschandl, P., Rosendahl, C., Kittler, H.: The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific data* 5(1), 1–9 (2018)
- [52] Yang, J., Shi, R., Wei, D., Liu, Z., Zhao, L., Ke, B., Pfister, H., Ni, B.: Medmnist v2- a large-scale lightweight benchmark for 2d and 3d biomedical image classification. *Scientific Data* 10(1), 41 (2023)
- [53] Yue, Y., Li, Z.: Medmamba: Vision mamba for medical image classification. *arXiv preprint arXiv:2403.03849* (2024)
- [54] Zhang, D., Ren, F., Li, Y., Na, L., Ma, Y.: Pneumonia detection from chest x-ray images based on convolutional neural network. *Electronics* 10(13), 1512 (2021)
- [55] Zhang, H., Zhu, Y., Wang, D., Zhang, L., Chen, T., Wang, Z., Ye, Z.: A survey on visual mamba. *Applied Sciences* 14(13), 5683 (2024)
- [56] Zhao, S., Chen, H., Zhang, X., Xiao, P., Bai, L., Ouyang, W.: Rs-mamba for large remote sensing image dense prediction. *arXiv preprint arXiv:2404.02668* (2024)
- [57] Zheng, Z., Wu, C.: U-shaped vision mamba for single

image dehazing. arXiv preprint arXiv:2402.04139 (2024)

[58] Zhu, L., Liao, B., Zhang, Q., Wang, X., Liu, W., Wang, X.: Vision mamba: Efficient visual representation learning with bidirectional state space model. arXiv preprint arXiv:2401.09417 (2024)

**■著者**



Shruti Phutke  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India

Amit Shakya  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India

Chetan Gupta  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India



Rupesh Kumar  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India

Lalit Sharma  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India

# Automatic Early Classification of Cassava Leaf Disease with Ensemble of Lightweight Models

Kaushik Vishwakarma Udaybhan Rathore Rupesh Kumar Amit Shakya  
Shruti Phutke Chetan Gupta Lalit Sharma

This paper, titled “Automatic Early Classification of Cassava Leaf Disease with Ensemble of Lightweight Models,” was presented at CVR-2025 (5th International Conference on Computer Vision and Robotics), held in hybrid mode at the National Institute of Technology Goa, India, from April 25–26, 2025.

Reprinted with permission. Copyright © 2025 CVR and SCRS.

Further use or distribution is not permitted without permission from CVR and SCRS.

## Abstract

The growing demand for healthy food, driven by population growth and the prevalence of plant diseases, is a significant concern. Farmers are increasingly forced to cultivate their fields continuously, often relying on pesticides that degrade soil quality and contribute to the spread of various diseases. Disruptive technologies, such as automated cassava leaf disease detection using deep learning models, can play a vital role in promoting agricultural sustainability. While numerous deep learning methods have been explored, challenges remain, particularly in developing lightweight models that can quickly and accurately identify disease class variations while also being suitable for deployment on electronic devices. This research focuses on evaluating lightweight deep learning models, specifically CNN and transformer networks, for cassava leaf disease detection. Based on our experiments, we propose an ensemble model consisting of lightweight models (CNN and transformer) for automated early disease detection using raw images. We have utilized the ResNeXt, EfficientNet-B5, and TinyViT lightweight models, optimizing the ensemble process through brute force approaches. Additionally, we experimented with centre-crop and multi-crop image transformations to test these models, aiming to enhance performance and classify images based on the global information derived from the whole image. Our combined approach achieved state-of-the-art results with an overall recall of 90.35% on the unseen test cassava leaf disease dataset. With high accuracy, fewer parameters, and low computation time, our model is well-suited for deployment on mobile devices.

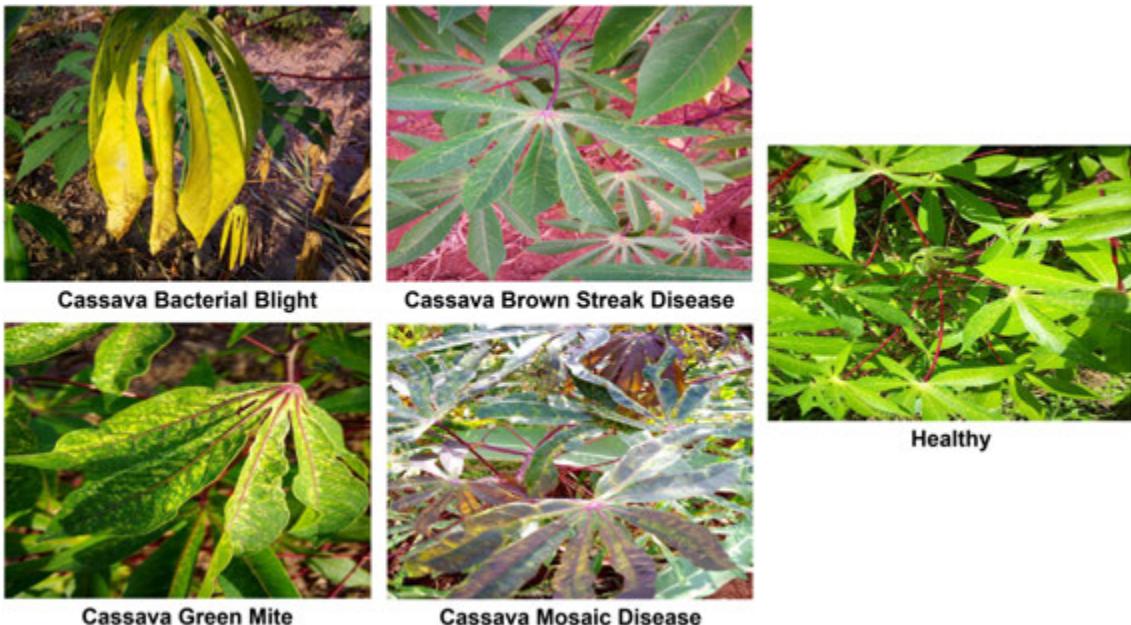
## 1

## INTRODUCTION

Cassava is a widely cultivated staple crop in sub-Saharan Africa, providing essential nutrition and energy. While the leaves and roots are both suitable for consumption, the roots are the preferred choice for their high energy content and versatility—they can be eaten raw, roasted, boiled, or processed in various ways<sup>[1][2]</sup>. The leaves and tender shoots are rich in protein and vitamins and are used as vegetables in many regions. Farmers across Africa grow cassava at different scales and under diverse geographical and weather conditions, contributing significantly to food stability and industrial development. Although widely cultivated, cassava is highly prone to

various diseases, including infections caused by lesser-known viral viruses. The four most common diseases, each with distinct signs and symptoms, are listed below and illustrated in Figure 1.

**Cassava Bacterial Blight (CBB):** CBB is one of the most devastating diseases, causing the highest yield losses among cassava crops worldwide. In the 1970s, it severely impacted Africa, with countries like Nigeria, Uganda, and Zaire suffering crop losses of 75% to 100%, leading to malnutrition and starvation in some regions. The disease thrives in moist conditions, affecting cassava plants in such areas. Symptoms include black leaf spots and blights, with affected leaves wilting, drying prematurely,



**Fig. 1** Provide sample images of the four major cassava leaf diseases together with healthy cassava leaf sample image from the dataset

and eventually shedding off.

**Cassava Green Mite (CGM):** CGM is a pest originating from South America, was introduced to Africa in the early 1970s and has now spread to 27 countries. It creates white spots on cassava leaves, beginning with small lesions that progressively expand, covering the entire leaf. This leads to chlorophyll loss and hinders photosynthesis. In cases of severe infestation, mottling symptoms may appear similar to those of cassava mosaic disease, making differentiation between the two difficult. Infested leaves commonly lose moisture, shrink, and break away from the plant, contributing to further damage.

**Cassava Brown Streak Disease (CBSD):** CBSD, which emerged in Eastern Africa in the 1930s, remains a major threat to cassava production, causing root yield losses of up to 70%. It is transmitted by whiteflies and is characterized by yellowing along the veins, which may expand into larger yellow patches. CBSD severely impacts tuber roots, forming dark-brown necrotic lesions that reduce root size and degrade the quality of the crop.

**Cassava Mosaic Disease (CMD):** CMD was first reported in 1894 by War-burg in Tanzania. It remains a significant

constraint to cassava production in Africa, contributing to famine and, in some regions, even death. The disease typically causes foliar symptoms like mottling, mosaic patterns, leaf deformation, and a significant decrease in both leaf and plant size. Affected leaves often display green patches mixed with yellow or white discoloration, diminishing the surface area for photosynthesis and resulting in poor growth and crop yields.

Distinguishing between different cassava diseases is challenging and often requires farmers' assistance, which is both time-consuming and resource-intensive<sup>[3]</sup>. The inability to detect infections early can lead to significant economic losses, highlighting the need for automated approaches to assist with timely disease detection and prevention. While several machine learning models have been explored in this area, recent deep learning models have delivered state-of-the-art (SOTA) results. However, these models typically require a large number of parameters, increasing their computational cost, and making them unsuitable for resource-constrained devices like mobile phones. In our research, we focused on lightweight deep-learning models that use fewer parameters while maintaining high accuracy and efficiency. We experimented with two architectures,

namely, convolutional neural networks (CNN) and transformers, and proposed an ensemble model that combines the strengths of both types of networks in lightweight constrained. During testing, the input image is cropped at the centre for specific model (CNNs/ transformer) input requirements. To mitigate information loss caused by centre-cropping when passing images through the model, a multi-crop approach is utilized. In this, each image is divided into five overlapping crops based on the model's input size, ensuring the entire image is covered. Each crop was then passed through the model to predict individual scores, which were subsequently merged to produce the final disease classification. Collectively, our ensemble model, combined with this image transformation approach, achieved an overall recall of 90.35% and an F1-score of 0.902 on the test dataset, setting a new state-of-the-art with fewer parameters and reduced computation time.

The primary contributions of this work are as follows:

- Conducted experiments on various lightweight models using two types of architectural, CNN-based and Transformer-based networks, for multi-class classification tasks.
- Analysis the performance of these models under two distinct testing strategies on the test datasets, namely centre-crop transformation and multi-crop transformation.
- We propose a novel ensemble model optimized using a brute-force algorithm, combined with multi-crop transformation testing strategies, which achieved state-of-the-art results on unseen test datasets.

## 2

## LITERATURE REVIEW

In artificial intelligence, two prominent types of vision-based frameworks—machine learning (ML) and deep learning (DL), are utilized for plant disease recognition. While ML-based models are generally lighter and less computationally intensive than DL-based models, their recognition accuracy often lags behind that of DL models<sup>[4]</sup>. Consequently, recent research has increasingly

shifted toward DL-based approaches, particularly for plant disease recognition. Among deep learning methods, Convolutional Neural Networks (CNNs) have demonstrated promising results in plant disease detection. This is evident in comparative studies evaluating various CNN architectures, such as VGGNet ResNet50<sup>[6]</sup>, DenseNet-BC-121-32<sup>[7]</sup>, and InceptionV3<sup>[8]</sup>, for cassava leaf disease classification. Among these architectures, DenseNet-BC-121-32 achieved the highest accuracy of 80.52%, followed by InceptionV3 at 77.25%, ResNet50 at 75.76%, while VGGNet exhibited comparatively lower performance. Further studies highlight the impact of dataset imbalance on model performance. For instance, experiments conducted on an imbalanced cassava disease dataset from Kaggle<sup>[9]</sup> demonstrated a 10% improvement in accuracy when employing Pulse- Coupled Neural Networks (PCNN) and Deep Residual Neural Networks (DRNN) compared to conventional CNN-based approaches. To enhance model robustness, the CNN-based Plant Disease Detection (PDD) framework incorporates data augmentation approaches and applies multilevel and multiscale feature extraction to establish a class- and scale-invariant architecture<sup>[10]</sup>. The PDD-Net model surpasses baseline architectures, achieving an overall of recall of 85.77%, precision of 86.41%, F1-score of 86.02%, and an accuracy of 86.98% on the cassava leaf disease dataset. Despite their success, CNN-based approaches are often computationally expensive and require substantial parameters, posing challenges for deployment on edge devices<sup>[11]</sup>. To address these limitations, an Enhanced Convolutional Neural Network (ECNN) was proposed for real-time cassava leaf disease identification<sup>[12]</sup>. The ECNN employs depth-wise separable convolutions to reduce both computational complexity and feature redundancy while maintaining classification accuracy.

The cassava dataset presents several challenges for training CNN models, primarily due to its small size and class imbalance, with certain disease classes being overrepresented. To address these issues and improve performance, a comparative study<sup>[13]</sup> is conducted to evaluate the Hybrid Ensemble Disease Detection Model

(HEDTM) against custom CNN architectures. This ensemble model, which integrates InceptionV3, Xception, and DenseNet-BC-121-32, achieved the highest accuracy of 88.83%, surpassing other ensemble configurations such as ResNet50V2 + DenseNet-BC-121-32 and ResNet50 + ResNet50V2. The results of this study highlight the potential of leveraging the strengths of individual models within a unified ensemble approach. Recently, researchers have also been attracted toward vision transformer-based models to address the limitations of CNN-based approaches. In the study by<sup>[14]</sup>, the Vision Transformer (ViT) was employed instead of traditional CNNs for classifying cassava leaf diseases. Experimental results demonstrated that ViT achieved at least 1% higher accuracy than popular CNN models like EfficientNet and ResNet50d on the Cassava Leaf Disease Dataset, highlighting its potential superiority in leaf disease analysis. Building on ViT, proposed a deep learning method for identifying viral diseases in cassava leaves, achieving a classification accuracy of 90.02% on the private test set after applying K-Fold cross-validation<sup>[15]</sup>. Training transformer models demands substantial computational resources and large datasets. While they achieve strong results, their deployment on mobile devices remains challenging. Consequently,

lightweight CNNs have gained traction in plant disease detection due to their efficiency on resource-constrained devices like smartphones and edge devices. For example, MobileNet-V2 enhanced using a class activation map to achieve better cassava leaf disease classification<sup>[16]</sup>.

Thus, our objective is to propose a novel model that integrates the strengths of both the networks, CNN and transformer-based to achieve superior classification performance with reduced computational complexity, making it suitable for mobile and edge applications through ensemble learning.

### 3 EXPERIMENTAL SETUP

#### 3-1. Dataset

This dataset was initially released for the Kaggle Cassava Leaf Disease Classification Challenge<sup>[17]</sup>. It contains a total of 21,397 training images, divided into five categories: CBB (1,087 images), CGM (2,386 images), CBSD (2,189 images), CMD (13,158 images), and Healthy (2,577 images). The dataset is allocated into 80% for training and 20% for validation, as illustrated in Figure 2. In the data processing step, the images were cropped to 512×512 and normalized to minimize background noise and

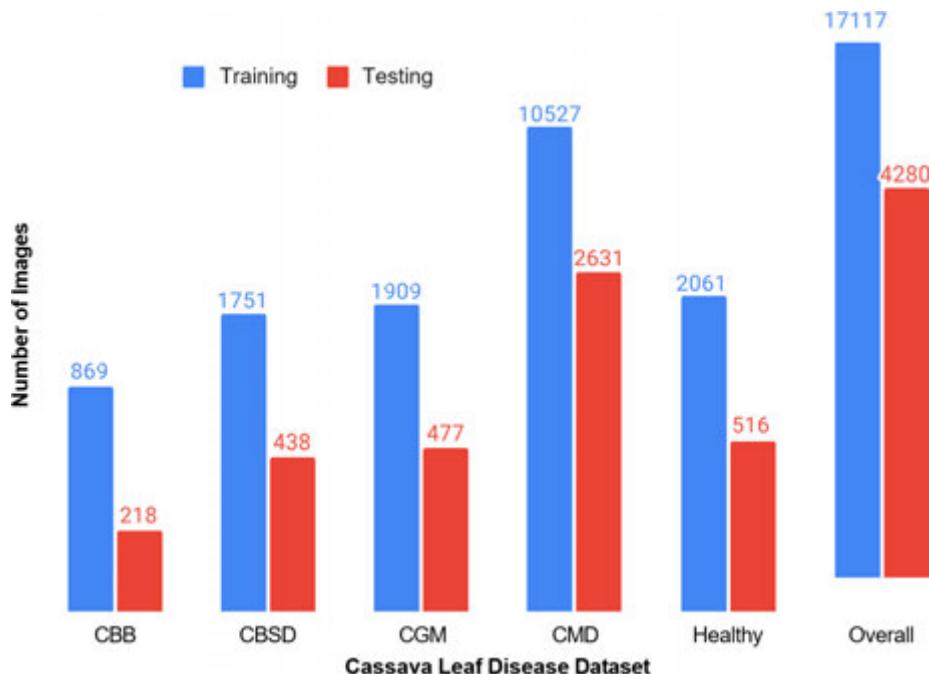


Fig. 2 The bar chart displays the distribution cassava leaf disease classification dataset

focus on critical features relevant to classification in training and testing. For model training, various transformations are applied to the training data to ensure generalized and stable model training, including transposition, horizontal flip, vertical flip, shift-scale rotation, and normalization. Since resizing from  $800 \times 600$  to  $512 \times 512$  can sometimes remove infected parts of the image affected by the disease. Therefore, we adopted a technique that was used to create five  $512 \times 512$  crops from each original image. To ensure the model had a comprehensive view for accurate disease detection during testing, five crops were taken from each image: four from the corners and one from the centre, as illustrated in Figure 3. The final prediction for each image was generated by averaging the class-wise probabilities across all five cropped regions. We evaluated the model's performance using both cropping strategies, which we refer to as centre-crop and multi-crop. Furthermore, to maximize the use of the training data, we implemented K-fold cross-validation using a 5-fold approach, dividing the dataset into five equal subsets. In each iteration, one subset was designated for validation, while the remaining four were used for training. This process was repeated five times, ensuring each subset served as the validation set once and as part of the training set four times. This method provides a comprehensive evaluation of each model's performance and generalization ability.

### 3-2. Models

In this research, we experimented with a lightweight deep-learning architecture designed to use fewer parameters while achieving higher accuracy in classifying cassava leaf diseases. Specifically, we utilized two types of models: CNN-based and transformer-based architecture, as detailed below:

**ResNeXt<sup>[18]</sup>:** ResNeXt, short for residual networks with external transformations, is a CNN-based architecture that builds the foundational principles of the ResNet (Residual Network) model. Unlike ResNet, which relies on multiple smaller paths, ResNeXt introduces groups of parallel paths. Each group contains multiple paths, with each path learning different features. This structured grouping allows the network to capture a broader range of features more effectively, eventually enhancing its overall classification of power. We fine-tuned ResNeXt50 using two distinct approaches. In the first approach, we directly added a classification layer, resulting in a model we refer to as ResNeXt50 23M. This model contains 22,990,149 parameters (approximately 23 million) and requires 90 MB of storage. In the second approach, we extended the fine-tuning by incorporating additional layers between ResNeXt50 and the final classification layer to facilitate more gradual transfer learning. This extended model, named ResNeXt50 26M, comprises 25,676,613 parameters (approximately 26 million) and occupies 100

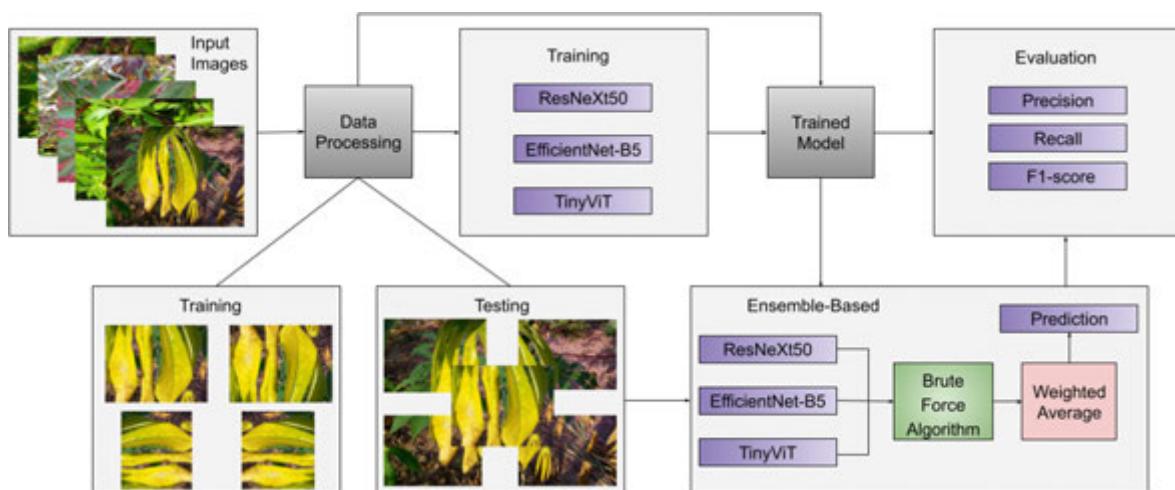


Fig. 3 Pipeline of the proposed ensemble model with a unique image transformations technique for training and testing images

MB of storage. These models were employed to evaluate the impact of parameter size on classification performance and efficiency.

**EfficientNet-B5<sup>[19]</sup>:** EfficientNet is a CNN architecture that uses a compound scaling method to proportionally adjust width, depth, and resolution, achieving both high accuracy and computational efficiency. The EfficientNet family provides a spectrum of models that balance complexity and accuracy, achieving strong classification performance with fewer parameters and reduced computational demands, while maintaining efficiency. Similarly, we utilized two variants of the EfficientNet B5 model to evaluate the impact of parameter size on inference time and overall efficiency. The first variant, similar to the fine-tuning approach of ResNeXt50 23M, is named EfficientNetB5 28M. It contains 28,351,029 parameters (approximately 28 million) and occupies 108.15 MB of storage. The second variant, similar to ResNeXt50 26M, is an extended version named EfficientNet-B5 31M. It features 31,037,493 parameters (approximately 31 million) and occupies 120 MB of storage. These models were used to assess the trade-off between parameter size and the lightweight model's performance in terms of speed and accuracy.

**TinyViT<sup>[20]</sup>:** TinyViT is a family of compact and efficient vision transformers designed for high performance with minimal computational resources. Pretrained on large-scale datasets, TinyViT employs a fast distillation framework, where a large pre-trained model (teacher) is scaled down into a smaller model (student) while preserving its core capabilities. This student-teacher distillation allows TinyViT to transfer knowledge effectively across various downstream tasks, delivering accurate results with fewer parameters and reduced memory requirements, making it ideal for resource-constrained environments. Similarly, we applied the same fine-tuning approaches to this model. The first approach is designated as TinyViT 20.6M, which contains 20,607,989 parameters and occupies 78.61 MB of storage. The second extended approach is referred to as TinyViT 20.8M, featuring 20,788,277 parameters and requiring 79.30 MB of storage. These experiments were carried out to assess the effect of small changes in

parameter size on performance and efficiency.

### 3-3. Proposed Ensemble Model

This research proposes an ensemble method that integrates CNN-based models (ResNeXt, EfficientNet-B5) with a transformer-based model (TinyViT) to leverage their respective strengths and improve disease classification accuracy as shown in Figure 3. Leveraging multiple models typically boosts accuracy and robustness over individual models.

In this work, each model was underwent trained on the cassava leaf disease dataset and was fine-tuned with different hyperparameter settings. Their classification accuracy was analyzed, and performance-based weights were assigned using the brute force algorithm<sup>[21]</sup>. The brute force approach methodically explores all possible combinations to determine the optimal weights for the ensemble, ensuring accurate and fast results, making it ideal for fewer model combinations.

The final predictions were made using a weighted average of the base models' predictions, with higher weights assigned to more accurate models. This ensemble approach mitigates the weaknesses of individual models by combining their predictions, as illustrated in Figure 3. Each model was trained separately, and their respective weights were used for evaluation. The ensemble prediction utilized various weight combinations to identify the optimal set that maximized accuracy. By incorporating the strengths of each model, this method enhances overall prediction performance rather than disregarding the limitations of any single model.

$$P = \arg \max \sum_{i=1}^n x_i^i y \quad (1)$$

Here,  $P$  represents the optimal prediction from the ensemble model, where  $x_i$  denotes the weight assigned to each model's output, which is multiplied by the corresponding weight vector  $y$ . The variable  $n$  refers to the total number of models included in the ensemble.

### 3-4. Evaluation Metrics

We evaluate the performance of the models for each

class using various metrics such as precision, recall, and F1-Score, which are calculated using Eq. (2), (3), and (4), respectively. Where True Positives, False Positives, False Negatives, and True Negatives are represented as TP, FP, FN, and TN, respectively. For the overall evaluation of the model's performance on multi-unbalanced classes, we employ the metrics from<sup>[22]</sup>, as described in Eq. (5) rather than the previous average evaluation method<sup>[10][13]</sup>. Where  $N$  refers to the total number of classes, and  $n_i$  represents the fraction of samples of particular class to the total number of samples.  $M_i$  denotes the value of the metric  $M$  for class  $i$ , where the metric  $M$  could be precision, recall, or F1-Score. To compare the performance of the different model variants, we examined the overall recall value. This metric is particularly important in this context, as the cost of missing an infected leaf (FN) is critical, potentially leading to the spread of disease to healthy leaves.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

$$\text{Overall} = \frac{\sum_{i=1}^N n_i \times M_i}{\sum_{i=1}^N n_i}, \quad (5)$$

We also measured the inference time of various models using two test strategies: centre-crop and multi-crop on the test images. The inference time, representing the duration required for model execution, was recorded in seconds. For the ensemble model, the inference time was measured to be approximately equal to that of the model with the maximum inference time.

Table 1 Comparison of the two ResNeXt50 model variants, with 23M and 26M parameters, in terms of overall recall (%) using two testing strategies on the multi-class test classification dataset

Fold	ResNeXt50 23M		ResNeXt50 26M	
	Centre-Crop	Multi-Crop	Centre-Crop	Multi-Crop
1	88.00	<b>88.43</b>	87.34	87.82
2	87.31	87.52	87.61	87.50
3	87.94	88.38	87.73	88.06
4	87.33	87.47	87.73	87.99
5	<b>88.10</b>	88.27	87.78	87.89

Table 2 Comparison of the two EfficientNet-B5 model variants, with 28M and 31M parameters, in terms of overall recall (%) using two testing strategies on the multi-class test classification dataset

Fold	EfficientNet-B5 28M		EfficientNet-B5 31M	
	Centre-Crop	Multi-Crop	Centre-Crop	Multi-Crop
1	87.32	88.24	88.48	88.80
2	88.00	88.43	87.96	88.17
3	87.82	88.43	<b>88.57</b>	88.83
4	87.59	88.48	88.36	<b>89.10</b>
5	87.25	87.85	87.96	88.31

Table 3 Comparison of the two TinyViT model variants, with 20.6M and 20.8M parameters, in terms of overall recall (%) using two testing strategies on the multi-class test classification dataset

Fold	TinyViT 20.6M		TinyViT 20.8M	
	Centre-Crop	Multi-Crop	Centre-Crop	Multi-Crop
1	89.00	88.99	88.50	88.76
2	<b>89.32</b>	89.32	89.00	88.41
3	89.27	<b>89.51</b>	88.78	89.23
4	88.69	88.69	89.29	89.00
5	88.85	89.15	88.71	88.41

Table 4 Comparison of various lightweight models and their parameter variants in terms of inference time (seconds) for the two testing strategies on the multi-class test classification dataset

Model	Inference Time (seconds)	
	Centre-Crop	Multi-Crop
ResNeXt50 23M	1.05734	5.52274
ResNeXt50 26M	1.10840	5.56961
EfficientNet-B5 28M	<b>0.94257</b>	5.20375
EfficientNet-B5 31M	1.35552	<b>4.81913</b>
TinyViT 20.6M	0.95797	5.87840
TinyViT 20.8M	0.99867	5.22033

## 4 RESULTS AND DISCUSSION

The results of experiments conducted with various lightweight models are detailed in this section. Table 1 compares the centre-crop and multi-crop classification performance on test images using two variants of the ResNeXt50 model with different parameter sizes. The smaller variant, ResNeXt50 23M, achieved the highest overall recall of 88.40% with the multi-crop strategy, outperforming the larger ResNeXt50 26M. However, when compared to ResNeXt, another CNN-based model, EfficientNet-B5 31M, showed an improvement of 0.83%,

achieving an accuracy of 89.10% using the multi-crop strategy, as presented in Table 2. Moreover, despite having more parameters, EfficientNet-B5 required less execution time than ResNeXt, as shown in Table 4. This demonstrates that EfficientNet-B5 not only outperforms ResNeXt in terms of accuracy but also offers greater computational efficiency for multi-class classification tasks.

**Table 5** Display the comparison results of the best-performing lightweight models and our ensemble models across different evaluation metrics for each class and the overall testing dataset

Method	Class	Precision	Recall	F1-score
ResNeXt	CBB	0.6762	0.6544	0.6651
	CGM	0.8112	0.7925	0.8017
	CBSD	0.8624	0.7443	0.799
	CMD	0.9444	0.9738	0.9588
	Healthy	0.7344	0.7287	0.7315
	Overall	0.8822	0.8843	0.8827
EfficientNetB5	CBB	0.7095	0.5853	0.6414
	CGM	0.8271	0.7925	0.8094
	CBSD	0.868	0.7808	0.8221
	CMD	0.9471	0.9802	0.9634
	Healthy	0.7319	0.7461	0.739
	Overall	0.8877	0.891	0.8884
TinyViT	CBB	0.6652	0.6959	0.6802
	CGM	0.7984	0.8134	0.8058
	CBSD	0.8722	0.8105	0.8402
	CMD	0.9581	0.9723	0.9651
	Healthy	0.773	0.7326	0.7522
	Overall	0.8943	0.8951	0.8945
Ensemble	CBB	0.7198	0.6866	0.7028
	CGM	0.8398	0.8134	0.8264
	CBSD	0.8909	0.8014	0.8438
	CMD	0.9556	0.9806	0.9679
	Healthy	0.7713	0.7713	0.7713
	Overall	<b>0.9019</b>	<b>0.9035</b>	<b>0.9023</b>

On the other hand, the vision transformer-based model, TinyViT, outperforms both CNN-based models in terms of inference time and accuracy when using the centre-crop testing strategy, as detailed in Tables 3 and 4. Using the multi-crop strategy, TinyViT achieved 0.41% and 1.08% higher accuracy than ResNeXt and EfficientNet-B5, respectively, with an accuracy of 89.51% using only 20.6M parameters. These results indicate that the lightweight transformer-based model for multi-class classification tasks outperforms the lightweight CNN-based models in terms of efficiency.

Based on the above results, we developed an ensemble model combining the best-performing variants of ResNeXt, EfficientNet-B5, and TinyViT. Interestingly, our multi-crop testing strategy consistently outperformed the centre-crop approach. This can be attributed to the fact that the diseased portion of the image may not always be located in the centre, and relying solely on the centre-crop could result in missing infected regions. To minimize the risk of undetected infections, which could lead to disease spreading to healthy plants, we employed the multi-crop strategy for testing our ensemble model. As shown in Table 5, our ensemble model achieved an F1-score of 0.902 and an overall recall of 90.35%, surpassing the performance of the individual models and setting a new SOTA on the test dataset. The enhanced performance is attributed from the fact that different models tend to make errors on different data samples. One model might excel in learning certain features that others may not. This diversity among the models enhances the robustness of the ensemble, effectively reducing variance in prediction errors and yielding more reliable results.

## 5 CONCLUSION

This research concentrated on building an automated framework to detect and classify cassava leaf diseases at an early stage. The existing methods ensemble either CNN or Transformer models leading to gain in either computational complexity or efficiency. To overcome the challenges of high memory usage and computational demands, we explored various CNN and Transformer-based lightweight models, including ResNeXt and EfficientNet-B5 (CNN), and TinyViT (Transformer). To achieve the best results, we applied centre and multi-crop transformation strategies to the testing dataset and utilized a cross-validation approach to fully utilize the entire dataset for training and testing. We optimized the ensemble method by using a brute force approach to determine the optimal weighted averaging of the models ideal to ensemble few models. Our experimental findings give a novel ensemble model that combines the strengths of transformation strategies with both CNN and Transformer architectures, delivering superior

performance with minimal memory and computational requirements. Although it involves a trade-off in computation time compared to individual models, it remains more efficient than earlier ensemble methods that relied solely on CNN or transformer-based models without considering computational complexity. In future work, we aim to validate this approach on devices like smartphones and agricultural robots for real-time detection of plant diseases in the field.

## REFERENCES

---

- [1] Huang, K. Y.: Application of artificial neural network for detecting phalaenopsis seedling diseases using color and texture features. *Computers and Electronics in agriculture* 57(1), 3–11 (2007)
- [2] Taiwo, K. A.: Utilization potentials of cassava in nigeria: the domestic and industrial products. *Food Reviews International* 22(1), 29–42 (2006)
- [3] Bock, C., Poole, G., Parker, P., Gottwald, T.: Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. *Critical reviews in plant sciences* 29(2), 59–107 (2010)
- [4] Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J., Hughes, D. P.: Deep learning for image-based cassava disease detection. *Frontiers in plant science* 8, 1852 (2017)
- [5] Wang, L., Guo, S., Huang, W., Qiao, Y.: Places205-vggnet models for scene recognition. *arXiv preprint arXiv:1508.01667* (2015)
- [6] Theckedath, D., Sedamkar, R.: Detecting affect states using vgg16, resnet50 and se-resnet50 networks. *SN Computer Science* 1(2), 79 (2020)
- [7] Zhang, K., Guo, Y., Wang, X., Yuan, J., Ding, Q.: Multiple feature reweight densenet for image classification. *IEEE access* 7, 9872–9880 (2019)
- [8] Xia, X., Xu, C., Nan, B.: Inception-v3 for flower classification. In: 2017, 2nd ICIVC. pp. 783–787. IEEE (2017)
- [9] Mwebaze, E., Gebru, T., Frome, A., Nsumba, S., Tusubira, J.: icassava 2019 fine-grained visual categorization challenge. *arXiv preprint arXiv:1908.02900* (2019)
- [10] Alghamdi, H., Turki, T.: Pdd-net: Plant disease diagnoses using multilevel and multiscale convolutional neural network features. *Agriculture* 13(5), 1072 (2023)
- [11] Zhang, S., Zhang, S., Zhang, C., Wang, X., Shi, Y.: Cucumber leaf disease identification with global pooling dilated convolutional neural network. *Computers and Electronics in Agriculture* 162, 422–430 (2019)
- [12] Lilhore, U. K., Imoize, A. L., Lee, C. C., Simaiya, S., Pani, S. K., Goyal, N., Kumar, A., Li, C. T.: Enhanced convolutional neural network model for cassava leaf disease identification and classification. *Mathematics* 10(4), 580 (2022)
- [13] Kiruthika, V., Shoba, S., Sendil, M., Nagarajan, K., Punetha, D.: Hybrid ensemble- deep transfer model for early cassava leaf disease classification. *Heliyon* (2024)
- [14] Thai, H. T., Tran-Van, N. Y., Le, K. H.: Artificial cognition for early leaf disease detection using vision transformers. In: 2021 International Conference on Advanced Technologies for Communications (ATC). pp. 33–38. IEEE (2021)
- [15] Zhuang, L.: Deep-learning-based diagnosis of cassava leaf diseases using vision transformer. In: Proceedings of the 2021 4th Artificial Intelligence and Cloud Computing Conference. pp. 74–79 (2021)
- [16] Chen, J., Zhang, D., Nanehkaran, Y. A.: Identifying plant diseases using deep transfer learning and enhanced lightweight network. *Multimedia tools and applications* 79, 31497–31515 (2020)
- [17] Mwebaze, M., Joyce, Julia, D. S.: Cassava leaf disease classification (2020)
- [18] Xie, S., Girshick, R., Dollár, P., Tu, Z., He, K.: Aggregated residual transformations for deep neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1492–1500 (2017)
- [19] Tan, M., Le, Q.: Efficientnet: Rethinking model scaling for convolutional neural networks. In: International conference on machine learning. PMLR (2019)
- [20] Wu, K., Zhang, J., Peng, H., Liu, M., Xiao, B., Fu, J., Yuan, L.: Tinyvit: Fast pretraining distillation for small vision transformers. In: European conference on computer vision. pp. 68–85. Springer (2022)
- [21] Morton, A. B., Mareels, I. M.: An efficient brute-force solution to the network reconfiguration problem. *IEEE Transactions on Power Delivery* 15(3), 996–1000 (2000)
- [22] Grandini, M., Bagli, E., Visani, G.: Metrics for multi-class classification: an overview. *arXiv preprint arXiv:2008.05756* (2020)

■著者



**Kaushik Vishwakarma**  
Indian Institute of Technology,  
Mandi/  
Emerging Technology and  
Innovation Lab, Yamaha Motor  
Solutions India



**Udaybhan Rathore**  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India



**Rupesh Kumar**  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India



**Amit Shakya**  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India



**Chetan Gupta**  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India



**Shruti Phutke**  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India



**Lalit Sharma**  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India

# A Novel Two-Stage Deep Learning Method for Enhanced Cell Nuclei Segmentation

Chetan Gupta Rupesh Kumar Amit Shakya Shruti Phutke Lalit Sharma

This paper, titled “A Novel Two-Stage Deep Learning Method for Enhanced Cell Nuclei Segmentation,” was presented at CVR-2025 (5th International Conference on Computer Vision and Robotics), held in hybrid mode at the National Institute of Technology Goa, India, from April 25–26, 2025.

Reprinted with permission. Copyright © 2025 CVR and SCRS.

Further use or distribution is not permitted without permission from CVR and SCRS.

## Abstract

Cell nuclei segmentation is essential for microscopic image analysis. It facilitates the detailed micro-environmental insights for clinical studies. Automated nuclei segmentation can simplify the work of pathologists and address the variability and subjectivity among them, thereby improving diagnostic consistency and accuracy. Although deep learning (DL) techniques usually offer better performance than traditional methods for nuclei segmentation, they still struggle with challenges especially when the nuclei are clustered and overlapped with each other. To address these challenges, many effective deep learning techniques have been developed. Nonetheless, these approaches still exhibit limitations, such as the tendency to overlook certain nuclei. To mitigate this issue, we propose a two-stage network to boost accuracy by incorporating an enhancement network (second stage) on top of the widely used encoder-decoder architectures (first stage). The enhancement network refines the results by utilizing the decoder’s output and original image. The primary goal of this proposed method is to segment those regions that might have been overlooked by the base model, by re-considering the original image within the network. The study employs popular models such as U-Net, Micro-Net, and U-Net++ as base models. The results illustrate improvement of  $0.8 \pm 1.8\%$  in Precision,  $1.5 \pm 3.0\%$  in Recall,  $1.4 \pm 0.5\%$  in Dice Score, and  $1.6 \pm 0.7\%$  in IoU across different datasets through the proposed method.

## 1 INTRODUCTION

Cancer is the second leading cause of death globally which make the early detection crucial for better treatment<sup>[1]</sup>. Morphological changes in nuclei served as the biomarker for cancer diagnosis<sup>[2]</sup>. Currently, pathologists use microscopes to examine tissue samples, a method that is slow, labor-intensive and dependent on human skills<sup>[3]</sup>. Recent Research shows that pathologists often disagree on diagnoses, with a 24.7% disagreement rate<sup>[4]</sup>. It highlights the critical necessity for computer-assisted pathology to enhance diagnostic consistency and precision. Cell nuclei segmentation is the fundamental task in computer-assisted pathology, involving the identification and outlining of cell nuclei in microscopic images<sup>[5]</sup>.

Traditional methods like Watershed segmentation<sup>[6]</sup>, K-means clustering<sup>[7]</sup>, and fuzzy C-means<sup>[8]</sup> have been used for this purpose but often struggle with blurry images<sup>[9]</sup>. As a result, researchers are seeking more effective and accurate solutions. Deep learning has proven to be a highly effective approach for segmenting cell nuclei, as it can automatically learn and extract features from images, resulting in more precise and reliable segmentation outcomes<sup>[10]</sup>. Unlike conventional methods, deep learning models excel in handling nuclei across a variety of image conditions. This flexibility is essential in biomedical imaging, where nuclei can differ significantly in size, shape, and intensity. Although single-stage training methods<sup>[11]-[15]</sup> are commonly used for medical image segmentation, some researchers have found the two-stage approach to be highly effective and promising. Existing methods often rely on specific

architectures, for instance, Jiang et al.<sup>[16]</sup> utilized a modified U-Net in a two-step process, where an initial simplified U-Net provides a preliminary result, followed by an enhanced model with additional decoding layers for refinement. In contrast, our method is flexible and independent of any particular architecture. It can integrate with any model in the first stage, offering adaptable and improved cell nuclei segmentation without being restricted to a specific design. In medical diagnostics, the structure and count of cell nuclei are crucial factors for decision-making<sup>[17]</sup>. For a cell nuclei segmentation model, one aspect is to accurately delineate the region of interest (cell nuclei) and another is to precisely segment the correct number of nuclei. However, there remains a gap between the quality and quantity of segmented nuclei. To bridge this gap, our paper introduces a two-stage cell nuclei segmentation approach falling under the category of representation enhancement. The proposed method uses existing encoder-decoder segmentation networks as the base model (first stage), followed by an enhancement network (second stage) to refine the initial prediction mask produced by the base network. This enhancement network is a lightweight model with approximately 77k trainable weights that can be easily integrated with the base model and produce better quantitative and qualitative results, as shown in the results section later. Notably, the proposed method not only captures initially missed cell nuclei but also refines the pixels along the boundaries of segmented nuclei. Our main contributions

are as follows:

- We propose a novel two stage deep learning approach for cell nuclei segmentation which effectively enhances the segmentation results of existing methods.
- The proposed pipeline is architecture independent, allowing it to integrate various advanced segmentation methods in the first stage, followed by an enhancement network in the second stage. To illustrate its effectiveness, we evaluated it using three state-of-the-art segmentation models: U-Net<sup>[12]</sup>, Micro-Net<sup>[13]</sup>, and U-Net++<sup>[14]</sup>.
- The comparison of proposed method is carried out on two publicly available histopathological datasets.

The structure of the paper is as follows: Section 2 explains the proposed method, Section 3 details about the dataset used, evaluation metrics, and implementation of proposed approach, Section 4 discusses the results, and finally, Section 5 concludes the paper.

## 2

## PROPOSED METHOD

The proposed method, as shown in Figure 1, consists of two stages that aims to enhance the performance of the cell segmentation. The method uses an encoder-decoder network, followed by an enhancement network that improves the quality of initial prediction mask produced by the first network. In this section, deep convolutional encoder-decoder network has been explained that

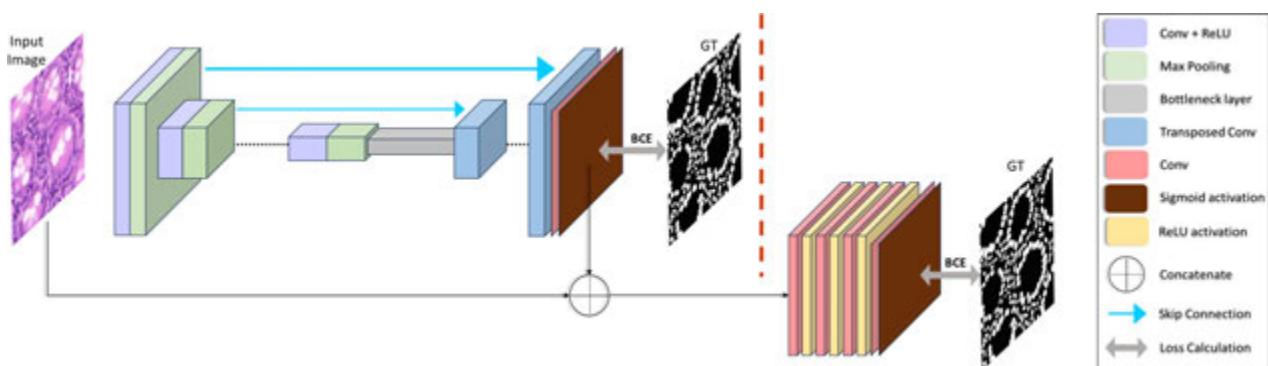


Fig. 1 Proposed method architecture. (a) first stage (architecture on the left side of red dashed line): an encoder-decoder based segmentation network. (b) second stage (architecture on the right side of red dashed line): enhancement network

constitutes the first stage of the proposed method. These networks take the original RGB image as input and produce a raw prediction mask. Subsequently, in the second stage, the enhancement network has been proposed, which refines the initial prediction mask, resulting in a more accurate and sharper edge mask.

### 2-1. Encoder-Decoder Stage

The first stage of our method comprises of a general segmentation network, as depicted in Figure 1(a). To comprehensively illustrate the effectiveness of the proposed method, we utilize three different popular segmentation networks U-Net, Micro-Net and U-Net++. U-Net<sup>[12]</sup> is a pioneering U-shaped encoder-decoder network specifically developed for efficient and accurate processing of biological images, even when trained on small datasets. It features a structured architecture with four encoder blocks that progressively down-sample the input image to capture feature representations, and four decoder blocks that up-sample these representations to produce the final segmentation. The skip connections between the encoder and decoder blocks help retain

spatial details and enhance segmentation accuracy. Building on the foundational design of U-Net, Micro-Net<sup>[13]</sup> introduces additional convolutional layers that bypass the traditional max-pooling layers. This modification allows Micro-Net to more effectively capture and learn both weak and strong features in fluorescence and H&E stained images, improving its ability to segment cells, nuclei, and glands. The inclusion of intermediate connections between layers helps to preserve contextual information and localization, further enhancing segmentation performance. U-Net++<sup>[14]</sup> takes the advancements of U-Net a step further by integrating deeply supervised learning. It employs a series of nested, dense skip pathways that connect the encoder and decoder sub-networks. These redesigned skip pathways aim to bridge the semantic gap between the feature maps generated by the encoder and those used in the decoder, resulting in more accurate and refined segmentation. Each of these models builds upon the previous innovations, contributing to the development of increasingly sophisticated and precise segmentation techniques.

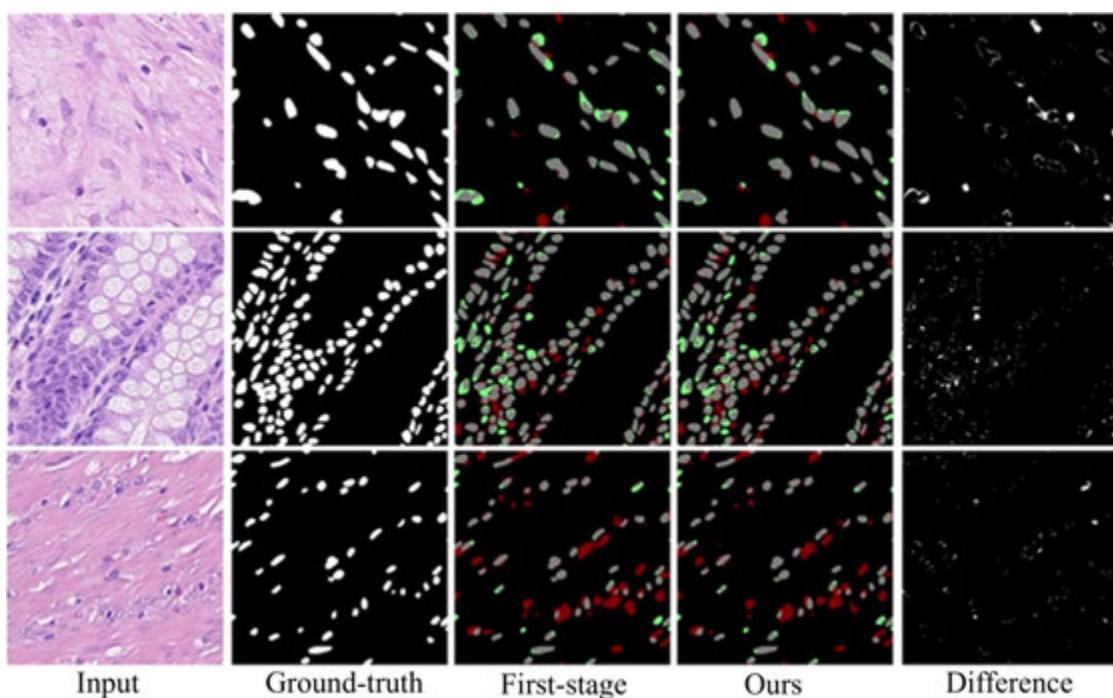


Fig. 2 Visual results of proposed refinement approach on Lizard dataset. **Red** indicates false positive pixels, **green** indicates false negative pixels and **gray** indicates true positive pixels. Fifth column - image obtained by difference between the refinement (ours) output and First-stage output. Images in lines 1, 2 and 3 are results from U-Net, Micro-Net, and U-Net++, respectively

## 2-2. Enhancement Stage

Enhancement stage is the second part of the proposed method, focusing on refining the initial prediction mask.

**Network structure:** The input image is concatenated channel-wise with the prediction from the first stage, creating a 4-channel input for the enhancement network. As illustrated in Figure 1(b), the network contains 4 convolutional layers, where each of the first three layers contains 64 filters of size  $3 \times 3$  followed by 'ReLU' activation layer. The final layer is a  $1 \times 1$  convolutional layer with 'sigmoid' activation. Notably, downsampling is deliberately omitted in each layer to preserve subtle structures that may have been missed in the first stage. As this enhancement stage is provided with the histopathological image and initial segmentation mask as input, the task here is to refine the segmentation results. This is done via extracting the features from the concatenated 4 channel input which already has localization of nuclei present in it. Processing this localized/referenced histopathological image with the proposed enhancement network helps to refine the earlier segmented nuclei by considering maximum receptive fields with the help of stack of convolution layers. The loss is calculated between the network's output and the corresponding ground truth.

**Table 1** Comparison of evaluation metrics of segmentation networks with and without the proposed method, trained on Lizard dataset<sup>[18]</sup> († is proposed two-stage approach)

	Precision	Recall	Dice	IoU
<b>U-Net</b>	0.810	0.768	0.786	0.665
<b>U-Net<sup>†</sup></b>	0.798	<b>0.818</b>	<b>0.806</b>	<b>0.689</b>
<b>Micro-Net</b>	0.790	0.774	0.780	0.655
<b>Micro-Net<sup>†</sup></b>	<b>0.804</b>	<b>0.780</b>	<b>0.790</b>	<b>0.664</b>
<b>U-Net++</b>	0.692	0.854	0.757	0.627
<b>U-Net++<sup>†</sup></b>	<b>0.715</b>	0.844	<b>0.769</b>	<b>0.642</b>

## 3

# EXPERIMENTATION

## 3-1. Dataset

We utilize two datasets to demonstrate the generalization potential of our approach.

**Lizard** We use Lizard dataset<sup>[18]</sup>, which is a large-scale dataset for colonic nuclear instance segmentation and classification. In total, there are 4981 image patches of size  $256 \times 256$  which contain nearly half a million labeled nuclei in H&E stained colon tissue. The dataset has nuclear class labels for epithelial cells, connective tissue cells, lymphocytes, plasma cells, neutrophils and eosinophils. However, for our binary segmentation task, we pooled these different cell types into a single class. We split the dataset into training set, validation set and test set, each consisting of 3842, 499, and 640 image-mask pairs, respectively.

**TNBC** Triple Negative Breast Cancer (TNBC) dataset<sup>[19]</sup>, consists of 50 images each of size  $512 \times 512$  with a total of 4022 annotated cells that include normal epithelial and myoepithelial breast cells (localized in ducts and lobules), invasive carcinomatous cells, fibroblasts, endothelial cells, adipocytes, macrophages and inflammatory cells (lymphocytes and plasmocytes). In our binary segmentation study, we group these diverse cell types together as a single foreground class. We divided dataset into three subsets: training set, validation set and test set, each containing 30, 10, and 10 image-mask pairs, respectively. After splitting the data, we addressed limitations of small dataset by applying data augmentation techniques like random horizontal flip, rotation and zoom, individually to each set.

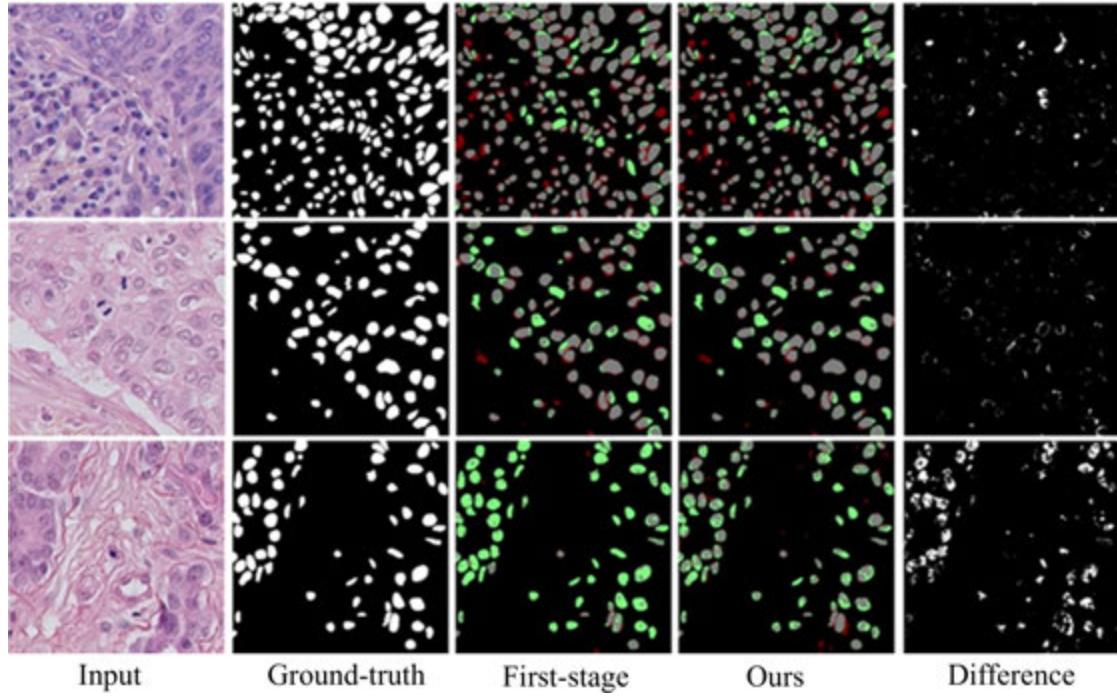


Fig. 3 Visual results of proposed refinement approach on TNBC dataset. **Red** indicates false positive pixels, **green** indicates false negative pixels and **gray** indicates true positive pixels. Fifth column - image obtained by difference between the refinement (ours) output and First-stage output. Images in lines 1, 2 and 3 are results from U-Net, Micro-Net, and U-Net++, respectively

### 3-2. Loss Function

During the training of both stages in the proposed method, we employ binary cross-entropy (BCE) loss.

$$BCE(y_i, p(y_i)) = -\frac{1}{N} \sum_i (y_i \log(p(y_i)) + (1 - y_i) \log(1 - p(y_i))) \quad (1)$$

We calculate pixel-wise binary cross-entropy loss between actual pixel value  $y_i$  and predicted pixel value  $p(y_i)$ .

Table 2 Comparison of evaluation metrics of segmentation networks with and with- out the proposed method, trained on TNBC dataset<sup>[19]</sup> († is proposed two-stage approach)

	Precision	Recall	Dice	IoU
<b>U-Net</b>	0.800	0.857	0.827	0.705
<b>U-Net<sup>†</sup></b>	<b>0.840</b>	0.826	<b>0.832</b>	<b>0.708</b>
<b>Micro-Net</b>	0.759	0.828	0.789	0.653
<b>Micro-Net<sup>†</sup></b>	<b>0.807</b>	0.805	<b>0.804</b>	<b>0.673</b>
<b>U-Net++</b>	0.782	0.665	0.649	0.519
<b>U-Net++<sup>†</sup></b>	<b>0.801</b>	<b>0.703</b>	<b>0.721</b>	<b>0.577</b>

### 3-3. Evaluation Metrics

To evaluate the performance of the proposed method, we

employ the commonly used metrics, including Precision, Recall, Intersection over Union and Dice coefficient. Precision reflects the quality of results, while recall indicates the quantity of relevant items identified. Intersection over Union (IOU) measures the ratio of overlapping pixels to the total pixels in two compared images, with a value ranging from 0 to 1, 0 indicates no overlap, while 1 signifies complete overlap while the Dice score shows how similar two binary images are, like comparing a predicted mask to the actual mask. It is calculated by taking twice the amount of overlap between them and dividing by the total number of pixels in both images. A score of 1 means they match perfectly, while 0 means no overlap.

$$\text{Precision} = \frac{tp}{tp + fp} \quad (2)$$

$$\text{Recall} = \frac{tp}{tp + fn} \quad (3)$$

$$\text{Dice Score} = \frac{2 \cdot tp}{2 \cdot tp + fp + fn} \quad (4)$$

$$\text{IoU} = \frac{tp}{tp + fp + fn} \quad (5)$$

Here,  $tp$  and  $tn$  denote the total counts of accurately identified positive and negative cases. On the other hand,  $fn$  and  $fp$  indicate the counts of incorrectly identified negative and positive cases, respectively.

### 3-4. Implementation Details

As mentioned in the section 3.1, our proposed method is trained on two different datasets. The RGB images were first normalized to a scale of 0 to 1 before being fed to the model. The entire network undergoes a two-stage training process. After both parts of the network converge, the entire network is fine-tuned in a single-stage training session.

During first stage training, we first trained the encoder-decoder network without the enhancement network. After the encoder-decoder part converges, we freeze it and then update the enhancement network. After the enhancement part also converges, finally we fine-tune the whole network together.

We use the Adam optimizer with a learning rate of 0.0001 and apply binary cross-entropy loss, mentioned in section 3.2, during each training session. Additionally, each training is facilitated by early stopping technique which halts the training once the validation loss stops improving.

## 4 RESULTS

---

Table 1 and Table 2 showcase the evaluation scores for precision, recall, dice coefficient and IoU on Lizard and TNBC dataset, respectively. The symbol  $\dagger$  on a segmentation network represents the combination of the encoder-decoder network followed by the enhancement network. In the tables, scores highlighted in blue indicate improvements in metrics with the proposed method. In each architecture, U-Net, Micro-Net and U-Net++, the integration of the enhancement network leads to improvements in at least three evaluation metrics. It can be noticed that the Dice coefficient improved in each case.

Figures 2 and 3 provide a clear visual analysis of the difference in the predictions of segmentation network with and without the enhancement network for Lizard and TNBC dataset respectively. The enhancements in the evaluation scores presented in the tables are visually reflected in Figure 2 and 3. The 'Difference' column is derived by subtracting the images in the 'First Stage Prediction' column from those in the 'Ours' column. White pixels in the images of the 'Difference' column indicate instances where the proposed method successfully identifies complete cells or parts of cells that were overlooked by the first stage network. It is important to note that some white pixels in these images may also result from false positive pixels generated by the proposed method, although they are relatively few in number.

## 5 CONCLUSION

---

Histopathological Images are complex in nature which makes segmentation of cell nuclei a challenging task. In this paper, we have proposed a method to integrate an enhancement network atop widely used existing segmentation networks. To showcase the generalizability of the proposed method, we apply it to diverse datasets, including Lizard and TNBC, across popular segmentation architectures like U-Net, Micro-Net, and U-Net++. The quantitative and qualitative results indicate that the proposed method not only refines pixels along the cell boundary but also segments cells that were initially missed by the first stage network. In the future, we would like to extend our experiments to encompass additional datasets across other popular segmentation architectures with the aim to improve the quality of output generated by the enhancement network.

## REFERENCES

---

- [1] Mortality, G., of Death Collaborators, C.: Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the global burden of disease study 2015. *The Lancet* 388, 1459-1544 (2016),

[https://doi.org/10.1016/S0140-6736\(16\)31012-1](https://doi.org/10.1016/S0140-6736(16)31012-1)

- [2] Singh, I., Lele, T. P.: Nuclear morphological abnormalities in cancer: a search for unifying mechanisms. In: Nuclear, chromosomal, and genomic architecture in biology and medicine, pp. 443-467. Springer (2022)
- [3] Onega, T., Barnhill, R. L., Piepkorn, M. W., Longton, G. M., Elder, D. E., Weinstock, M. A., Knezevich, S. R., Reisch, L. M., Carney, P. A., Nelson, H. D., et al.: Accuracy of digital pathologic analysis vs traditional microscopy in the interpretation of melanocytic lesions. *JAMA dermatology* 154(10), 1159-1166 (2018)
- [4] Hsu, W., Han, S. X., Arnold, C. W., Bui, A. A., Enzmann, D. R.: A data-driven approach for quality assessment of radiologic interpretations. *Journal of the American Medical Informatics Association* 23(e1), e152-e156 (2016)
- [5] Alom, Z., Asari, V. K., Parwani, A., Taha, T. M.: Microscopic nuclei classification, segmentation, and detection with improved deep convolutional neural networks (dcnn). *Diagnostic Pathology* 17(1), 38 (2022)
- [6] Beucher, S.: The watershed transformation applied to image segmentation. *Scanning microscopy* 1992(6), 28 (1992)
- [7] Ahmed, M., Seraj, R., Islam, S. M. S.: The k-means algorithm: A comprehensive survey and performance evaluation. *Electronics* 9(8), 1295 (2020)
- [8] Bezdek, J. C., Ehrlich, R., Full, W.: Fcm: The fuzzy c-means clustering algorithm. *Computers & geosciences* 10(2-3), 191-203 (1984)
- [9] Thomas, R. M., John, J.: Detection and segmentation of mitotic cell nuclei in breast histopathology images. In: Proceedings of the 2017 International Conference on Networks & Advances in Computational Technologies (NetACT). pp. 246-250. IEEE (2017)
- [10] Gudhe, N. R., Kosma, V. M., Behravan, H., Mannermaa, A.: Nuclei instance segmentation from histopathology images using bayesian dropout based deep learning. *BMC Medical Imaging* 23(1), 162 (2023)
- [11] Xing, F., Yang, L.: Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: a comprehensive review. *IEEE Reviews in Biomedical Engineering* 9, 234-263 (2016)
- [12] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III* 18. pp. 234-241. Springer (2015)
- [13] Raza, S. E. A., Cheung, L., Shaban, M., Graham, S., Epstein, D., Pelengaris, S., Khan, M., Rajpoot, N. M.: Micronet: A unified model for segmentation of various objects in microscopy images. *Medical image analysis* 52, 160-173 (2019)
- [14] Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., Liang, J.: Unet++: A nested u-net architecture for medical image segmentation. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings* 4. pp. 3-11. Springer (2018)
- [15] Ibtehaz, N., Rahman, M. S.: Multiresunet: Rethinking the u-net architecture for multimodal biomedical image segmentation. *Neural networks* 121, 74-87 (2020)
- [16] Jiang, Z., Ding, C., Liu, M., Tao, D.: Two-stage cascaded u-net: 1st place solution to brats challenge 2019 segmentation task. In: *Proceedings of the International MICCAI Brainlesion Workshop*. pp. 231-241 (2019)
- [17] Xing, F., Yang, L.: Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: a comprehensive review. *IEEE reviews in biomedical engineering* 9, 234-263 (2016)
- [18] Graham, S., Jahanifar, M., Azam, A., Nimir, M., Tsang, Y. W., Dodd, K., Hero, E., Sahota, H., Tank, A., Benes, K., et al.: Lizard: a large-scale dataset for colonic nuclear instance segmentation and classification. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. pp. 684-693 (2021)
- [19] Naylor, P., Láe, M., Reyal, F., Walter, T.: Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE transactions on medical imaging* 38(2), 448-459 (2018)

■著者



Chetan Gupta  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India

Rupesh Kumar  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India

Amit Shakya  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India



Shruti Phutke  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India

Lalit Sharma  
Emerging Technology and  
Innovation Lab,  
Yamaha Motor Solutions India

## ヤマハ発動機 技報企画委員会

委 員 荒木 治 ヤマハモーターエンジニアリング(株) コーポレートデザイン部  
 磯田 直也 モーターサイクル車両開発本部 MC 車両開発統括部 MC 開発戦略部  
 岡田 芳郎 アウトドアランドビークル事業本部 LSM 事業部 LSM 開発部  
 奥山 高志 マリン事業本部 開発統括部 技術管理部  
 菅崎 拓真 クリエイティブ本部 プロダクトデザイン部  
 神田 大 ソリューション事業本部 UMS 事業推進部 開発部  
 楠木 寿幸 ソリューション事業本部 ロボティクス事業部 技術統括部  
 近藤 俊介 技術・研究本部 AM 開発統括部 第1技術部  
 澤淵 敦志 パワートレイン開発本部 プロダクト開発統括部 第3PT 設計部  
 高木 秀明 ヤマハモーターハイドロリックシステム(株) 開発統括部 技術戦略部  
 土居 航介 生産技術本部 材料技術部  
 原 以起 マリン事業本部 開発統括部 先行開発部  
 増島 雄三 ランドモビリティ事業本部 SPV 事業部 プロダクト開発部  
 村松 恒行 技術・研究本部 共創・新ビジネス開発部

編集事務局 池田 厚司 グローバル RC・法務・知財本部 知財戦略部  
 田中 裕 グローバル RC・法務・知財本部 法務戦略部  
 三宅 英典 クリエイティブ本部 ブランドマーケティング部  
 菅崎 拓真 クリエイティブ本部 プロダクトデザイン部  
 加藤 隆輔 技術・研究本部 技術戦略部  
 片桐 健太郎 技術・研究本部 技術戦略部  
 片川 純 技術・研究本部 技術戦略部

## ヤマハ発動機 技報 第60号

YAMAHA MOTOR TECHNICAL REVIEW 2025 No.60

発 行 2025年12月1日

発 行 所 ヤマハ発動機 株式会社  
 〒438-8501 静岡県磐田市新貝2500  
 発 行 人 丸山 平二  
 デザイン レタープレス株式会社  
 印 刷 所 レタープレス株式会社  
 〒739-1752 広島県広島市安佐北区上深川町809-5  
 お問い合わせ 技報編集事務局 TEL 0538-37-1864  
 (技術・研究本部 技術戦略部)  
 ホームページ [https://global.yamaha-motor.com/jp/design\\_technology/technical/](https://global.yamaha-motor.com/jp/design_technology/technical/)

※無断転載を禁じます。

※本号の表紙は、創立70周年と技報 No.60の節目を記念し、ヤマハ発動機の歩みを振り返る写真を組み合わせたデザインとしました。そのため、表紙には許可を得たエリアやクローズドエリアでのプロライダーによる走行シーン、当時の社会背景や基準に基づく表現・写真、試作モデルやコンセプトモデルが含まれることがあります。現行の交通法規や製品の安全基準とは異なることがありますので、ご注意ください。交通法規を守り、安全運転・安全航行を心がけましょう。



ヤマハ発動機株式会社  
〒438-8501 静岡県磐田市新貝2500