Print ISSN 0285-9831
Online ISSN 1348-0316
CD-ROM ISSN 2186-716X
278

Visual Computing, Devices & Communications

画像電子学会誌

The Journal of the Institute of Image Electronics Engineers of Japan

【ビジュアルコンピューティング論文特集号】

◆【特集論文】Degrad2vec: 劣化埋め込みベクトルを用いた対話的画像復元
◆【特集論文】かぎ針編みにおける動作センシングと技能習得支援
◆【特集ショートペーパー】ストックの八重鑑別における日齢の影響検証
◆【特集ショートペーパー】小規模データセットにおける ニューラルネットワークと Hu モーメントを用いた小松菜の日齢推定
◆【特集ショートペーパー】Gaia 星表を用いた天体写真撮像系の内部標定
◆【一般論文】ハイパースペクトル画像を用いた ki-67 陽性核推定法の開発
◆【一般論文】改良型 StarGAN を用いたデータ拡張による近代書籍文字認識精度の向上

一般社団法人 画像電子学会 Vol.55 No.4 2026

画像電子学会賛助会員　（2026年9月30日現在）

池上通信機株式会社
科学技術振興機構
公益財団法人画像情報教育振興協会
株式会社ゲネシスコンマース
コニカミノルタ株式会社
ＺＡＺＡ株式会社
大日本印刷株式会社
日本電信電話株式会社
日本放送協会
株式会社日立製作所
富士フイルムビジネスイノベーション株式会社
古野電気株式会社
三菱電機株式会社
株式会社リコー

編集委員会

（委員長）　内田　理
（副委員長）　小林直樹，竹島由里子，石川雅浩．
（編集顧問）　安田靖彦，富永英義，小宮一三，小野文孝，羽鳥好律，松本充司，加藤茂夫，田中　清，児玉　明．
（編集理事）　石川雅浩，今給黎　隆，佐野　卓，董　然．
（編集幹事）　河村尚登，森谷友昭，山田雄一郎．
（編集・査読委員）　荒井良徳，池田充郎，井尻　敬，石川知一，五十嵐悠紀，上平員丈，大井　翔，大木眞琴，木村俊一，久下哲郎，倉掛正治，小池崇文，櫻井快勢，佐藤甲癸，佐藤周平，白川真一，新谷幹夫，関野雅則，田中賢一，田中　清，新田高庸，Ｎ．Ｐ．チャンドラシリ，プレーマチャンドラ・チンタカ，豊浦正広，長谷川まどか，濱本和彦，藤澤　誠，藤代一成，牧田孝嗣，向井智彦，吉田典正，吉川　浩，Chee Seng CHAN, Paramesran RAVEENDRAN, KokSheik WONG.

（査読委員）　荒川賢一，荒木昭一，有川智彦，伊藤貴之，五十嵐悠紀，岩切宗利，岩橋政宏，大澤秀史，尾上孝雄, 金井 崇, 金子俊一, 金森由博, 金子 格，金田和文, 北郷正輝，勝間ひでとし，　栗原恒弥，黒沢俊晴, 洪　博哲，小館亮之，駒形英樹，小町祐史, 今間俊博，斎藤隆文，齋藤　豪，斉藤文彦，佐藤真知子，篠原克幸，島村　潤，下馬場 朋禄，白井啓一郎，杉崎栄嗣，瀬崎　薫，瀬政孝義，高島洋一，高野邦彦，田中芳樹，高橋時市郎，谷口行信，田村　徹，辻　宏行，鉄谷信二，中村康弘，納富一宏，包　躍，林　正樹，福江潔也，堀田裕弘，本宮隆広，茅　暁陽，松木　眞，松田浩一，三田雄志，三ツ峰秀樹，六浦光一，森島繁生，柳原政弘，藪下浩子, 山口隆二, 山崎龍次, Hernan AGUIRRE, Yoong Choon CHANG，Robin Bing-Yu CHEN, Mochamad HARIADI，Pizzanu KANONGCHAIYOS, Teck Chaw LING，Keat Keong PHANG，Nordin BIN RAMLI.

（事務担当）　大嶽英宗，福島理恵子，本田京子．

入会のご案内

入会ご希望の方は下記ご参照の上，学会ホームページよりお申込頂くか，事務局にその旨ご連絡ください。

○会員の種別
正会員：本会の目的に賛同する個人
学生会員：本会の目的に賛同する学生
賛助会員：本学会を援助する個人または法人
特殊会員：本学会の目的に賛同する個人以外の図書室，研究室など

○入会金および年会費
入会金：正会員　1,000円　　学生会員 500円
年会費：正会員　10,000円
学生会員　3,000円
賛助会員　50,000円（1口）
特殊会員　12,500円

○ご連絡先

〒116-0002　東京都荒川区荒川3-35-4
ライオンズマンション三河島第二　101号
TEL (03)5615-2893　　FAX（03）5615-2894
E-mail: hensyu@iieej.org（編集）
kikaku@iieej.org（研究会・会員情報）
hyoujun@iieej.org（テストチャート）
http://www.iieej.org/
http://www.facebook.com/IIEEJ

画 像 電 子 学 会 誌

第55巻 第4号　通巻278号　(2026年10月)

目　　次

随 想

257　研究会の楽しみ方　坂東幸浩

ビジュアルコンピューティング論文特集号

258　ビジュアルコンピューティング論文特集号に寄せて　土橋宜典

論　文

259　Degrad2vec: 劣化埋め込みベクトルを用いた対話的画像復元　竹内英輔，青木義満

270　かぎ針編みにおける動作センシングと技能習得支援　木下 梓，五十嵐悠紀

ショートペーパー

279　ストックの八重鑑別における日齢の影響検証　堀越雄太，奥田萌莉，日浦慎作

284　小規模データセットにおける ニューラルネットワークと Huモーメントを用いた小松菜の日齢推定　奥田萌莉，日浦慎作

289　Gaia星表を用いた天体写真撮像系の内部標定　松尾陽矢, 山北彩未, 前原秀明

一般論文

292　ハイパースペクトル画像を用いたki-67陽性核推定法の開発　石川雅浩，吉田結真，中野夏澄，黒田真代，茅野秀一，小林直樹

300　改良型StarGANを用いたデータ拡張による近代書籍文字認識精度の向上　竹本有紀，安田敏悠，陳 延偉

報 告

309　第10回アレキサンダー・ベイン賞 授賞報告　ベイン賞選定委員会

311　2026 年度 第54 回 画像電子学会年次大会報告　藤澤 誠

会　告・ニュース

314　理事会だより
315　シクミネットへの登録のお願い
316　画像電子学会研究会等予定
317　画像電子学会 第319回研究会 In 沖縄 発表募集
319　ビジュアルコンピューティング論文募集
320　論文投稿の手引き
324　アレキサンダー･ベイン賞候補推薦募集
326　画像電子技術賞候補の推薦のお願い
329　フェロー候補推薦募集
331　代議員候補推薦のお願い
332　日本画像学会誌･日本写真学会誌･日本印刷学会誌目次
337　会　報
337　編集後記

The Journal of
the Institute of Image Electronics Engineers of Japan
Vo1.55 No.4 October 2026

CONTENTS

Foreword

257 Greetings from the Planning Committee Chair — Yukihiro BANDOH

Special Issue on Visual Computing

258 Introduction to the Special Issue on Visual Computing — Yoshinori DOBASHI

Contributed Papers

259 Degrad2vec: Interactive Image Restoration via Degradation Embedding Vectors — Eisuke TAKEUCHI , Yoshimitsu AOKI

270 Supporting Crochet Skill Learning through Motion Sensing — Azusa KINOSHITA, Yuki IGARASHI

Short Papers

279 Verification of the Effect of Seedling Age on Double-Flower Discrimination in Stock (Matthiola incana) — Yuta HORIKOSHI, Moeri OKUDA, Shinsaku HIURA

284 Age Prediction of Komatsuna Using Hu Moments with Neural Networks for Small Datasets — Moeri OKUDA, Shinsaku HIURA

289 Intrinsic Calibration of Astrophotography System Using Gaia Catalogue — Haruya MATSUO, Ayami YAMAKITA, Hideaki MAEHARA

Regular Section

Contributed Papers

292 Development of Ki-67 Positive Nuclei Estimation Method Using Hyperspectral Images — Masahiro ISHIKAWA, Yuma YOSHIDA, Kasumi NAKANO, Masayo KURODA, Hidekazu KAYANO, Naoki KOBAYASHI

300 Character Recognition for Early-Modern Japanese Printed Books with Data Augmentation by Improved StarGAN — Yuki TAKEMOTO, Toshihisa YASUDA, Yen-Wei CHEN

Reports

309 The Report of the Tenth Presentation of Alexander Bain Award — Alexander Bain Award Committee

311 The Report of IIEEJ Media Computing Conference 2026 — Makoto FUJISAWA

Published four times a year by the Institute of Image Electronics Engineers of Japan（Gazo Denshi Gakkai）
101, 3-35-4, Arakawa, Arakawa-ku, Tokyo 116-0002, Japan
Tel：+81-5615-2893 Fax：+81-5615-2894 E-mail：hensyu@iieej.org https://www.iieej.org/

IIEEJ 随想

研究会の楽しみ方



坂東　幸浩（企画委員長，下関市立大学）

Greetings from the Planning Committee Chair

Yukihiro BANDOH (Planning Committee Chair, Shimonoseki City University)

今年度より，研究会等の運営を担当する役目を務めることになった．そこで，この機会に，研究会のような学会発表の場にどのような意義があるのか，自分なりに整理してみたいと思った．早速，これまでの自分の学会発表を振返ってみると，まず思い出されたのは，発表そのものよりも，むしろ発表以外の出来事であった．少し余談になるが，本題に入る前に，思いつくままに書き連ねてみたい．

初めての学会発表は，研究室で借りたワゴン車で参加した．運転は全員で交代した．発表を控えた往路は，楽しい旅ではあったものの，どこか気の張る道中だったように思う．一方，発表を終えた帰路では，運転を終えて交代するたびに，ビール缶から「プッシュ」という景気の良い音が車内に響いた．発表も終え，運転の責務を果たした後の味は格別だった．じゃんけんに負けて最後の運転者になった友人には申し訳なかったが．

ポーランドのある街で会議に参加した時，ホテルの蛇口から出てきたのは茶色の水だった．飲用しなければ大丈夫だろうと高を括り，シャワーはもちろん，洗面・歯磨きなどにもその水を使っていたら，因果関係は不明だが，発熱した．近くに薬局などは無く，ホテルの近くにあった小さな店でウォッカを買い求めて，飲んで寝た．お陰様で，その後，三日ほどで解熱した．このことがあって，海外では水に注意しようと心に決めたが，幸い，その後は，蛇口から茶色の水が出る機会には遭遇せずに済んでいる．

初めての国際会議に参加したカナダで，忘れられない思い出がある．私と友人の各発表に対して，現地大学の教授が質問してくれた．発表後に挨拶を交わし，下手な英語でどうにかこうにか会話をしていると，一緒にランチに行こうと誘って下さり，素敵なレストランで私達二人に美味しいランチをご馳走して下さった．英語も下手で，発表内容も粗削りであった異国の学生に対してである．当時は「ホスピタリティ」という言葉すら知らなかったが，異国の地で触れた，温かな人情であった．

いささか寄り道をしたが，ここから本題について考えてみたい．研究会という場には，どのような意義があるのだろうか．まず，発表の機会としての意義について考えてみよう．一般的に，研究に関する発表内容は，発表者以外にとっては，難しい内容となる．特定の分野を深く考究した結果であるためである．したがって，研究発表では，こうした難しい内容を，限られた時間内で，分かり易く伝える力が必要となる．「一流の研究者は，難しい内容を分かり易く伝える」とは，かつての上司の言葉である．絶対的な正解は無いため，試行錯誤を繰返し，自分なりの伝え方を磨いていくことになる．こうした「伝える力」は，研究発表に限らず，社会の様々な場面で必要とされるスキルの一つであると感じている．そして，興味深いのは，取り組む姿勢と結果の関係である．十分でない準備の発表から得られるものは，それなりのものに留まる．しかし，渾身の準備を注いた発表には，その努力に相応しい反応が返ってくる．仮に思ったようにいかなかったとしても自身の気づきが深い．作用・反作用の法則のような原理が働いているように思われてならない．

次に，対話の機会としての意義について考えてみたい．私は，発表後の質疑応答を通じて，自分にはなかった視点を得て，視野が広がった経験が多い．見えていなかった部分に気づいたり，前提としていたことの妥当性を改めて考えたり，新たな発展の方向性を見出す糸口を得たりした経験である．こうした気づきを得るには，肩肘張らず，自然体で対話することが大切なのではないかと感じている．その意味でも，本会議中の質疑に加えて，コーヒーブレイクでの何気ない雑談や，懇親会での交流の持つ意義は大きいと思う．私の場合，コーヒーを片手にすると，会話は不思議と自然に弾むことが多いと感じている．

最後に，縁をつなぐ機会としての意義について考えてみたい．見ず知らずの人と新たに知り合うことは，一般には容易ではない．しかし，研究会などの場では，特別な工夫をしなくても，自然と新たな知己を得る機会がある．発表者であれば，質問をしてくれた人に対して，発表後にお礼を伝え，挨拶を交わせばそれでよい．発表者でなくとも，興味をもった発表があれば，質問をしてみればよい．そして，発表後に発表者のもとへ行き，「興味深い発表でした」と声をかけ，挨拶を交わせば，それだけで立派な面識となる．次に会った時には，自然と挨拶をかわせる関係になっているであろう．人との縁は，人生における貴重な財産だと思う．本会の研究会が，参加される皆様にとって，新たな縁をつなぎ，次の一歩につながる場となれば嬉しく思う．そうした場を少しでも育んでいけるよう，これから微力ながら努めていきたい．

ビジュアルコンピューティング論文特集号に寄せて

ゲスト編集委員長
ビジュアルコンピューティング研究会委員長　土橋　宜典

2026 年度のビジュアルコンピューティング（VC）論文特集号をお届けします．VC 特集号は，本学会 VC 研究会のメンバーが中心となり，コンピュータグラフィックスやコンピュータビジョン，画像処理，可視化領域を核とする視覚計算融合領域に広く関わる論文の募集・査読・編集・出版を行うもので，本号で 32 回目を迎えます．近年，生成 AI 技術の爆発的な進化や NeRF・3D Gaussian Splatting に代表される三次元復元技術の発展，さらには XR やメタバースの普及などにより，ビジュアルコンピューティングがカバーする領域は急速に拡大しています．本分野は，単なる画像処理の枠を超え，学際的かつ社会の基盤を支える重要技術としてその存在感をますます高めています．

今年度の VC 特集号への投稿件数は 12 件（内，英語論文 1 件，ショートペーパー5 件）でした．投稿締切の 3 月はちょうど学会シーズンでもあり，ご多忙の中投稿していただいた方々に感謝の意を表します．

厳正な審査の結果，最終的に 6 件の論文（内，英語論文 1 件，ショートペーパー3 件）が採択にいたりました．このうち本号には，2 件のフルペーパーと 3 件のショートペーパーを掲載しております．具体的には，画像復元，身体動作のセンシングと技能支援，植物の品種・生育状態の推定，天体写真撮像系の標定など，多様な対象を扱う研究が収録されています．機械学習，画像処理，センシング，コンピュータビジョンなどの技術を組み合わせ，それぞれの対象に応じた新たな画像・視覚情報処理の可能性を示す成果となっています．また，今回同時に採択となりました質感探索のための合成画像補間に関する英語論文の 1 件につきましては，英文誌 12 月号への掲載となります．本号の掲載論文と合わせて，ぜひご一読いただけますと幸いです．これらは会員の皆様にとっても大変示唆に富む内容であると確信しております．一方，今回残念ながら返戻となりました論文につきましても，優れた要素が少なからず含まれておりました．査読コメントを参考に研究を継続され，ぜひとも再投稿していただけますよう心よりお待ち申し上げております．

VC 論文特集号は，日本国内の学会誌としては数少ない，主にビジュアルコンピューティング技術にフォーカスした論文特集を毎年行っております．査読は国内の一流の研究者を選定し，しっかりと行っておりますので，掲載される論文は一定の品質が保証されたものとなっております．また，博士課程の学生が学位取得の対象となる論文を投稿しやすいスケジュールで出版時期を設定している点も大きな特徴です．さらに，掲載論文は本学会の論文賞だけでなく，「西田賞」の選定対象にもなります．西田賞は，本学会誌に掲載された過去 2 年の論文から，ビジュアルコンピューティング関係の論文のみを対象として隔年で選出される栄誉ある表彰であり，受賞の大きな機会となります．このように，本特集への掲載は多くの特典がございますので，次年度以降もぜひお気軽にご投稿いただけますと幸いです．

VC 研究会の活動としては，他にビジュアルコンピューティングワークショップがあります．これは 1 泊 2 日の合宿形式で議論が中心となる会議です．昨年は 12 月 7〜8 日に北海道札幌市の定山渓温泉で開催し，ビジュアルコンピューティングの未来について熱く議論が交わされました．今年は愛知県豊田市にて，12 月中旬での開催を予定しています．ぜひ対面で皆様とお会いし，昨年にも勝る議論を交わせることを楽しみにしております．

最後になりましたが，本論文特集号の出版にあたり，ゲスト編集幹事や編集理事・委員の先生方には，論文の査読に関する作業や校正・編集業務等で献身的にご尽力いただきました．また，学会事務局の皆様にも，論文の募集，編集，校正の過程で大変お世話になりました．この場をお借りして心から御礼申し上げます．

ゲスト編集委員長：土橋 宜典（北海道大学）
ゲスト編集幹事：古屋 貴彦（山梨大学），謝 浩然（北陸先端科学技術大学院大学），藤堂 英樹（拓殖大学）

IIEEJ 論文 -- ビジュアルコンピューティング特集 --

Degrad2vec: 劣化埋め込みベクトルを用いた対話的画像復元

竹 内 英 輔† 青 木 義 満†(正会員)

†慶應義塾大学

Degrad2vec: Interactive Image Restoration via Degradation Embedding Vectors

Eisuke TAKEUCHI† , Yoshimitsu AOKI†(*Member*)

†Keio University

〈あらまし〉 対話的画像復元は，外部から制御変数を入力することによって画像復元の強さを調整可能にすることを目指すものである．また劣化復元タスクはもとより，実世界に存在する複雑な劣化を仮定して学習することが難しいという課題を抱えている．我々は対話的画像復元と実世界の複雑な劣化への対応を同時に達成するために，劣化情報をベクトルに埋め込み，そのベクトルのノルムを操作することで対話的な画像復元を行う手法である Degrad2vec を提案する．モデルは劣化をベクトルに埋め込みつつ特徴量を抽出する前半部分と，そのベクトルを利用して復元を行う後半部分に分かれる．アテンションモジュールの働きが手伝い，エンドツーエンドの学習によって劣化ベクトルは自動的に獲得される．本来定量化が難しい劣化情報を多次元ベクトルに埋め込むことによって多様な劣化に対応できるようになるほか，そのベクトルを劣化の分析に用いることもできる．実験により本手法で獲得したベクトルが実際に劣化情報を高度に埋め込むこと，またそのノルムを操作することで復元強度が連続的に調整されることを実証する．

キーワード：ニューラルネットワーク, アテンション, 画像復元, 対話的画像復元, 埋め込み表現

<**Summary**> Interactive image restoration aims to enable adjustable restoration intensity through external control variables. However, a conventional challenge in this field is the difficulty of modeling and learning from the complex degradations encountered in real-world scenarios. To simultaneously address interactive restoration and real-world degradation complexity, we propose Degrad2vec, a novel framework that embeds degradation information into a latent vector and performs interactive restoration by manipulating its norm. The proposed architecture consists of two primary components: a front-end module for joint feature extraction and degradation embedding, followed by a back-end module that executes restoration guided by the degradation vectors. Facilitated by a vector-oriented attention mechanism, these embeddings are automatically acquired through end-to-end training. By mapping inherently unquantifiable degradation information into a multi-dimensional latent space, our method not only handles diverse degradations but also enables downstream degradation analysis. Experimental results demonstrate that the learned vectors effectively capture high-fidelity degradation characteristics and that manipulating their norms allows for continuous, fine-grained control over restoration strength.

Keywords: neural networks, attention mechanism, image restoration, interactive image restoration, embedding representation

1. はじめに

画像復元（Image Restoration, IR）は，劣化した画像から高品質な画像を再構築することを目的とした技術である．劣化とは，ノイズ，ぼかし（ブラー），圧縮によるアーティファクト，低解像度など，さまざまな要因による画質の劣化を含む．初期の手法では主に劣化プロセスを数理的にモデル化し，それを逆問題として解くことに焦点が当てられていた．近年では深層学習が着目され，CNN（畳み込みニューラルネットワーク）や Transformer を用いて自動的な学習によって逆関数を近似しようとする試みが主流である．

ところがほとんどの深層学習による復元手法は，データセットからパターンを学習するという性質上理想的な条件下を前提とし，単純な劣化モデルに基づいて設計されているものが

多い．実世界の劣化の大半はこのような仮定から外れた複雑な原因によるもので，その乖離が従来手法の性能に対する大きな制約となる．そのため近年では，RWSR（Real-World Super-Resolution，実世界超解像）を対象とした手法[1)2)3)4)]が提案されており，より複雑な劣化モデルへの対応や汎用性の向上を目指した研究が進んでいる．

また，深層学習ネットワークは本質的に関数であり，一対一の固定的なマッピングにより一つの入力に対して一種類の出力を出すものがほとんどである．ところが写真編集などの一部のアプリケーションやカメラ開発の領域では復元強度をユーザーが調整する柔軟性が要求される場合がある．例えばデノイジングとデブラーの二つの典型的な復元操作は，しばしばトレードオフの関係にある．互いに画像の高周波成分を抑制，強調するという側面があるためだ．このようなトレードオフを考慮し最適なバランスを新たなモデルの学習なしで探索するために，復元結果を動的かつ対話的に調整できる手法が必要となる．

しかし従来の対話的画像復元を実現する先行手法[5)6)7)]ではユーザーが指定，あるいはモデルが推測した各種劣化スコアのスカラー値を，そのままネットワークに入力する．劣化スコアとは，本来定性的な度合いである画像にかかる劣化強度を定量化し，基準としてモデルに与えるための数値である．先行手法や本手法では，$[0,1]$ の範囲で定義されるスカラー値となる．線形性といったような劣化の性質はブラーとノイズの二種間で異なるのはもちろんのこと，複雑な原因によって発生する実世界劣化ではブラーやノイズ自体の種類も多岐にわたるといえる．その場合劣化強度の連続的な推移も劣化の種類によって多様であると考えることができる．したがって多様な劣化の強度推移を単純なスカラー値として表現すると，その「単純化」によって失われる情報がモデルの精度に悪影響を与えると思われる．

本研究では劣化情報を連続的な値をとる二本の多次元ベクトルに埋め込み，そのベクトルのノルムを劣化スコアによって調整することで対話的な復元を実現することを目指す．二本のベクトルはそれぞれブラー劣化とノイズ劣化の情報を持つ（ブラーベクトル，ノイズベクトル）．多次元な劣化埋め込みベクトルを導入することによって，強さ調整の次元を拡張することができる．ユーザーによって指定された劣化スコアが同じであっても別種の劣化画像であれば劣化ベクトルの内容が異なり，調整の方向性が異なることになる．そのためラベルとして与えられる劣化スコアがある程度不正確なものであっても，学習によって獲得されるベクトルがその不安定さを補正する役割を持たせることが可能である．

提案するモデルとして入力画像の劣化特徴量を抽出しベクトルとするネットワークを含み，ノルム調整されたベクトルと入力画像を受け取って対話的に復元を行うネットワークである Degrad2vec を提案する．後半の復元部分に画像用 Self-Attention[8)]である Neighborhood Attention[9)10)]を採用した上で，空間エンコーディングによってベクトルと特徴量マップを Cross-Attention することに対応させる．

2. 関 連 研 究

2.1 CResMD[6)]

CResMD は，ニューラルネットワークを用いた対話的画像復元の先駆けとなる研究である．本フレームワークは画像復元を担う基盤ネットワークと，復元強度を制御する条件ネットワークの 2 枝構造で構成される．

核となる機構は，以下の式で定義される制御可能残差接続（Controllable Residual Connection）である：

$$Y = \alpha \cdot f(X, W_i) + X \tag{1}$$

ここで，X 及び Y は入力と出力，$\alpha \in [0,1]$ は条件ネットワークから出力される変調パラメータである．

条件ネットワークは，入力された劣化スコアを二つの $[0,1]$ スカラー値で構成されたベクトル $z \in \mathbf{R}^2$ として受け取り，各層の α を動的に生成する．その α をもとに基盤ネットワークが挙動を変化させることで，ユーザーの要求に応じた復元画像を生成する．

しかし，この研究の復元ネットワークでは内部的に二種の劣化は明示的に分離されていない．二種の劣化スコアから α を計算するネットワークが分離した復元を暗黙的に学習するのみである．またモデルのアーキテクチャも単純な畳み込み層の計算結果を線形にスケールさせたものであり，実世界の非線形な劣化に対する対応力は限定的である．我々の手法ではアーキテクチャ内で二種の復元それぞれを担当するモジュールが分かれており，各劣化復元の独立が明示的である．更に画像の劣化情報を劣化ベクトルという高次の特徴量にまとめること，それを受け取った復元モジュールが画像の領域ごとに重みづけした計算を行う Attention 機構を採用していることにより，複雑で非線形な劣化や，モーションブラーといった同じ画像内の領域ごとに劣化強度が変動するタイプの劣化にも対応しやすいアーキテクチャになっている．

2.2 MMRealSR[7)]

MMRealSR は，メトリック学習を用いて定量化困難な実世界の劣化強度を相対的な劣化スコアへとマッピングし，対話的な画像復元を実現する．実世界の劣化はランダム順序劣化[4)]や高次の劣化プロセス[3)]によってシミュレーション可能だが，劣化の強さを定量的に計算することが難しい．対話可能なモデルは学習時に劣化スコアをラベルとして与える必要があるため，そのことが RWSR と対話的な復元を両立することの障壁となる．この研究ではメトリック学習によって劣化スコアマッピングをモデルで獲得することでこの問題を解決するとした．

条件ネットワークは，汎用ぼかし（Gblur）と汎用雑音（Gnoise）の 2 系統の予測ブランチから構成される．各ブラン

チは入力画像から劣化の相対強度を反映したスコア $s \in [0,1]$ を推定する．

劣化スコアの妥当性を保証するため，以下のスキームが導入されている：

- **対比学習**：順序付けられた対比グループ（g_c）を用い，サンプル間の距離に基づきメトリックスペース内でのランク付けを学習する．
- **アンカーポイント戦略**：最大劣化（a_1）と無劣化（a_2）のサンプルからなるアンカーグループ（g_a）により，スコア分布を正規化する．

基盤ネットワークは，低解像度画像と条件ネットワークから得られる条件ベクトルを入力として受け取る．ベクトルは二種の s からなる．ユーザーは別途このベクトルを調整することで，リアルタイムかつ連続的に復元強度を制御することも可能である．

しかし劣化スコアをスカラー値として予測するモデルをトレーニングする方法では，多様な様態が想定される実世界の劣化に関する情報が，一種の劣化につき一次元という低次な量に圧縮されてしまうといえる．我々の手法では劣化情報を多次元なベクトルに埋め込むことで劣化次元を拡張し，劣化に関わる様々な特徴を含んだ量を条件付けとして対話的画像復元を行うことができる．

3. 手　　法

本手法で提案するネットワークは U-net[11]的構造をなす．U-net 的構造とは，受け取った入力が特徴量を取り出すネットワークを通過していくごとにその中間特徴量の解像度が減少させていく前半部分（エンコーダー）と，縮小した中間特徴量を段階的に元の大きさに拡大していき出力を推論する後半部分（デコーダー）からなるものである．詳細な局所的特徴と，大まかな大域的特徴の両方を高精度で処理するために，このような構造をとる画像処理ネットワークは多い．本手法では，エンコーダーの最後のブロックに主に畳み込み層で構成された小さいヘッドを二つ接続し，それぞれがノイズベクトル，ブラーベクトルを推論する．デコーダーはそれらベクトルと特徴量マップを受け取り，出力画像を推論する．劣化画像からのマッピングをネットワークが学習する過程で，エンコーダーからのベクトルは自然と劣化を復元するために必要な情報を持つようになる．エンコーダーとデコーダーの間でベクトルはユーザーの入力した劣化スコアによってノルムを調整され，それによって復元強度が変調する．

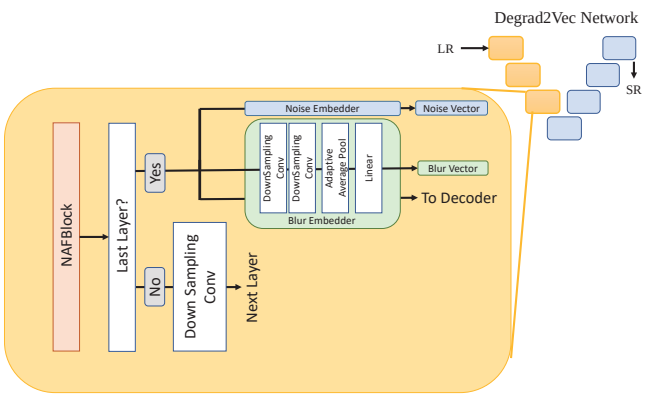


図 1　提案手法 Degrad2vec エンコーダーの図解
Fig. 1　Figure of Degrad2vec encoder

3.1　エンコーダー

エンコーダーの基本構造を**図 1** に示す．まず，入力画像 $I_{input} \in \mathbf{R}^{3\times H\times W}$ が，3×3 の畳み込み層によって初期特徴量に変換される．その後ネットワークは先行研究である NAFNet[12]の基本ブロック（NAFBlock）と，ストライド 2 の 2×2 畳み込み層からなるダウンサンプリング層を交互に配置した階層構造を経る．空間解像度を段階的に縮小しながら，チャンネル数を倍増させることで（本実装では $C_n \in \{64, 128, 256\}$），局所的及び大局的なコンテキストを抽出する．抽出された各階層の中間特徴量は，後段のデコーダーにおけるスキップ接続のために保持される．最深部におけるボトルネック特徴量を $F_{mid} \in \mathbf{R}^{C_N\times \frac{H}{4}\times \frac{W}{4}}$ とする．この F_{mid} に最後の NAFBlock を適用した後，ネットワークはノイズとブラーという 2 つの異なる劣化要因を独立してエンコードするために，分岐した 2 つの埋め込みモジュールへと特徴量を入力する．ノイズの表現ベクトル $v_n \in \mathbf{R}^{D_{emb}}$ を抽出する Noise Embedder は，空間情報を圧縮しつつ特徴の次元を目標の埋め込み次元 D_{emb}（本設定では 512）へとマッピングする．具体的には，ストライド 2 の 2×2 畳み込み層を 2 回連続して適用し，チャンネル数を D_{emb} と変換する．その後，Global Average Pooling (GAP) によって空間次元を 1×1 に集約し，テンソルを平坦化することでノイズベクトルを得る．この過程は次式で表される．

$$v_n = \mathrm{Squeeze}(\mathrm{GAP}(\mathrm{Conv}_{n2}(\mathrm{Conv}_{n1}(F_{\mathrm{mid}})))) \quad (2)$$

同様に，ブラーの表現ベクトル $v_b \in \mathbf{R}^{D_{emb}}$ を抽出する Blur Embedder も，同様のダウンサンプリングとプーリングの構造を採用している．

エンコーダーによって抽出されたノイズベクトル v_n 及びブラーベクトル v_b は，デコーダーへ渡される直前に，ユーザーが与えるプロンプト（劣化スコア）によって変調される．これにより，ユーザーは各劣化要因に対する修復の度合いを明示的に制御することが可能となる．

ユーザーからの入力プロンプトを $S = [S_n, S_b] \in \mathbf{R}^2$ としたとき，ノルムが調整されたベクトル $\tilde{v}_n, \tilde{v}_b$ は次のように計算される．

$$\tilde{v}_n = v_n \cdot (S_n + \epsilon) \quad (3)$$

$$\tilde{v}_b = v_b \cdot (S_b + \epsilon) \quad (4)$$

ここで，S_n 及び S_b はそれぞれノイズとブラーの修復強度

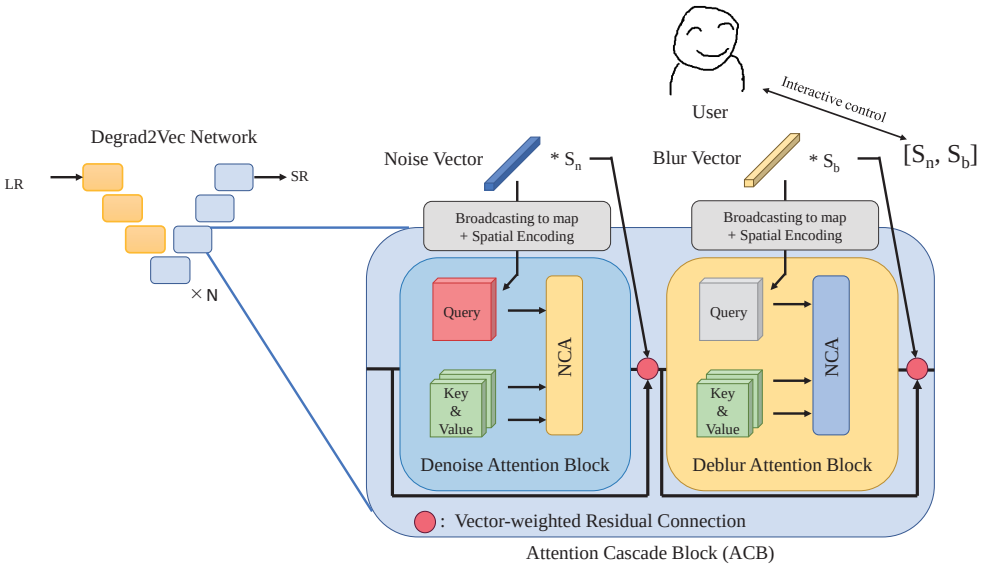


図 2 提案手法 Degrad2vec デコーダーの図解
Fig. 2 Figure of Degrad2vec decoder

を指定するスカラー値である．$\epsilon = 10^{-6}$ はベクトルが完全にゼロベクトルとなることを防ぎ，計算を安定させるための微小値である．このようにスケーリングされた $\tilde{v}_n$ と $\tilde{v}_b$ が，後段のデコーダーにおいて特徴量を適応的に変調するための補助として入力される．

3.2 デコーダー

デコーダーは，エンコーダーからスキップ接続を介して渡される階層的な空間特徴量と，ユーザーによって変調された劣化表現ベクトル $(\tilde{v}_n, \tilde{v}_b)$ を受け取り，段階的に解像度を拡大しながら最終的な出力を計算する．図解を**図 2** に示す．

ネットワークの各解像度レベルには Attention Cascade Block (ACB) が配置されている．初層（ボトルネック層）では，N 個 (本実装では $N = 10$) の ACB が積層され，各ブロックの Neighborhood Attention[9]における dilation（拡張率）は層ごとに 1 と 3 が交互に適用される．続く解像度レベル では dilation 1 及び 7 を持つブロック 2 つが，最終レベルでは dilation 9 を持つブロック 1 つがそれぞれ配置されている．このように解像度の復元過程において段階的に dilation を $1 \to 3 \to 7 \to 9$ と増加させることは受容野を効果的に拡大し，広範囲のコンテキストを考慮した復元の実現を意図している．

各解像度レベルでの処理後，特徴量は 1×1 畳み込み層と `PixelShuffle` によってアップサンプリングされ，エンコーダーからの同解像度の特徴量と加算（スキップ接続）される．最後に，入力画像と同一の解像度となった特徴量に対して 3×3 畳み込み層による調整を行い，最終的な出力層を通すことで 3 チャネルの結果画像を得る．

3.2.1 Attention Cascade Block (ACB)

ACB の内部では，ノイズ修復とブラー修復の処理が独立し，カスケード（直列）接続される．入力特徴量を X_{in} としたとき，まずノイズベクトル $\tilde{v}_n$ を条件とした処理が行われ，続いてその出力に対してブラーベクトル $\tilde{v}_b$ による処理が行われる．この処理フローは次式で定式化される．

$$\begin{aligned} F_n &= \mathrm{NCA}(X_{in}, \tilde{v}_n) \\ F'_n &= \mathrm{LN}(\mathrm{DWConv}(F_n)) \\ X_{mid} &= X_{in} + \mathrm{SCA}_v(F'_n, \tilde{v}_n) \\ F_b &= \mathrm{NCA}(X_{mid}, \tilde{v}_b) \\ F'_b &= \mathrm{LN}(\mathrm{DWConv}(F_b)) \\ X_{out} &= X_{mid} + \mathrm{SCA}_v(F'_b, \tilde{v}_b) \end{aligned} \tag{5}$$

ここで，NCA は劣化ベクトルをクエリとする Neighborhood Cross-Attention による変換，LN は Layer Normalization[13]，DWConv は 3×3 の Depthwise Convolution である．また SCA_v は，本研究で独自に導入したモジュールで，劣化ベクトルを入力として特徴量をスケーリングする Spatial Channel Attention である．SCA_v の演算は以下の式で表される．

$$\mathrm{SCA}_v(F, \tilde{v}) = \mathrm{Conv}(F \odot \mathrm{LeakyReLU}(\mathrm{Conv}(\tilde{v}))) \tag{6}$$

重要な機構として，SCA_v 内のベクトル空間からの投影層（Linear 層や Conv 層）でバイアス項を無効としている．そのためユーザーがプロンプトを通じて特定の劣化ベクトル（例：$\tilde{v}_n$）のノルムを 0 に設定した場合，SCA_v の出力はゼロ行列となる．結果，残差接続における加算が $X_{mid} = X_{in} + 0$ となり，該当の修復ブロックは恒等写像として振る舞う．これにより，ユーザーが修復を指示していない (=劣化スコアを 0 とした) 劣化を担当するブロックは，数理的に無効化されることとなる．SCA_v によるベクトルによって比率が変わる残差接続を，Vector-weighted Residual Connection(VRC) と呼称する．

この機構はネットワークがブラーとノイズを復元する機能を独立して学習するために重要となる．学習時に確率でブラーかノイズどちらかの劣化のみ含まれる画像が入力されると，含まれない方のベクトルがゼロベクトルとなる．したがってこのとき，ネットワークはその画像を片方の NCA ブロックのみで復元しなければならず，勾配も片方のブロックにのみ流れる．学習時にこのような片方の劣化のみの画像を復元するプロセスが繰り返し挟まれることによって，カスケード接続された NCA ブロックの分業が学習される．また，分業されたブロックが参照する劣化ベクトルもまた，片方の劣化のみの情報が含まれるような暗黙的な学習が進む．

3.2.2 Neighborhood Cross-Attention (NCA)

Neighborhood Attention は Query，Key，Value を同じ特徴量からマッピングして算出する Self-Attention である．しかし本手法では劣化ベクトルと画像特徴量を関連付けた演算が必要であるため，これを Cross-Attention に拡張する．

Neighborhood Cross-Attention は，劣化ベクトル $\tilde{v}$ と前段のモジュールからの画像特徴量 X の二種を受け取る．劣化ベクトル $\tilde{v}$ から動的に Query を生成し，画像特徴量から Key 及び Value を生成することによってベクトルを復元に関連付ける Cross-Attention 機構である．

ベクトルから 2D の特徴量マップである Query を生成するために，まず劣化ベクトル $\tilde{v}$ をバイアスなしの線形層で 4 つの部分ベクトル $[V_1, V_2, V_3, V_4]$ に分割・拡張する．これらに対して，2 次元の Positional Encoding（PE）を用いて空間的な位置情報を付与し，合算することで画像特徴量と同じ空間解像度を持つ Query 特徴量 Q を生成する．

$$\begin{aligned} Q &= \sum_{i=1}^{4} (\mathrm{PE}_i \odot V_i), \\ &\text{where } [V_1, V_2, V_3, V_4] = \mathrm{Linear}(\tilde{v}) \end{aligned} \tag{7}$$

ここで用いられる Positional Encoding の係数 (PE) は，画像の 4 つの領域（左上，右上，左下，右下）に対する滑らかな空間的勾配を提供する 4 チャンネルの重みマップである．画像の高さ及び幅方向の正規化された座標をそれぞれ $y \in [0,1]$ 及び $x \in [0,1]$ としたとき，各チャンネルのマップ $\mathrm{PE}_i(y,x)$ は三角関数を用いて次のように定式化される．

$$\begin{aligned} \mathrm{PE}_1(y,x) &= \frac{1}{2}\left(\sin\left(y\frac{\pi}{2}\right) + \cos\left(x\frac{\pi}{2}\right)\right) \\ \mathrm{PE}_2(y,x) &= \mathrm{PE}_1(y, 1-x) \\ \mathrm{PE}_3(y,x) &= \frac{1}{2}\left(\sin\left((1-y)\frac{\pi}{2}\right) + \cos\left(x\frac{\pi}{2}\right)\right) \\ \mathrm{PE}_4(y,x) &= \mathrm{PE}_3(y, 1-x) \end{aligned} \tag{8}$$

この設計により，PE_1 から PE_4 はそれぞれ画像の異なる方向に向かって値が減衰する，スケールされた係数として機能する．劣化ベクトルから生成された $V_1, \ldots, V_4$ は，この PE と要素ごとに乗算（$\odot$）される．結果として，単一のグローバルな劣化情報であるベクトルからでも，空間的に多様性のある適応的な Query 特徴量 Q を構築できる．

一方，画像特徴量 X からはバイアス付きの畳み込みによって Key (K) 及び Value (V) が抽出される．

$$\begin{aligned} K, V &= \mathrm{Conv}_{kv}(X) \\ \mathrm{NCA}(X, \tilde{v}_n) &= \text{Neighborhood Attention}(Q, K, V) \end{aligned} \tag{9}$$

Q, K, V を算出した後は，先行研究[9)10)]と同様に Neighborhood Attention を計算する．計算の詳細はこれらを参照されたい．

3.3 損失関数

本研究のネットワークは，生成画像 I_{pred} と正解画像 I_{gt} との間に，3 つの異なる評価指標を組み合わせた損失を用いて End-to-End で学習される．デブラーとデノイズを両方同時に行う本手法では，ピクセルベースの損失だけだとモデルがデノイズに対して偏って適応してしまう現象が確認できた．そのことを避けるため，ブラーに対して鋭敏に反応する傾向のある LPIPS[14)]，DISTS[15)]という二種の知覚的損失を導入する．

全体的な損失関数 $\mathcal{L}_{total}$ は，L_1 損失（実装上は最適化を安定させる Charbonnier 損失を使用），LPIPS 損失，及び DISTS 損失の重み付き和として次式で定義される．

$$\mathcal{L}_{total} = \mathcal{L}_{l1} + \lambda_1 \mathcal{L}_{lpips} + \lambda_2 \mathcal{L}_{dists} \tag{10}$$

本実装においては，経験的に $\lambda_1 = 0.1$, $\lambda_2 = 0.15$ と設定した．

4. 実　　験

4.1 データセット及びデータ処理

本研究では，画像復元アルゴリズムの学習及び評価に 4 つのデータセットを用いた．高解像度画像群からなる Flickr2K[16)]に加え，リアルなモーションブラーを含む RealBlur[17)]，RealBlur データセットの論文内で同著者が BSD500 データセット (BSD300[18)]の拡張版として後日公開されたもの) をベースに作成した BSD_B_Centroid[17)]，及び近遠距離の屋外シーンを網羅した HIDE[19)]を導入した．なお，各データセット間の規模の不均衡を是正するため，BSD_B_Centroid については，他のデータセットに対して 1/3 の確率で学習セットに選ばれるものとした．

入力画像サイズは固定 (256×256) とし，データセットから取り出された学習データは 70%の確率でリサイズ，30%の確率でランダムクロップ（サイズ不足の場合はリサイズ）によって調整した．

学習時の劣化付与プロセスは以下のとおりである．実ブラー劣化画像の場合は元のブラー強さを算出した後，60%の確率でノイズを付与した．クリーン画像の場合は，60%の確率で合成ブラーを付与した後，必ず合成ノイズを付与した．

4.1.1 合成劣化モデル

合成劣化はブラー，ノイズの順に付与され，それぞれ以下の手法から等確率で選択される．

表 1　定量的指標による精度比較

Table 1　Quantitative comparison

Method	**PSNR ↑**	**SSIM ↑**	**LPIPS ↓**	**DISTS ↓**	**NIQE ↓**
BSD300					
BSRGAN	26.67	0.696	0.392	0.240	5.938
RealCUGAN	27.87	0.759	0.301	0.276	6.308
RealESRGAN	25.64	0.696	0.345	0.210	6.130
RealSR	28.00	0.758	0.362	0.183	7.901
MMRealSR	25.89	0.715	0.312	0.196	6.010
MMRealSR(oracle)	28.56	0.775	0.307	0.178	6.971
Ours	**28.65**	**0.781**	**0.215**	**0.134**	**5.929**
DRealSR					
BSRGAN	31.81	0.876	0.226	0.201	7.934
RealCUGAN	31.65	0.881	0.240	0.246	9.004
RealESRGAN	31.18	0.866	0.226	0.189	7.985
RealSR	34.47	0.909	0.194	0.156	8.889
MMRealSR	32.56	0.893	0.199	0.177	8.545
MMRealSR(oracle)	**35.35**	**0.918**	0.166	0.144	7.692
Ours	34.69	0.913	**0.118**	**0.114**	**6.934**
RealSR					
BSRGAN	27.54	0.837	0.146	0.129	5.911
RealCUGAN	27.54	0.849	0.159	0.181	7.086
RealESRGAN	26.57	0.831	0.155	0.133	6.461
RealSR	28.34	0.848	0.203	0.163	6.771
MMRealSR	26.39	0.844	0.193	0.170	6.828
MMRealSR(oracle)	28.02	0.849	0.166	0.135	5.531
Ours	**30.99**	**0.903**	**0.134**	**0.115**	**5.355**

表 2　計算量・実行時間の比較

Table 2　Comparison of computational complexity and inference time

Model	**Params (M) ↓**	**FLOPs (G) ↓**	**Time (ms) ↓**
BSRGAN	16.7	9398.8	648.7
RealCUGAN	1.29	420.5	34.8
RealESRGAN	16.7	9398.8	646.9
RealSR	15.6	2221.6	630.0
MMRealSR	26.1	10038.5	688.0
MMRealSR(Restoration only)	23.5	9398.8	649.2
Ours	28.8	490.2	179.2

- **合成ノイズ:**　ガウシアンノイズ（分散 $\sigma^2 \sim \mathrm{Uniform}(0.3,1)^2$），ポアソンノイズ（スケール $s_p \sim \mathrm{Uniform}(0.02,0.15)^2$），ソルトアンドペッパーノイズ（確率 $P \sim \mathrm{Uniform}(0.01,0.1)$，確率で画像のピクセルが 255 か 0 の値をとるノイズ）．更に 70%の確率で，1 種目と異なる 2 種類目のノイズが追加される．

- **合成ブラー:** ガウシアンブラー，メディアンブラー（カーネルサイズ $\in \{3,5,7\}$），及び JPEG 圧縮ノイズ（品質 15 ∼ 30）．JPEG 圧縮は高周波成分を損なう性質があるため，本研究ではブラーの一種として扱う．

4.1.2　劣化スコアの算出

本手法における対話的画像復元の学習には，0 から 1 に正規化された劣化スコアのラベルが必要となる．それぞれの劣化スコアは，劣化前後の画像ペアの差分から以下のように計算される．

ノイズスコア　ノイズスコアは絶対平均誤差（MAE）を定数 $\lambda_n = 30$ で正規化して算出する．

$$S_{\mathrm{noise}} = \frac{1}{\lambda_n}\mathrm{MAE}(x,y) \tag{11}$$

ブラースコア　カーネルサイズ 5 のガウシアンブラーを用いて抽出した高周波成分 $\mathbf{H} = \mathbf{x} - \mathrm{GaussianBlur}(\mathbf{x},5)$ を用いる．劣化前後の高周波成分 $\mathbf{H}_x, \mathbf{H}_y$ の MAE を定数 $\lambda_b = 7$ で正規化して算出する．

$$S_{\mathrm{blur}} = \frac{1}{\lambda_b}\mathrm{MAE}(H_x(i), H_y(i)) \tag{12}$$

4.2　実験条件

学習時の主なハイパーパラメータは以下のとおり設定した．入力画像のパッチサイズは 256×256 ピクセル，バッチサイズは 2 とした．ネットワークの最適化アルゴリズムには

AdamW[20]を採用し，重み減衰（Weight Decay）を 0.001 に設定した．

学習率の制御に余弦アニーリング（Cosine Annealing）スケジューラ[21]を導入した．初期学習率を 1e-4 とし，全 100 エポックにかけて滑らかに減衰させ，学習終盤における最小学習率を 1e-6 となるように設定した．演算精度は 32 ビット浮動小数点数（FP32）とし，7 基の GTX TITAN X で学習を実行した．

4.3 定量的結果

本手法の実世界劣化に対する復元性能を検証するため，BSD300[18],RealSR[22],DRealSR[23]という 3 つのベンチマークセットを対象に BSRGAN[4]，MMRealSR[7]，RealCUGAN[24]，RealESRGAN[3]，RealSR[22]といった先行手法と本手法で各種指標を算出した．結果を**表 1** に示す．

ベンチマークデータセットに含まれる画像を全て 512×512 ピクセルにリサイズし入力した．本手法は，比較対象とする一部の既存手法と異なり，復元時に入力として劣化スコアを指定する必要がある．そのため，評価においては 1 枚の入力に対し，2 種の劣化スコアを 0 から 1.0 まで 0.1 刻みで変化させた計 121 パターンの出力を生成し，各指標において最良となった値を採用した．また，先行手法の MMRealSR は劣化スコアの値を 0 から 1 で推論するネットワークを持ち，復元ネットワークには推論された値を入力することも，任意の値を入力することもできる．そのため推論された値をそのまま用いたものを MMRealSR の結果とし，MMRealSR(oracle) という形で本手法の検証と同様にグリッドサーチによって最良の値を探索した場合の実験結果も記載した．

その結果，MMRealSR(oracle) の一部を除いて全ての指標及びデータセットにおいて本手法が既存手法を上回った．この結果は MMRealSR(oracle) との比較を除き，各指標に最適化された調整値を選択したという「オラクル（Oracle）的な上界」の結果であることに留意する必要がある．したがって，本結果は本手法のモデルが潜在的に (繰り返しパラメータ調整を目的に合わせ行えば) 既存手法と同等以上の復元能力を保持していることを示したほか，推論値をそのまま用いる標準運用の MMRealSR に対しても優位な復元性能を持っていることを示したといえる．

更に，本手法を用いて BSD300 データセットの 300 枚の画像に対し先ほどと同様の 121 通りのパラメータ出力（計 36,300 サンプル）を集計し，その画像の最良値（Oracle）との PSNR 値の差分を検証したところ，ΔPSNR の平均値は −0.42 dB，中央値は −0.27 dB であり，探索空間全体の 90.48% が最良値から −1.0 dB 以内に収まる結果となった．oracle でない MMRealSR と本手法の PSNR の差が，BSD300 データセットでは 2.76 dB であることを踏まえるとこの統計は，ユーザーが極めて厳密なパラメータ探索を行わずとも先行手法と比較して良好な復元結果を得られる可能性が高いことを示しており，対話的運用における調整の負担が小さいことを裏付けるものである．

また，本手法の主眼は復元精度の絶対値ではなく，出力の柔軟性であると考える．既存手法である MMRealSR とは異なり，本手法は劣化スコアの推定値を出力しないが，対話的な復元を前提とする本モデルにおいて，ユーザーによる調整操作を必要とする設計は妥当である．調整を必要としない復元を目的とするならばそもそも対話的モデルを採用する必要はない．他の調整可能でない手法と異なり，優れた復元結果を得るために繰り返しの調整を要求したとしても，本手法の特性を考慮すればこの制約は許容しうると考える．

4.4 復元画像についての定性的結果

復元画像の定性的な評価結果として，以下の 2 つの比較を示す．

図 3 は，RealSR データセット[22]の画像に対して各手法による復元を適用した結果である．提案手法（Ours）における入力劣化スコアは，4.2 節で述べた手法により 121 通りの組合せについてグリッドサーチを行い，LPIPS[14]が最良となったパラメータを採用している．先行手法と比較して，提案手法による復元結果は細部のテクスチャやエッジ情報を良好に保持していることが確認できる．一方で，他手法に比べて僅かにノイズが残存する傾向も見られるが，これは LPIPS の最小化に最適化した結果であって，細部の保持とノイズ低減のトレードオフは適用用途に応じて柔軟に制御可能である．

図 4 は，CelebA データセット[25]の画像に対して合成的な劣化を付加し，プロンプトとなる劣化スコアを段階的に変化させた際の復元結果を示したものである．同図 (a) は合成ノイズを付加した画像に対してノイズに関する入力劣化スコアのみを増加させた結果であり，(b) は合成ブラーを付加した画像に対してブラーに関する入力劣化スコアのみを増加させた結果である．入力する劣化スコアの増加に伴い，対象とする劣化に対する復元強度が連続的かつ漸進的に強まっていることが確認できる．

4.5 劣化ベクトルについての定性的結果

提案手法により獲得された劣化ベクトルが，実際に劣化に関連する情報を保持しているかを実証する．まず，学習及び評価で用いたデータセットから 44 枚の実ブラー画像と 100 枚の自然画像を抽出した．次に，自然画像に対して学習時に用いた各種ノイズ又は合成ガウシアンブラーを一定の割合で付加した．これらを提案手法のネットワークに入力し，ノルム調整前の劣化ベクトルを抽出して t-SNE[26]により 2 次元空間へ圧縮・可視化した．結果を**図 5** に示す．同図 (a) はブラーベクトル，(b) はノイズベクトルのプロット結果である．なお，ソルトアンドペッパーノイズは特有の集団を形成していたため，他のノイズと区別して描画している．

図 5(a) のブラーベクトルにおいて，ブラーの種類ごとに空間的なクラスターを形成していることが確認できる．これ

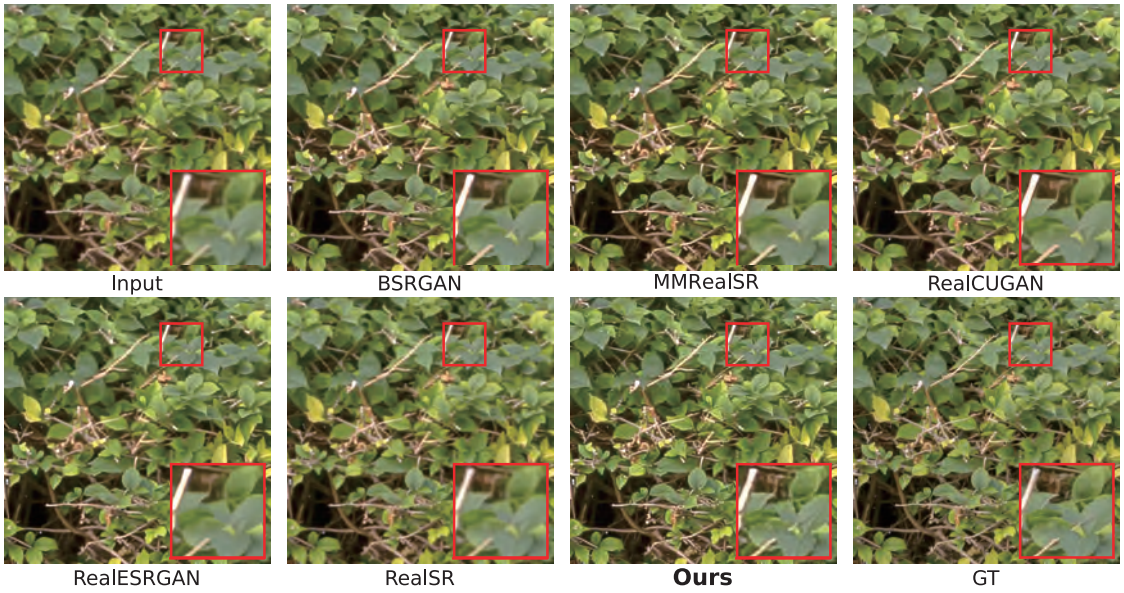
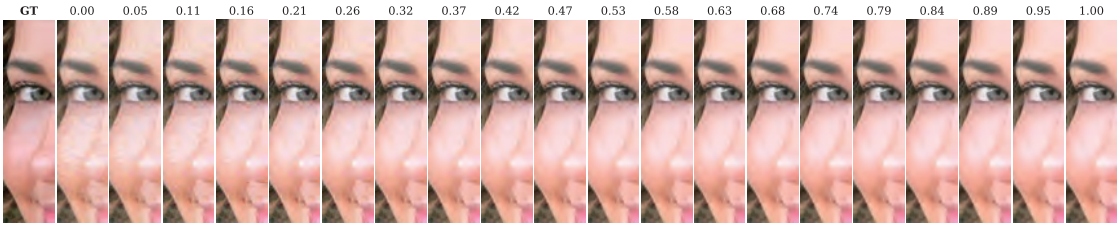


図 3　復元性能の定性的比較
Fig. 3　Qualitative comparison of restoration performance



(a) Transition of Denoising (Control Vector: 0.0 to 1.0)



(b) Transition of Deblurring (Control Vector: 0.0 to 1.0)

図 4　制御性能の定性的比較
Fig. 4　Qualitative comparison of controllability

は，ブラーベクトルが当該のブラー劣化に関わる情報を保持していることを示している．一方で，ノイズのみが付加された画像とクリーン画像は同一領域に混在しており分離されていないことから，ブラーベクトルがノイズに関する情報を含んでいないことが示唆される．これは劣化情報の分離という本手法の設計意図に合致する結果である．

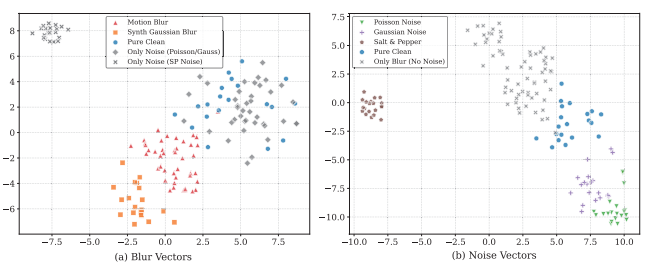


図 5　合成劣化に対する劣化ベクトルの t-SNE 可視化
Fig. 5　t-SNE visualization of degradation vectors under various synthetic degradations

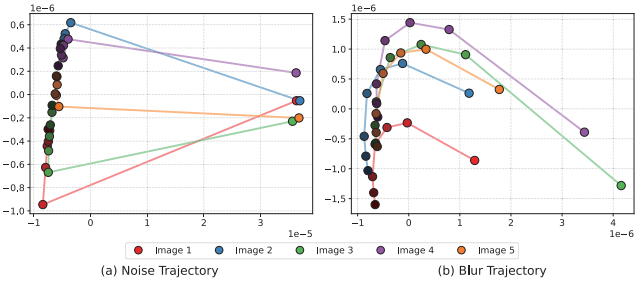


図 6　劣化強度の変化に伴う劣化ベクトルの軌跡
Fig. 6　Trajectories of degradation vectors with varying degradation intensities

図 5(b) のノイズベクトルについても，ノイズの種類ごとにクラスターを形成しており，ノイズベクトルが対象の劣化情報を獲得できていることが示される．ただし，ブラーのみが付加された画像とクリーン画像の間でベクトルが分離する傾向が見られ，ノイズベクトルがブラーに関する情報をある程度保持してしまっていることが示唆される．なお，ソルトアンドペッパーノイズは (a), (b) のいずれにおいても固有のクラスターを形成しており，当該ノイズの復元には他とは異なる特異な特徴表現が必要とされている可能性が考えられる．

続いて，劣化強度の変化に対する劣化ベクトルの挙動を検証する．CelebA データセット[25]から抽出した 5 枚の画像に対し，加法的ノイズ及びブラーの強度を段階的に変化させ，各条件下における劣化ベクトルを抽出した．これらのベクトルを PCA により低次元空間へ投影し，劣化強度の推移に伴う軌跡を可視化した結果を**図 6** に示す．同図 (a) は合成ノイズ付加時のノイズベクトルの軌跡，(b) は合成ブラー付加時のブラーベクトルの軌跡であり，各プロットは暗色かつ低彩度になるほど劣化強度が大きいことを表している．

いずれの劣化においても，劣化強度に相関した軌跡が確認できる．また，劣化が強くなるにつれて各画像の劣化ベクトルが一定の領域へ収束する傾向が見られる．これは，劣化ベクトルが画像の内容（被写体やテクスチャ）に関わる不要な情報を持たず，劣化の特性に関する情報のみを純粋に保持していることを示唆している．

一方で，劣化のないクリーン画像は劣化した画像群から空間的に離れており，画像間でのばらつきも大きい．本手法の学習時には，確率的に劣化を含まない画像が入力された場合，対応する劣化スコアとして 0 が与えられる．この際，3.2.1 節で述べたように当該劣化ベクトルに対する損失の勾配は 0 となるため，クリーン画像に対する劣化ベクトルは学習が行われず，初期値に近いランダムな状態に留まる．得られたプロットの分散は，この学習挙動を裏付ける結果であると考えられる．

4.6　計算量・実行時間の比較

4.3 節と同条件で，本手法と全ての比較手法のモデルに 512 × 512 の画像一枚に相当するダミーデータを与え，その際にかかる計算量と実行時間を計測し，**表 2** に示す．MM-RealSR[7]は劣化スコア推論ネットワークを使用せず，任意の劣化スコアを入力して運用する場合も想定し，後段の復元ネットワークのみを計算に用いた場合の各種計測結果も記載する．

本手法はパラメータ数こそ最大であるものの，計算量や実行時間が RealCUGAN[24]を除いた比較手法に対して最も良好な結果を示した．特に最も精度で拮抗し，同じく対話的画像復元手法である MMRealSR と比較して FLOPs は約 1/20，計算時間は約 1/4 となっている．本手法が Attention 構造を中心としたアーキテクチャであり，(局所的 Attention である Neighborhood Attention[9]と比較し) 計算量の大きい CNN の採用割合を抑えていること，ベクトルの推論と特徴量の圧縮を同時に行い全体として一つの U-Net 構造にまとめたこと，U-Net 構造の段数が多く計算のほとんどが元画像から 1/4 まで縮小したサイズの特徴量に対して行われていること，採用した Neighborhood Attention の計算効率が高く，カーネルが専用の実装によって最適化されていることなどが要因として考えられる．なお，FLOPs の削減幅（約 1/20）に対して実行時間の削減幅（約 1/4）が小さい点や，同等 FLOPs の RealCUGAN より計算時間を要している点については，Attention 機構特有のメモリ帯域ネックやカーネル起動オーバーヘッドによるものと考えられる．

対話的画像復元は，ユーザーが復元結果を確認しながら逐次パラメータを調整・再推論する運用形態を想定している．実際の対話における試行回数 N はケースバイケースで不定であるが，処理系全体のコストは単一推論コストの N 倍として線形に累積する．したがって，1 回あたりの推論時間を約 1/4 に短縮した本手法は，実運用におけるユーザーの総待機時間や反復操作に伴うストレスを大幅に軽減できる．参考として，4.3 節のオラクル評価で用いた 121 パターンの探索に要した総時間を比較すると，MMRealSR の 98.76 秒に対し本手法は 28.68 秒と約 3.4 倍の高速化を確認できた．この結果からも，試行回数がかさむ探索的な運用において本手法が高い実用性を有していることが示されたといえる．しかし実行時間が単純に一枚当たりの実行時間の 121 倍より大きいことが示すように，この結果には検証プロセスにおける定量的指標を計算する時間なども秒数に含まれており，この比は保守的な見積りである．また，単一推論の計算コスト削減に本手法独自のアーキテクチャ設計がどの程度寄与しているかは不明であり，大部分が Neighborhood Attention 自体の計算効率によるものである可能性に留意する必要がある．

5. む　す　び

本稿では，実世界の複雑な劣化のかかった画像を対話的に復元するため，2 種の劣化の情報を含んだベクトルを埋め込み，それを補助入力として調整可能な復元を実現する Degrad2vec ネットワークを提案した．学習時，一定の確率でブラーかノイズ片方しか含まれない画像が入力されることにより，VRC によってデノイズとデブラーを機能的に分離させる学習が促進される．その有効性が定性的結果より実証された．獲得されたノイズベクトル，ブラーベクトルはそれぞれノイズ，ブラーに関わる情報を持ち，逆に画像の内容に関する情報は比較的含まれていないことを示唆する結果が得られた．定量評価においても，既存の RWSR（Real-World Super-Resolution）特化型手法に対して競争性を持つ復元性能が，潜在的には存在することが確認できた．また，Neighborhood Cross-Attention 及び VRC の設計により，入力ベクトルのノルム，すなわちユーザープロンプトに応じた復元強度の連続的な制御を実現した．更に類似の先行手法に比べ，計算量や実行時間を大きく

削減することにも成功した．しかし一方で，ソルトアンドペッパーノイズがブラーベクトルにも深く影響を与えていることが確認されたほか，ノイズベクトルにブラー画像とクリーン画像を区別できてしまう情報が含まれている可能性が示されるなど，両ベクトルの分担が完全でないことは今後の課題である．また本手法で獲得されたベクトル埋め込み器は，様々な劣化を効果的に符号化するため事前学習済みネットワークとして他のタスクに転用できる可能性がある．その可能性の検証も今後検討する予定である．

参 考 文 献

1) S. Bell-Kligler, A. Shocher, M. Irani: "Blind Super-Resolution Kernel Estimation Using an Internal-GAN", Proc. of Advances in Neural Information Processing Systems (NeurIPS 2019), pp. 284–293 (2019).
2) S. Maeda: "Unpaired Image Super-Resolution Using Pseudo-Supervision", Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2020), pp. 288–297 (2020).
3) X.-T. Wang, L.-B. Xie, C. Dong, Y. Shan: "Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data", Proc. of IEEE/CVF International Conference on Computer Vision Workshops (ICCVW 2021), pp. 1905–1914 (2021).
4) K. Zhang, J. Liang, L. Van Gool, R. Timofte: "Designing a Practical Degradation Model for Deep Blind Image Super-Resolution", Proc. of IEEE/CVF International Conference on Computer Vision (ICCV 2021), pp. 4791–4800 (2021).
5) J. He, C. Dong, Y. Qiao: "Modulating Image Restoration with Continual Levels via Adaptive Feature Modification Layers", Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2019), pp. 11056–11064 (2019).
6) J.-W. He, C. Dong, Y. Qiao: "Interactive Multi-Dimension Modulation with Dynamic Controllable Residual Learning for Image Restoration", arXiv preprint arXiv:1912.05293 (2020).
7) C. Mou, Y.-Z. Wu, X.-T. Wang, C. Dong, J. Zhang, Y. Shan: "MM-RealSR: Metric Learning Based Interactive Modulation for Real-World Super-Resolution", arXiv preprint arXiv:2205.05065 (2022).
8) A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N Gomez, L. Kaiser, I. Polosukhin: "Attention Is All You Need", Proc. of Advances in Neural Information Processing Systems (NeurIPS 2017), pp. 5998–6008 (2017).
9) A. Hassani, S. Walton, J. Li, S. Li, H. Shi: "Neighborhood Attention Transformer", Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2023), pp. 6185–6194 (2023).
10) H.-Z. Liu, B.-H. Li, C.-K. Liu, M. Lu: "DeblurDiNAT: A Generalizable Transformer for Perceptual Image Deblurring", arXiv preprint arXiv:2403.13163 (2024).
11) O. Ronneberger, P. Fischer, T. Brox: "U-Net: Convolutional Networks for Biomedical Image Segmentation", Proc. of International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015), pp. 234–241 (2015).
12) L. Chen, X. Chu, X. Zhang, J. Sun: "Simple Baselines for Image Restoration", Proc. of European Conference on Computer Vision (ECCV 2022), pp. 17–33 (2022).
13) J. L. Ba, J. R. Kiros, G. E. Hinton: "Layer Normalization", arXiv preprint arXiv:1607.06450 (2016).
14) R. Zhang, P. Isola, A. A. Efros, E. Shechtman, O. Wang: "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric", Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2018), pp. 586–595 (2018).
15) K. Ding, K. Ma, S. Wang, E. P. Simoncelli: "Image Quality Assessment: Unifying Structure and Texture Similarity", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 44, No. 5, pp. 2567–2581 (2022).
16) R. Timofte, E. Agustsson, L. Van Gool, M.-H. Yang, L. Zhang, et al.: "NTIRE 2017 Challenge on Single Image Super-Resolution: Methods and Results", Proc. of IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW 2017), pp. 1110–1121 (2017).
17) J. Rim, H. Lee, J. Won, S. Cho: "Real-World Blur Dataset for Learning and Benchmarking Deblurring Algorithms", Proc. of European Conference on Computer Vision (ECCV 2020), pp. 184–201 (2020).
18) D. Martin, C. Fowlkes, D. Tal, J. Malik: "A Database of Human Segmented Natural Images and Its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics", Proc. of IEEE International Conference on Computer Vision (ICCV 2001), Vol. 2, pp. 416–423 (2001).
19) Z. Shen, W. Wang, J. Shen, H. Ling, T. Xu, L. Shao: "Human-Aware Motion Deblurring", Proc. of IEEE/CVF International Conference on Computer Vision (ICCV 2019), pp. 5572–5581 (2019).
20) I. Loshchilov, F. Hutter: "Decoupled Weight Decay Regularization", Proc. of International Conference on Learning Representations (ICLR 2019) (2019).
21) I. Loshchilov, F. Hutter: "SGDR: Stochastic Gradient Descent with Warm Restarts", Proc. of International Conference on Learning Representations (ICLR 2017) (2017).
22) J. Cai, H. Zeng, H. Yong, Z. Cao, L. Zhang: "Toward Real-World Single Image Super-Resolution: A New Benchmark and a New Model", Proc. of IEEE/CVF International Conference on Computer Vision (ICCV 2019), pp. 3086–3095 (2019).
23) P.-X. Wei, Z.-W. Xie, H.-N. Lu, Z.-Y. Zhan, Q.-X. Ye, W.-M. Zuo, L. Lin: "Component Divide-and-Conquer for Real-World Image Super-Resolution", Proc. of European Conference on Computer Vision (ECCV 2020), pp. 101–117 (2020).
24) bilibili, Real-CUGAN: Real Cascade U-Nets for Anime Image Super Resolution, https://github.com/bilibili/ailab/tree/main/Real-CUGAN (2021).
25) Z. Liu, P. Luo, X. Wang, X. Tang: "Deep Learning Face Attributes in the Wild", Proc. of IEEE International Conference on Computer Vision (ICCV 2015), pp. 3730–3738 (2015).
26) L. van der Maaten, G. Hinton: "Visualizing Data Using t-SNE", Journal of Machine Learning Research, Vol. 9, No. 86, pp. 2579–2605 (2008).

(2026 年 6 月 22 日 受付)
(2026 年 8 月 24 日 再受付)



竹 内 英 輔
2025 年 慶応義塾大学理工学研究科修士課程修了．現在慶応義塾大学理工学研究科博士課程に在籍．機械学習に関する研究に従事．



青 木 義 満 （正会員）
1996 年 早稲田大学理工学部応用物理学科卒業．2001 年 同大大学院理工学研究科物理学及応用物理学専攻博士課程修了．博士 (工学)．2002 年芝浦工業大学工学部情報工学科講師，准教授を経て，2008 年より慶應義塾大学理工学部電子工学科准教授，2017 年より教授．画像センシング技術に関する研究に従事．電子情報通信学会，画像電子学会，映像情報メディア学会，IEEE 会員．日本顔学会理事．画像センシング技術研究会組織委員．

IIEEJ 論文 -- ビジュアルコンピューティング特集--

かぎ針編みにおける動作センシングと技能習得支援

木下 梓† 五十嵐 悠紀†† (正会員)

†お茶の水女子大学，††東京大学

Supporting Crochet Skill Learning through Motion Sensing

Azusa KINOSHITA†, Yuki IGARASHI†† (*Member*)

†Ochanomizu University，††The University of Tokyo

〈あらまし〉 編み物は，動画や書籍など多様な学習資源が存在する一方で，初心者がそれらを参照しながらお手本通りに制作することは依然として難しい．特に，編み動作の質を客観的に評価し，技能習得を支援する仕組みは十分に確立されていない．本論文では，かぎ針編みを対象として，かぎ針にセンサを取り付けて編み動作を計測・解析するためのシステムを提案する．提案システムを用いて初心者及び経験者の動作データを記録する実験を行った．さらに，センサ値に基づいて閾値を設定することで，鎖編みを一目編むごとにリアルタイムで音によるフィードバックを提示する機能を実装した．実験及び分析の結果に基づき，適切なセンサ軸及び閾値設定について検討を行い，編み動作に対するリアルタイムフィードバックの有効性を示した．本研究は，手工芸における技能の定量化とフィードバックを通じて，初心者の効率的な技能習得を支援する新たなアプローチを提供する．

キーワード：手工芸, かぎ針編み, 動作センシング, リアルタイムフィードバック, 技能習得支援

<**Summary**> Knitting has become increasingly popular in recent years, supported by abundant learning resources such as online videos and books. However, beginners often struggle to reproduce patterns accurately, even when following these materials. In particular, methods for objectively evaluating knitting motions and providing feedback for skill acquisition remain underexplored. In this paper, we propose a sensing system for crochet that captures and analyzes hand movements by attaching sensors to a crochet hook. Using the proposed system, we collected motion data from both beginners and experienced users and analyzed the differences in their knitting behaviors. Furthermore, we implemented a real-time auditory feedback mechanism that generates a sound for each completed chain stitch based on predefined thresholds derived from sensor data. Based on experimental results and data analysis, we examine appropriate sensor axes and threshold settings for effective feedback. Our findings demonstrate the feasibility of real-time feedback based on motion sensing and provide insights into designing interactive systems for supporting skill acquisition in handcraft activities. This work contributes a novel approach to quantifying and augmenting knitting skills, with the aim of improving learning efficiency and output quality for beginners.

Keywords: handicrafts, crochet, motion sensing, real-time feedback, skill learning support

1. はじめに

編み物は，手作業による創作活動の一つとして親しまれてきたが，近年若者を中心に再び注目を集めている．コロナ禍での外出自粛により自宅でできるという点で人気を集め，普及する動画 SNS で個人の作品や制作過程を気軽に目にするようになったことや，K-POP アイドルや水泳選手が編み物を趣味として発信し話題となったことで関心の高まりを後押ししたと報告されている[1)]．編み物は，毛糸や編み針などの材料・道具が安価かつ容易に入手可能であり，初心者でも手軽に始めやすい．また，制作方法に関する動画や書籍も多数あり，帽子やマフラー・バッグなどの私たちが身に付けられるものから，あみぐるみなどの作品，人形の着せ替えなどさまざまなものを制作可能である．さらに，多様な太さの編み針があり，毛糸は太さだけでなく素材や色の種類も豊富にあるため，自分好みにアレンジするなど，高い表現の自由度を有する．編み物を含む Creating Arts and Crafting(CAC) 分野はウェルビーイングへの効

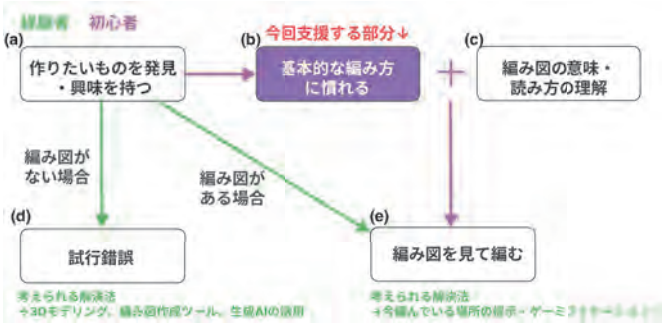


図 1　作品を編むまでの流れ
Fig. 1　Flow of the knitting process

果がある[2]と示唆されている．加えて，インターネット上では minne[3]をはじめとするハンドメイドマーケットやフリマサイトを通じて作品の売買も行われており，創作活動は自己表現の手段としても機能している．

一方で，作品をお手本通りに制作するためには，書籍や画像を参考にする場合，基本的な編み方の習得に加えて，図 1 に示すような設計図である編み図の読み方を知る必要がある．動画を参照する場合は編み図の理解は必須ではないものの，編み物は繰り返しの動作が多いため「特定の編み方で何目編む」(「目」は繰り返される編み目の単位) といった表現で説明されることが一般的であり，少なくとも基本的な編み方の習得が前提となる．しかしながら，初心者にとっては編み目を安定して整えながら継続的に編み進めることが難しく，技能習得の過程においてつまずきやすいという課題がある．

本論文では，かぎ針編みを対象とし，かぎ針にセンサを取り付けて編み動作を記録し解析するためのシステムを提案する．提案システムを用いて，初心者及び経験者の編み動作データを収集した．さらに，センサ値に基づいて閾値を設定することで，編み目を一目編むごとにリアルタイムで音によるフィードバックを提示する機能を実装した．閾値として用いるセンサ軸及び基準値は，記録したセンサの値を読み込んでグラフを作成し，目視で検討した．経験者にはシステムを用いて閾値を決定してもらい，その閾値を利用した実験を行い，これを元に閾値として用いるセンサの軸の方向を検討した．

本研究の貢献は以下の 3 点である．(1) かぎ針にセンサを装着することで編み動作を計測し，編み動作の進行状況及び熟練度を反映する特徴を定量的に捉える手法を提案した点，(2) 編み動作に基づくリアルタイム音フィードバックを実現し，技能習得を支援するインタラクティブなシステムを構築した点，(3) 初心者と経験者の動作データの比較を通じて，有効な特徴量及びフィードバック設計に関する知見を示した点である．これにより，本研究は編み物技能の習得支援に新たなアプローチを提示し，作品の品質向上及び効率的な上達に寄与することを目指す．

2. 関 連 研 究

本章では，編み物に関する関連研究を，(1) 設計・制作支援，(2) シミュレーション，(3) 動作解析，(4) フィードバックの観点から整理する．

編み物の設計・制作支援に関して，Igarashi らは，既存の 3D モデルから毛糸を用いて作成するぬいぐるみである「あみぐるみ」をかぎ針編みで作成するための編み図を自動生成する手法[4]を提案した．また，あみぐるみの概形をスケッチ入力でデザインすると，編み図と完成予想の 3 次元モデルを提示するシステム Knitty[5]を提案した．Wu ら[6]は，取手のあるポットなどの複雑な 3D モデルから，棒針編み及び機械編みで制作可能な編み目メッシュを生成する手法を提案した．これらの研究は，デザインの支援や編み図の生成，完成予想の提示に焦点を当てており，実際の編み動作に対する支援は対象としていない．

毛糸のシミュレーションに関して，Twigg-Smith ら[7]は，機械編みにおける柄表現において，引き上げ編みやすべり目を活用することで糸の使用量を削減する手法を提案した．ユーザがグリット状のエディタで各目の糸の色や編み目の種類を指定すると，システムが自動で機械編みした際のシミュレーション 3D 画像を生成する．Yukcel ら[8]は毛糸レベルのモデルを構築する手法を提案し，バーチャルキャラクターの衣服表現として，毛糸のセーターの表現を可能にした．また，かぎ針編みは自由度が高くコンピュータで扱うことが難しいが，Guo ら[9]はこれを数学的・幾何学的に定義可能なメッシュ構造として表す手法を提案した．これらの研究は主にシミュレーションや仮想環境での表現を対象としており，実際の製作時の手の動きを支援するものではない．

かぎ針編みの手指の動作に関して，木村ら[10]は棒針編みを対象とし，機械編みと初心者の手編みの完成品を画像解析により比較するとともに，毛糸を保持する左手人差し指に圧力センサを装着して動作を解析した．髙澤らはかぎ針編みの細編みを対象として，モーションキャプチャによる 3 次元計測や，各指に装着した 6 軸力センサによる指先反力の測定を行い，動作解析を実施した[11]．さらに，肘の固定の有無による比較から，編み動作の安定性や作業時間に差異が生じることを示した．また，アイトラッカーを用いた分析により，糸の引き量を視覚的に確認しながら調整していることを明らかにした[12]．加えて，LSTM を用いた動作モデルの構築により，異なる条件下での編み動作の比較が可能であることを示した[13]．川端ら[14]は鎖編みを対象として，手指及びかぎ針の動きと完成品の評価を行い，熟練者と初心者の差異や練習による上達過程を分析した．センサを用いてかぎ針編みを認識する研究としては，Wang らによる，IMU センサを用いてかぎ編み動作を認識し，編み手順を自動的に記録・編集・共有するシステム StitchFlow

が提案されている[15]．このシステムは創作過程の記録やパターン共有，デザイン探索を支援することを目的としている．一方，本研究では，一目ごとの編み動作をリアルタイムに検出し，音やゲーム画面によるフィードバックを提示することで，初心者の技能習得支援を目的としている点が異なる．これらの研究は，編み動作の計測・認識や事後的な分析を主な目的としており，編み動作の認識結果を初心者の技能習得のためのリアルタイムフィードバックに利用するものではない．

編んでいる際にフィードバックを提示する研究として，Felice ら[16]は，編み針に内蔵したセンサから動作を取得し，編み図と照合して誤りを検出し，音声または視覚的に警告を提示するシステムを提案している．しかし，このシステムは編み図の理解を前提としており，また専用の機器を用いる必要がある．また，Leap Motion Controller を用いて手のひらの向きを取得し音にマッピングする研究[17]や，棒針編みの動作を楽器入力として利用する研究[18]も提案されているが，いずれも技能習得支援を主目的としたものではない．センサを用いた繰り返し動作の支援としては，筆記音の強調フィードバックによる影響を検討した研究[19]をもとに，筆記行為の促進を目的とした製品 Write More が開発されている．

以上のように，編み物に関する研究は，設計支援，シミュレーション，動作解析，及び一部のフィードバック提示に関するものが提案されている．しかし，編み動作を一目単位で認識し，その結果を初心者の技能習得を支援するリアルタイムフィードバックへ直接利用する研究は十分に行われていない．

3. 動作解析システム

本研究では，かぎ針編みの動作を計測・解析し，技能習得を支援するための動作解析システムを開発した．本システムは，かぎ針に装着したセンサにより編み動作を取得し，そのデータに基づいて編み目の単位で動作を検出するとともに，リアルタイムに音によるフィードバックを提示する．また，ユーザが自身の動作データを基に閾値を設定することで，個人差や使用環境に適応可能な設計とした点に特徴がある．以下では，本システムの設計要件及び構成について述べる．

3.1 動作解析システムの要件

かぎ針編み初心者の動作を解析及び支援をするための要件を，以下のように設定した．

- **要件 1．軽量・簡便で動作を妨げない設計**
  手軽に導入可能なシステムとするため，小型・軽量で安価，かつ着脱が容易なセンサを用いる．また，編み動作への影響を最小限に抑える構造とする．
- **要件 2．作業中の視線を妨げないこと**
  編む作業では手元と参考資料（書籍・動画）を交互に確認するため，視覚的な負荷を増やさないことが重要である．そのため，作業中は聴覚フィードバックを行い，作業後に視覚的な振り返りを可能とする．
- **要件 3．技能習得支援と継続性**
  初心者の技能向上と継続的な練習を促すため，動作と出来栄えの関係を把握し，反復練習を支援する仕組みを提供する．
- **要件 4．ユーザ主導の分析方法**
  かぎ針の種類や糸の条件，個人差により動作が大きく異なるため，大規模な学習データを前提とする機械学習は適用が難しい．そこで，ユーザがセンサデータを基に閾値を設定する方式を採用する．



図 2 システム要件を満たす構成
Fig. 2 System configuration that satisfies the design requirements

3.2 動作解析システムのハードウェア構成

要件 1 を満たすため，小型で無線通信が可能なセンサデバイスである M5StickCPlus を採用した．本デバイスは加速度・角速度センサを内蔵しており，軽量で取り付けが容易である．プログラムの書き込みには Arduino 言語を用い，データ処理には Python を用いた．かぎ針への装着方法として，四角い鉛筆キャップを 3 つ連結した構造を用い，中央のキャップにかぎ針上部を差し込むことで着脱を容易にした．これにより，異なる環境でも柔軟に利用可能である．用いるかぎ針のサイズは，持ち手の上部の太さが鉛筆キャップに合っている 7/0 号とした．また，フットペダルを用いて記録の開始・停止を行うことで，作業中に手を離す必要をなくし，要件 2 を満たす設計とした．

以上により，本システムは図 2 に示す構成となる．

3.3 ユーザインタフェース

ユーザインタフェースの利用フローを図 3 に示す．図中の赤字は閾値を設定する場合のみの処理を表し，括弧内はシステムの挙動，括弧外はユーザの動作を示す．

まず，ユーザは作り目を終えた後にプログラムを実行する (a)．その後，Bluetooth 接続が確立されると，フットペダルを踏むことで記録開始のタイミングを入力する (b)．記録開始時及び終了時にはそれぞれ異なる音が鳴るため，ユーザは入力が正しく行われたことを確認できる．

記録中，ユーザは編み動作を行う (c)．編み終えた後，

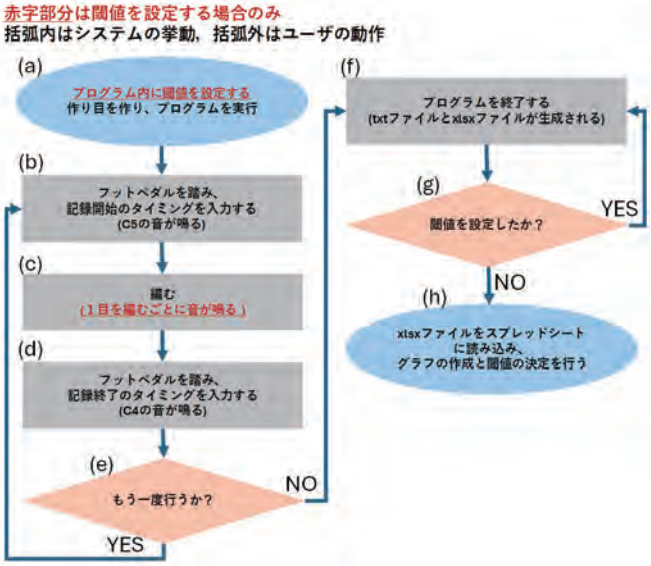


図 3　システム利用時のフローチャート
Fig. 3　Flowchart of system usage

フットペダルを再度踏むことで記録終了のタイミングを入力する (d)．その後，続けて記録を行うかを選択し (e)，継続する場合は再びフットペダルを踏んで記録を再開し，終了する場合はプログラムを終了する (f)．一方，事前に閾値 T_h 及び T_l を設定している場合には，編み動作中に各編み目のタイミングで音が提示される (c)．このとき，記録開始・終了時とは異なる音を用いる．この閾値は，記録後にデータをスプレッドシートに読み込み，グラフを作成して決定する (h)．

このように，本システムはフットペダル操作と音フィードバックを組み合わせることで，視線を妨げることなく直感的な操作を可能にした．

3.4　プログラム実装

本システムでは，3 軸加速度，3 軸角速度，3 軸姿勢角の時系列データを記録する．加えて，フットペダル操作により記録開始・終了のタイミングを取得する．記録データに基づき，鎖編み動作を検出するための閾値を設定する．ここで正の閾値を T_h，負の閾値を T_l とする．これらの閾値を目視で算出し，これらの閾値を同じシステムに設定して使用すると，一目編むごとに音を鳴らすことができる．一目編むごとに音を鳴らすか否かもプログラム内のパラメータで設定可能である．これらは要件 1 を満たす．

鎖編みにおいて，編んだことを判定するタイミングは図 4 に示すように，一度負の閾値を下回った後正の閾値を超えた場合を high，一度正の閾値を上回った後に負の閾値を下回った場合を low とすると high と low の組み合わせ 1 つにつき一目編んだと判断する．

この手法により，個人差や環境差に対応しつつ，要件 4 を満たすユーザ適応型の解析が可能となる．さらに，検出された編み動作に基づいて音を提示することで，リアルタイムフィードバックを実現する．

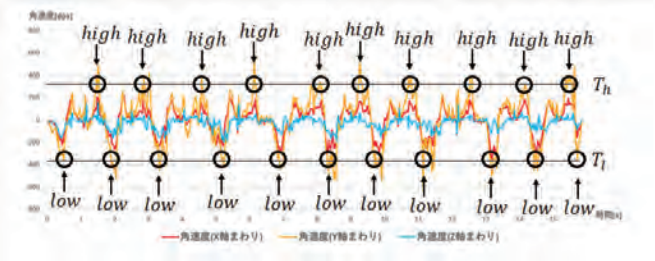


図 4　閾値に基づく編み目検出の判定方法
Fig. 4　Stitch detection method based on thresholds

4.　ゲーミフィケーションを利用した技能習得支援

本研究では，要件 3 である技能習得の継続性を高めるために，動作解析システムにゲーミフィケーション要素を付加したアプリケーションを設計・実装した．本アプリケーションは，動作解析システムと同一のアルゴリズム及びハードウェア構成を用いながら，フィードバックの提示方法を変えることで，ユーザのモチベーション向上を図るものである．

本システムでは，センサから取得した値に基づき，鎖編み 1 目の動作を検出するアルゴリズムを用いる点は動作解析システムと共通である．一方で，従来は音によるフィードバックを提示していたのに対し，本アプリケーションではその検出結果をゲームの入力として利用する．

具体的には，ユーザが編み動作を行うと，1 目編まれたと判定されるごとに画面上の編み目が増加する（図 5）．編み目は下方向に積み重なるように表示され，ユーザは自身の動作に応じて視覚的な進捗を確認できる．クリア条件は一定数の編み目を達成することである．本実装では 10 目と設定したが，この値はユーザが任意に変更可能であり，習熟度や練習目的に応じた調整が可能である．

本アプリケーションは，動作解析システムと同様に Arduino 言語及び Python を用いて実装し，ユーザインタフェースには Tkinter 及び Pygame を用いた．ハードウェア構成も動作解析システムと同一である．

このように，本研究では同一の動作解析アルゴリズムに対して，音フィードバックとゲーム型フィードバックという異なる提示手法を設計することで，技能習得支援の新たな可能性を示す．

5.　実　　験

要件 3 である上達支援及び継続性の確保に関する有効性を検証するとともに，経験の有無や個人差による動作の違いを調査するための実験を行った．

5.1　実験方法

複数人の結果を比較するため，実験参加者全員が同一のかぎ針 (7/0 号) 及び毛糸を用いる条件とし，同一のハード



(a) ゲームスタート時　　(b) 6 目編んだところ

図 5　センサにより検出された編み動作に応じて編み目が増加するゲーム型フィードバックの画面

Fig. 5　Game-based feedback interface in which stitches increase based on knitting motions detected by the sensor

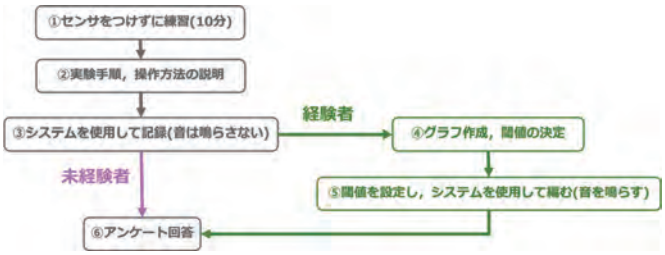


図 6　予備実験の流れ

Fig. 6　Preliminary experimental procedure

ウェア構成で実験を行った．以降，本論文の第 1 著者であり，かぎ針編み経験者である実験主催者は「実験者」と表記する．また，鎖編み 1 本あたりの目数は 10 目とした．

予備実験と本実験に共通して，未経験者にはセンサを装着していないかぎ針を用い，書籍及び動画を参照しながら 10 分間の練習を行ってもらった．書籍及び動画は右利き用であるため，左利きの参加者には画像及び動画を左右反転して提示した．経験者には，動作の確認を目的とした短時間の練習を行ってもらった．なお，予備実験及び本実験のいずれにおいても，実験参加者には大学の規定に基づく謝金を支払った．

5.2　予備実験

一目編みの判定に用いる閾値に適したセンサ軸を検討するため，図 6 に示す手順で予備実験を実施した．

実験参加者の属性は女性 4 名 (経験者 3 名，未経験者 1 名，平均年齢 23.75 歳，SD=1.26) で，以降 e1，e2，e3，b4 とする．e が経験者，b が未経験者を表す．本実験の設計を検討する目的から，実験参加者ごとに編む本数や編み方を適宜調整しながら実施した．アンケートの結果，未経験者 b4 からは「作り目が難しい」との回答が得られた．また，経験者 e1, e2, e3 に対して音フィードバックを有効にした状態での編み動作を記録した結果，X 軸及び Y 軸周りの角速度が，一目編みの判定に有効な特徴量であることが示唆された．

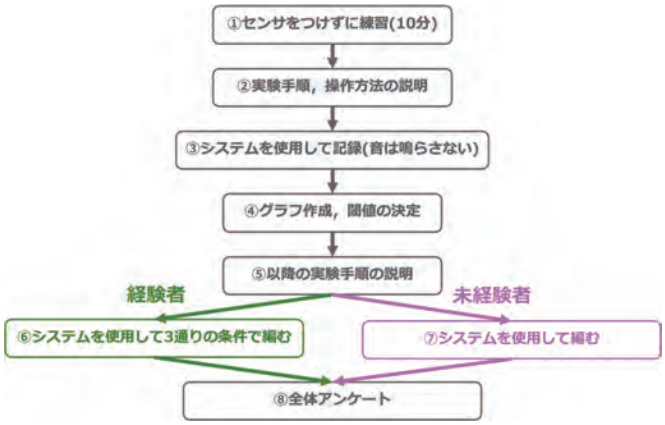


図 7　本実験の流れ

Fig. 7　Flow of the experiment

表 1　経験者でのフィードバック提示順序

Table 1　Counterbalanced feedback sequences for experienced participants

参加者 ID	条件	手法 1・2 の順番
E5	1: 音のみ	手法 1 → 手法 2
	2: 画面のみ	手法 1 → 手法 2
	3: 音と画面両方	手法 2 → 手法 1
E6	1: 音のみ	手法 2 → 手法 1
	2: 音と画面両方	手法 2 → 手法 1
	3: 画面のみ	手法 1 → 手法 2
E7	1: 画面のみ	手法 1 → 手法 2
	2: 音のみ	手法 2 → 手法 1
	3: 音と画面両方	手法 1 → 手法 2
E8	1: 画面のみ	手法 2 → 手法 1
	2: 音と画面両方	手法 1 → 手法 2
	3: 音のみ	手法 2 → 手法 1
E9	1: 音と画面両方	手法 1 → 手法 2
	2: 音のみ	手法 2 → 手法 1
	3: 画面のみ	手法 2 → 手法 1
E10	1: 音と画面両方	手法 2 → 手法 1
	2: 画面のみ	手法 1 → 手法 2
	3: 音のみ	手法 1 → 手法 2
E11	1: 画面のみ	手法 1 → 手法 2
	2: 音のみ	手法 2 → 手法 1
	3: 音と画面両方	手法 1 → 手法 2

5.3　本実験

予備実験の結果に基づき，閾値の設定及びゲーミフィケーションを用いたフィードバックの有用性を調査するため，本実験を図 7 に示す手順で実施した．

図 7 のステップ (4) では，各実験参加者が閾値を設定できるように，角速度 3 軸のうち変化量の大きい軸に着目し，その極大値と極小値が可視化されたスプレッドシートを提示した．実験参加者はこれを参照しながら閾値を決定した．

実験参加者は予備実験とは異なる女性 14 名 (経験者 7 名，未経験者 7 名，平均年齢 21.86 歳，SD=2.21) であり，以降それぞれ E5〜E11, B12〜B18 と表記する．ここで，E が経験者，B が未経験者を表す．

5.3.1　経験者

経験者に対しては，図 7 でのステップ (7) において，フィードバック方法として「音のみ」「画面のみ」「音・画面両方」

の 3 条件と，システムが設定した閾値を用いる「手法 1」，実験参加者が設定した閾値を用いる「手法 2」を組み合わせ，表 1 に示すように実験を行った．実験条件の提示順による影響を抑えるため，E5〜E10 の 6 名について，「音のみ」「画面のみ」「音と画面両方」の 3 条件の提示順をカウンタバランスした．また，各条件における手法 1 及び手法 2 の提示順についてもカウンタバランスした．なお，フィードバック条件と手法 1・2 の組合せを含めた完全なカウンタバランスは行っていない．7 人目の E11 の条件は，かぎ針の持ち方が他の実験参加者と異なっていた E7 と同じ条件とした．E7 はナイフ持ち，その他の実験参加者はペン持ちであった．画面はノート PC の画面を提示した．

また，実験参加者には各フィードバック条件を対応する記号で提示し，「音のみ」「画面のみ」「音・画面両方」という具体的な内容は明示しなかった．さらに，「手法 1」と「手法 2」の違いについても事前には説明せずに実験を実施した．

実験終了後，質問紙による事後アンケートを実施した．アンケートでは，利き手，かぎ針編みの経験，センサの重量に対する印象，ユーザ自身による閾値設定機能の必要性，音・画面によるフィードバックの初心者への有用性，3 種類のフィードバック手法（音のみ，画面のみ，音と画面両方）の有用性の順位とその理由，システムに対する改善点，及び実験全体に関する感想などを尋ねた．また，一部の設問では自由記述による回答を求めた．

実験終了後に手法 1 及び手法 2 の違いを説明した上で，システムが設定した閾値からユーザ自身が閾値を調整できる機能 (手法 2) の必要性について尋ねた．その結果，7 名中 6 名が必要であると回答した．一方，1 名は「音が鳴るタイミングがわかりにくく，どのように活用すればよいか分からなかった」と回答した．

また，「音のみ」「画面のみ」「音・画面両方」の 3 条件について，初心者に対する有用性の観点から 1 位から 3 位の順位付けを求めた．その結果，「音のみ」の平均順位は 2.29，分散は 0.49，「画面のみ」の平均順位は 2.00，分散は 0.86，「音・画面両方」の平均順位は 1.71，分散は 0.49 となった．「音・画面両方」が平均順位が最も高く，分散も小さいため，有用性が高いと評価された．

初心者であった頃にこのようなシステムがあったらどう思うか尋ねたところ，6 名が肯定的な意見を述べた．1 名は「編み目を数えるのが面倒かつわからなくなりやすいので，そこを正確に数えて支援してもらえるシステムがあれば嬉しい」と述べた．また，今回の実験ではゲーミフィケーション要素を付加したアプリケーションは鎖編みにのみ対応していたが，他の編み方でも利用できたらどうかと尋ねたところ，6 名が肯定的な意見を示した．1 名は「同じ編み方を繰り返すときは有用だと考える．鎖編みはわかりやすいが，長編みとかになると逆に複雑に感じてしまう可能性もある」と述べ，編み方の種類によるシステムの有用性の違いに言及した．

音や画面に関する改善点については，2 名が音質，2 名が音の変化の無さ，1 名が画面の大きさ，1 名が精度について言及した．音質については「音が少しびっくりした．もう少し柔らかい音で良い」「『ぽこぽこ』という音だったので，もう少し優しい音が望ましい」といった意見があった．音の変化については「何回編んだか忘れやすいため，単純な音でなく『1，2，3』などの回数を示す音声フィードバックがあると有用」「音が何を表しているかわかる音であればさらにわかりやすい」との意見があった．画面の大きさについては「手元から目を離して画面を見ることが難しいため，表示がより大きければ自然と視界に入る」との意見があった．

さらにシステム全体についての改善点として，「値の変化が音や画面にどのように影響するか，初心者でもわかりやすいと良い」との意見が得られた．

5.3.2 未経験者

予備実験における未経験者 b4 のアンケート結果を踏まえ，未経験者のセンサデータの記録では，1 本目及び 3 本目は実験者が作り目を行い，2 本目及び 4 本目実験参加者自身が作り目を行う条件とした．しかし，実験者と実験参加者による作り目の違いによる顕著な差は確認されなかった．

また，経験者を対象とした本実験のアンケート結果に基づき，未経験者に対するフィードバック提示方法は「音・画面両方」とした．

6. 考　　察

本章では，実験結果及びアンケート結果に基づき，提案手法の有効性と課題について考察する．

6.1 動作センシングと編み目検出の妥当性

予備実験の結果より，かぎ針編みの一目の動作において，X 軸及び Y 軸周りの角速度が有効な特徴量となることが示唆された．これは，かぎ針編みの動作が主に手首の回転運動を伴うことと一致しており，センサによる動作取得が編み動作の本質的な特徴を捉えていることを示している．

また，本実験においても閾値に基づく編み目検出が一定程度機能しており，単純なルールベース手法であっても，個人の動作に適応させることで実用的な検出が可能であることが確認された．特に，大規模な学習データを必要とせずにユーザごとに調整可能である点は，本手法の実用性を示すものである．

一方で，動作のばらつきにより，編み目が閾値を超えず未検出となる場合も確認された．図 8 にその一例を示す．図中では，検出された編み目を緑の実線，実際に編み動作が行われているものの閾値を超えず未検出となった編み目を

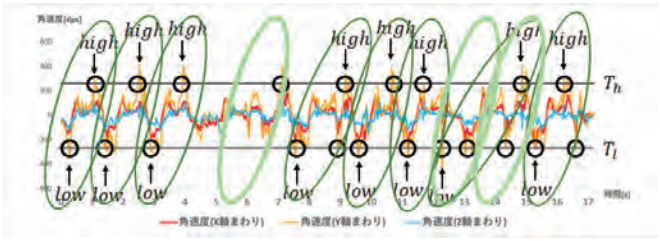


図 8　動作のばらつきにより一部の編み目が未検出となった例（緑実線：検出された編み目，薄緑マーカー線：未検出の編み目）

Fig. 8　Example of missed stitch detection due to motion variations (solid green lines: detected stitches, light green marker lines: undetected stitches)

薄い緑のマーカー線で囲んだ．実際の編み動作では 10 目編んでいるのに対して，設定した閾値に基づく検出では 8 目として検出されている．例えば，3 つ目と 4 つ目の high の間に現れるピークは実際の編み動作に対応するものの，正の閾値 T_h を超えていないため，編み目として検出されていない．このような未検出は，動作のばらつきによって生じるものであり，単純に閾値を下げると他の動作を編み目として検出する可能性がある．したがって，編み目検出精度の向上には，個人差や動作のばらつきを考慮した特徴量や検出方法の改善が今後の課題である．

6.2　ユーザ適応型閾値設定の有効性

経験者を対象とした実験では，7 名中 6 名がユーザによる閾値調整機能の必要性を指摘した．これは，編み方やリズム，かぎ針の持ち方といった個人差が大きい作業において，固定的な閾値では十分に対応できないことを示している．

また，「音が鳴るタイミングがわかりにくい」との意見も得られており，閾値とフィードバックの関係が直感的に理解しにくい可能性がある．そのため，動作とフィードバックの対応関係を可視化・説明するインタフェース設計が重要であると考えられる．

6.3　フィードバック手法の違いと技能習得支援

フィードバック手法の比較では，「音・画面両方」が最も高く評価された．この結果は，聴覚と視覚を組み合わせたマルチモーダルなフィードバックが，編み動作の理解及び進捗把握に有効であることを示唆している．

一方で，「画面を見る余裕がない」「音が作業の妨げになる場合がある」といった指摘もあり，最適なフィードバックはユーザの熟練度や状況に依存する．特に初心者においては認知負荷の増加を避けるため，フィードバックの種類や強度を調整可能とする適応的な設計が望ましい．

6.4　ゲーミフィケーションの可能性

ゲーム型フィードバックについては，画面上に編み目がリアルタイムに生成されることに対して，「モチベーションが上がりそう」「編んだ達成感が得られる」といった肯定的な意見が得られた．また，「楽しかった」とする初心者の意見も得られたことから，編み目の生成を視覚的に提示することが，達成感や楽しさを伴う練習体験につながる可能性が示唆された．自らの動作が即座に結果として反映される点は，反復練習の促進に有効であると考えられる．

一方で，編み方が複雑になる場合には単純なカウントベースのフィードバックでは対応が難しくなる可能性があり，技能の複雑さに応じた表現方法の設計が必要である．

6.5　ハードウェア設計と使用感の影響

センサの重量について回答した 18 名のうち 11 名が，「重い」「少し重い」「やや重い」など，センサの重量を気にした回答をした．一方で，「短時間だからそこまで気にならなかった」「編んでいるときの違和感はあまり感じなかった」といった意見もあり，重量を感じたことが必ずしも実験中の大きな負担につながったわけではなかった．また，「動きが引っ張られる感じがした」「遠心力が増すように感じた」「センサが端に取り付けられていたので重さを感じやすかった」といった意見も得られ，センサの重量だけでなく，装着位置や重心が編み動作に影響する可能性が示唆された．

この結果から，センサの小型化・軽量化に加え，装着位置や重心バランスの最適化が重要な課題である．

6.6　初心者支援への適用可能性

初心者に関する結果からは，作り目や糸の保持といった初期段階の技能に困難があることが確認された．また，編み目数の把握の難しさに対して，本システムのカウント機能が有用である可能性が示された．

一方で，音フィードバックの意味が理解しにくい場合や，リズムに慣れるまでに時間を要する点も指摘されており，画面を注視しなくても音のみで進捗を把握できるフィードバック設計が求められる．

以上より，本システムは初心者の技能習得を支援するインタフェースとして有用である可能性を示した．さらに，本研究で取得した動作データは，編み目検出にとどまらず，技能の質的評価へと拡張可能である．例えば，動作の周期性や一貫性といった特徴量を定量化することで，技能指標の設計につながると考えられる．

7.　まとめと今後の課題

本研究では，かぎ針編みにおける動作センシングと技能習得支援を目的として，かぎ針に装着したセンサにより編み動作を計測・解析するシステムを提案した．提案システムでは，センサデータに基づいて編み目単位の動作を検出し，音によるリアルタイムフィードバックを提示することで，初心者の反復練習を支援する．さらに，ユーザが閾値を設定することで個人差や使用環境に適応可能な仕組みを実現した．また，同一の動作解析アルゴリズムを用いて，音によるフィードバックに加え，ゲーム型フィードバック

を提示するアプリケーションを実装し，技能習得支援における提示方法の多様性を示した．これにより，楽しさを伴う継続的な練習支援の可能性を示した．

本研究では，個人差や環境差（糸の種類，針の太さ，持ち方）が大きい編み物の特性を踏まえ，機械学習ではなくユーザ調整可能な閾値ベース手法を採用した．これにより，少量のデータでも即時適応が可能となる点に利点がある．

一方で，本研究にはいくつかの制限がある．センサのバッテリー持続時間や Bluetooth 接続の安定性といった運用上の課題に加え，フィードバックが本来の編み動作に影響を与える可能性がある．これらが技能習得に与える影響については，今後の詳細な検証が必要である．また，本研究では予備実験において最も特徴的な変化が観測された Y 軸角速度を用いて編み目を検出した．しかし，かぎ針の持ち方や個人差によって支配的な動作軸は変化する可能性があり，複数軸を統合した認識手法については今後検討が必要である．また，本研究では編み目検出精度を定量的には評価しておらず，今後は動画との比較などにより認識精度を評価する予定である．

さらに，本研究では短時間の利用実験を実施しており，技能向上そのものを客観的に評価したものではない．本研究で示したのは，利用者がリアルタイムフィードバックを有用と感じたこと，及び技能習得支援への可能性である．今後は長期間の利用実験を実施し，編み速度，編み目の均一性，動作の安定性などを指標として，技能向上を定量的に評価する予定である．さらに，角加速度や移動平均など他の特徴量を用いた編み目検出についても，認識精度向上の観点から今後検討する予定である．

本手法は経験者の動作分析や技能評価指標の構築への応用も期待される．また，初心者における糸の保持や作り目といった初期技能への支援についても今後の課題である．さらに，今後は複数の実験参加者のデータを用いた統計的手法や機械学習による閾値推定・自動最適化を検討する．

本研究のアプローチは編み物に限らず，書字や楽器演奏などの繰り返し動作を伴う技能にも適用可能である．特に，「動作単位の検出」と「リアルタイムフィードバック」を組み合わせた枠組みは，身体技能の習得支援における汎用的な設計指針となり得る．

このように，本研究は，手工芸における暗黙的な身体技能をセンシングとフィードバックにより外在化し，習得支援へと接続する試みとして位置付けられる．

謝　　辞

本研究を進めるにあたり，センサについてご助言いただいた石田こなつ氏，的場やすし氏に感謝する．また，実験にご協力くださった皆様に感謝する．本研究の一部は JSPS 科研費 JP25K15207，髙田弘子奨学金，JST BOOST の支援を受けて行われたものである．

参 考 文 献

1) 矢野きくの: “なぜ今、空前の「編み物ブーム」？ その火付け役や品薄の現状は？”, Yahoo ニュース（オンライン）, https://news.yahoo.co.jp/expert/articles/04f07cc14a3a9e09785fd780cc58332f6e2a2da3 (2025). (閲覧日：2026 年 7 月 13 日)
2) H. Keyes, S. Gradidge, S.-E. Forwood, N. Gibson, A. Harvey, E. Kis, K. Mutsatsa, R. Ownsworth, S. Roeloffs, M. Zawisza. “Creating Arts and Crafting Positively Predicts Subjective Wellbeing”, *Frontiers in Public Health*, Vol. 12, p. 1417997 (2024).
3) minne. https://minne.com/
4) Y. Igarashi, T. Igarashi, H. Suzuki. “Knitting a 3D Model”, *Computer Graphics Forum*, Vol. 27, No. 7, pp. 1737–1743 (2008).
5) 五十嵐悠紀, 五十嵐健夫, 鈴木宏正: “あみぐるみのための 3 次元モデリングと製作支援インタフェース”, コンピュータソフトウェア, Vol. 26, No. 1, pp. 51–58 (2009).
6) K. Wu, H. Swan, C. Yuksel. “Knittable Stitch Meshes”, *ACM Tran. on Graphics*, Vol. 38, No. 1 (2019).
7) H. Twigg-Smith, E. Whiting, N. Peek: “Knitscape: Computational Design and Yarn-Level Simulation of Slip and Tuck Colorwork Knitting Patterns”, In Proc. of the CHI Conference on Human Factors in Computing Systems (CHI ’24) (2024).
8) C. Yuksel, J.-M. Kaldor, D.-L. James, S. Marschner: “Stitch Meshes for Modeling Knitted Clothing with Yarn-Level Detail”, *ACM Trans. on Graphics*, Vol. 31, No. 4 (2012).
9) R. Guo, J. Lin, V. Narayanan, J. McCann: “Representing Crochet with Stitch Meshes”, In Proc. of the ACM Symposium on Computational Fabrication (SCF ’20) (2020).
10) 木村優希, 戸田沙也香, 藤井浩光: “編み物における巧拙評価のための編み目模様と手先動作の解析”, ロボティクス・メカトロニクス講演会講演概要集, Vol. 2021, pp. 1P1-H06 (2021).
11) 髙澤怜里, 藤井浩光: “編み物における技能理解のための編み目形状生成に関わる引き動作中の視線解析システム”, ロボティクス・メカトロニクス講演会講演概要集, Vol. 2022, pp. 2A2-E02 (2022).
12) 髙澤怜里, 藤井浩光, 山野辺夏樹: “編み物における技能理解のための編み目形状生成に関わる引き動作中の視線解析システム”, 計測自動制御学会システムインテグレーション部門講演会, 第 23 巻, pp. 1A2-C12 (2022).
13) 髙澤怜里, 上田隆一, 山野辺夏樹, 藤井浩光: “編み物における技能理解のための LSTM による手先動作の解析”, ロボティクスシンポジア予稿集, 第 30 巻, p. 2C1 (2025).
14) 川端博子, 中谷俊裕: かぎ針編みの学習意識と上達からみる家庭科の製作題材提案に関する研究＜教育科学＞. 埼玉大学紀要 教育学部, Vol. 66, No. 2, pp. 117–128 (2017).
15) Z. Marciniak, P. Lertjaturaphat, A. Bianchi: “StitchFlow: Enabling In-Situ Creative Explorations of Crochet Patterns With Stitch Tracking and Process Sharing”. In Proc. of the 38th Annual ACM Symposium on User Interface Software and Technology (UIST ’25), Article 7, 1–15 (2025).
16) K. Felice: “Interactive Knitting and Crocheting System”, US Patent 8,615,319 (2011).
17) J.-E. Bruen, H. Kwon, M. Jeon: “Cro-Create: Weaving Sound Using Crochet Gestures”, In *Proceedings of the Conference on Creativity and Cognition*, pp. 334–337 (2023).
18) Enrique Encinas, Konstantia Koulidou, Robb Mitchell: “The Kraftwork and the Knittstruments: Augmenting Knitting with Sound”, In Proc. of the Augmented Human International Conference, pp. 49–52 (2015).
19) 金ジョンヒョン, 橋田朋子, 大谷智子, 苗村健: “筆記音のフィードバックが単純な筆記作業に及ぼす影響の検討”, 日本バーチャルリ

アリティ学会論文誌, Vol. 17, No. 3, pp. 289–292 (2012).
(2026 年 4 月 6 日 受付)
(2026 年 9 月 5 日 再受付)



木　下　　梓
2024 年 お茶の水女子大学理学部情報科学科卒業．2026 年 お茶の水女子大学大学院人間文化創成科学研究科理学専攻情報科学コース修了．在学中，HCI に関する研究に従事．



五十嵐　悠　紀（正会員）
2010 年 東京大学大学院工学系研究科博士課程修了．博士（工学）．日本学術振興会特別研究員 DC2, PD, RPD を経て 2015 年より明治大学総合数理学部先端メディアサイエンス学科専任講師，2018 年より同准教授．2022 年よりお茶の水女子大学理学部情報科学科准教授．2025 年より東京大学先端科学技術研究センター准教授（クロスアポイントメント）．コンピュータグラフィックス及びユーザインタフェースに関する研究に従事．

IIEEJ ショートペーパー -- ビジュアルコンピューティング特集--

ストックの八重鑑別における日齢の影響検証

堀越雄太† 奥田萌莉† 日浦慎作†(正会員)

†兵庫県立大学

Verification of the Effect of Seedling Age on Double-Flower Discrimination in Stock (*Matthiola incana*)

Yuta HORIKOSHI†, Moeri OKUDA†, Shinsaku HIURA†(*Member*)

†University of Hyogo

〈あらまし〉 ストック (*Matthiola incana*) は切り花として利用される重要な園芸作物であり，一重咲きと八重咲きが存在する．特に八重咲きは市場価値が高いため，苗段階で将来の開花型を予測する「八重鑑別」が行われている．しかし，目視による鑑別は熟練を要し，判定の属人化が課題となっている．本研究では，画像情報に加えて播種後日数（日齢）という時間的情報を活用し，生育段階の個体差を考慮した安定的な分類手法の有効性に関し，その検証を行った．姫路市産業局農業振興センターで撮影したセンターデータセット（152 枚）と，既存の MiLIC データセット（1,350 枚）を用い，(1) train/test 間の日齢不一致の影響，(2) 日齢ベクトルによる明示的日齢情報の導入，(3) 3 時系列画像による擬似的日齢情報の導入，の 3 点について評価を行った．その結果，日齢が一致しているほど分類性能が向上する傾向が確認された．日齢ベクトルの導入は，train/test 間で日齢が一致している条件下において性能向上に寄与した．一方，3 時系列画像の利用は日齢の一致・不一致にかかわらず最も高い分類性能を示した．これらの結果は，日齢情報を適切に活用することで，八重鑑別の性能および安定性の向上が期待できる可能性を示唆している．
キーワード：ストック, 日齢, 八重鑑別

＜***Summary***＞ *Matthiola incana* (stock) is an important ornamental cut flower species that exhibits single- and double-flowered phenotypes. Because double-flowered plants have higher market value,“double-flower discrimination”is performed at the seedling stage to predict flowering type. However, visual inspection requires considerable expertise and suffers from subjectivity. In this study, we investigated a more stable classification approach that incorporates temporal information—specifically, seedling age (days after sowing)—in addition to image data, in order to account for individual differences in growth stage. Experiments were conducted using a Center dataset (152 images) collected at the Himeji City Agricultural Promotion Center and the existing MiLIC dataset (1,350 images). We evaluated (1) the effect of seedling age mismatch between training and test data, (2) the introduction of explicit seedling age information using an age vector, and (3) the introduction of pseudo age information using three time-series images. The results showed that classification performance tended to improve when seedling ages were matched between training and test data. The age vector improved performance under the matched-age condition, whereas the use of three time-series images achieved the highest classification performance regardless of age matching. These findings suggest that appropriately incorporating seedling age information can enhance the accuracy and robustness of double-flower discrimination in stock.
Keywords: matthiola incana, seedling age, double-flower discrimination

1. はじめに

ストックとは切り花などに用いられる園芸用の花の一種であり，**図 1**[1)]に示すように一重咲きと八重咲きが存在する．八重咲きはその華やかさから市場価値が高く，効率的な生産を目的として八重鑑別が行われている[1)]．八重鑑別とは苗の段階で八重咲きとなるか推測，厳選することである．しかし，目視による鑑別は熟練の経験を要し，判定基準の属人化が課題となっている．これに対し,CNN や ViT など近年の機械学

習技術の発展により, 医用画像等の様々な分野で人と同等もしくはそれを凌駕する鑑別が可能になりつつあり, これを八重鑑別に応用することで効率的な生産と省力化の両立が可能になると考えられる.

八重咲きと一重咲きは同じ日齢であっても苗の大きさが異なることが知られており，八重咲きのほうが一重咲きよりも大きい傾向がある[2)]．一方，苗は生育に伴い大きくなるため，学習データと鑑別対象の苗の日齢が一致しない場合，日齢に起因する大きさの違いが八重・一重の大きさの違いと混同され，八重・一重の判別性能が低下するという仮説が立てられる．この仮説のもと，本研究では画像情報に加えて日齢に関する時間的情報を活用し，生育段階の個体差を考慮したより安定した分類手法の検証を行った．なお，日齢とは播種日（種をまいた日）から経過した日数を指す.

先行研究[1)]では，ストックの八重咲き・一重咲きを含むMiLIC データセットが構築され，画像認識モデルを用いた二値分類の検証が行われた．Apricot Quartet（以下，Apricot と呼ぶ）という品種を対象とした検証において，Efficient-NetV2 を用いたモデルが約 8 割の分類性能を示すことが報告されている[1)]．その一方で，学習データと鑑別対象苗の日齢の不一致が分類性能に与える影響については検討されておらず，日齢を考慮した分類手法の構築が課題として残されている.

2. 用いたデータセットとリサーチクエスチョン

2.1 節で本研究に用いたデータセットについて，2.2 節で 3 つのリサーチクエスチョンについて，2.3 節で用いたモデルの概略について述べる.

2.1 データセット

本研究では，姫路市産業局農業振興センターにて「ミニクロ レッド」という品種のストックを 1 回撮影し，ストック画像 152 枚（八重：109 枚，一重：43 枚）を作成した．この画像は日齢が 6 日目である．熟練者により行われた八重鑑別の結果に基づき一重もしくは八重のラベルが付与されており，開花については確認されていない．これをセンターデータセットと呼ぶこととする．また，先行研究[1)]で作成された MiLIC データセットに含まれる Apricot という品種のストック画像 1,350 枚も併せて用いた．MiLIC データセットのストック画像は，グループ A およびグループ B の 2 個体群を 3 日間連続で撮影したものであり，各グループ内で苗の日齢は 1 日ずつ異なる.



図 1 一重咲き (左) および八重咲き (右) のストック[1)]

Fig. 1 Single-flowered (left) and double-flowered (right) stock[1)]

センターデータセットと MiLIC データセットの日齢は異なる．具体的には，MiLIC データセットの Apricot の苗はグループ A・B ともに日齢が約 2 週間とされているが[1)]，詳細な日数は不明であり，生育状況から A と B の日齢は異なるものと考えられる．センターデータセットはそれよりも早期の段階である日齢 6 日目に撮影されたものである.

2.2 リサーチクエスチョン

検証は以下の 3 点について行った.

RQ1：学習データの日齢に近い日齢の画像が入力された場合，分類性能はどの程度向上するか（3.1 節）

センターデータセットを test データとして固定し，MiLIC データセットのグループ A を撮影日（撮影 1 日目・2 日目・3 日目）ごとに切り替えて train データとして使用した．グループ A は播種から約 2 週間後に 3 日間連続で撮影されているため，撮影日が後になるほど test データである日齢 6 日目のセンターデータセットとの日齢差が拡大する．この日齢差の段階的な変化を利用することで，train/test 間の日齢のずれが分類性能に与える影響を検証した.

RQ2：日齢ベクトルの導入は分類性能を向上させるか，また，その効果は train/test 間の日齢が一致する場合と一致しない場合のどちらで有効に機能するか (3.2 節)

パターン 1（日齢不一致）：MiLIC データセットのグループ A （八重：408 枚，一重：291 枚）を train データ，グループ B（八重：396 枚，一重：255 枚）を test データとして用いた．グループ A とグループ B は同一品種であるが日齢が異なるため，train/test 間で日齢が一致しない条件に相当する.

パターン 2（日齢一致）：MiLIC データセットのグループ A を均等に分割し，八重：204 枚，一重：146 枚を train データ，八重：204 枚，一重：147 枚を test データとして用いた．同一グループ内での分割であるため，train/test 間で日齢が一致する条件に相当する.

RQ3：日齢ベクトルを用いない場合において，時系列画像の導入は日齢ベクトルの代替手段として分類性能の向上に寄与するか (3.3 節)

MiLIC データセットには各グループにつき 3 日分の画像が存在することを活用し，パターン 1（日齢不一致）とパターン 2（日齢一致）を検証した．3 日分のうちそれぞれの画像を個別のモデルに入力してセットとして処理する並列構成を採用した．比較のため，1 時系列および 2 時系列についても同様に検証した.

2.3 実験設定

3.1 節および 3.2 節において，ImageNet-1k で事前学習された Vision Transformer の一種である ViT-B/16 を基盤としたモデルを活用した．全体図を**図 2** に示す．ViT-B/16 は，入力画像を 16 × 16 ピクセルのパッチに分割し，各パッチを全結合層を通して 768 次元のベクトル（トークンと呼ぶこととする）に変換する．先頭には，クラストークンと呼ばれる特殊トークンが追加される．クラストークンには画像全体の情報が集約される．学習の際には，patience を 10 とし，Adam を最適化関数として 15epoch で早期終了した．

3.3 節において，ViT-B/16 を 3 並列で用いることが理想であったが，GPU メモリの制約上，本検証では timm の ImageNet-1k で事前学習された Vision Transformer の一種である ViT-Small（vit_small_patch16_224）を ViT-B/16 の軽量代替として採用した．ViT-Small は，ViT-B/16 と同様のパッチ分割方式を採用しつつ，モデルパラメータ数を削減した軽量版である．学習の際には，patience を 10 とし，Adam を最適化関数として 20epoch で早期終了した．

3. 検　　証

3.1 RQ1：学習データの日齢に近い日齢の画像が入力された場合，分類性能はどの程度向上するか

センターデータセットと MiLIC データセットの日齢が相違し，MiLIC データセットのほうが日齢が大きいことを利用して検証する．そのために本検証では，センターデータセットを test データとして固定し，MiLIC データセットのグループ A を日齢（撮影 1 日目・2 日目・3 日目）ごとに切り替えて train データとして使用した．グループ A は 3 日間連続で撮影されているため，撮影日が 3 日目に向かうほど，test データとして設定したセンターデータセットとの日齢差が拡大する．

センターデータセットを test データとして固定し，train データを MiLIC データセットのグループ A の日齢を変化させた場合の結果を**表 1** に示す．撮影 3 日目に向かうほど，つまり train/test 間の日齢差が拡大するにつれ，分類性能が低下する傾向が確認された．具体的には，撮影 1 日目の画像を学習させた場合（正解率 80.2%）は，撮影 3 日目の画像を学習させた場合（正解率 72.3%）に比べて性能が約 8%向上した．また ROC-AUC, PR-AUC については 2 日目の値が最も高くなっているが 1 日目に対してその差はわずかであり，3 日目では大きく低下していることから，やはり日齢がずれるほど分類性能が低下する傾向があると言える．

この結果より，本研究の RQ1 に対し，学習データの日齢に近い日齢の画像が入力された場合のほうが，学習データの日齢と異なる日齢の画像が入力された場合に比べ正解率が向上することを示した．

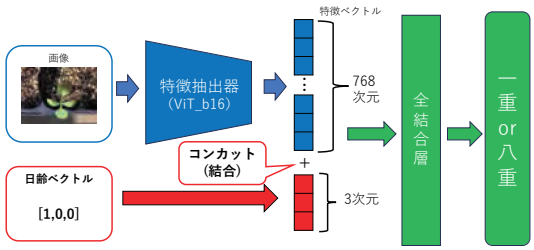


図 2　日齢ベクトルの入力方法
Fig. 2　Method for inputting the seedling age vector

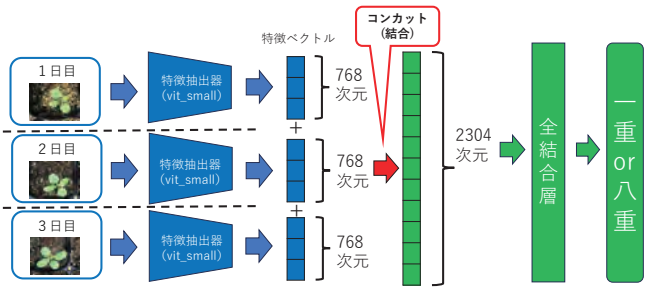


図 3　3 時系列画像の入力方法
Fig. 3　Method for inputting three time-series images

表 1　日齢の不一致による影響の検証結果
Table 1　Verification results of the effect of seedling age mismatch on classification performance

train データ	test データ	Accuracy (%)	ROC-AUC	PR-AUC
MiLIC (1 日目)	農業振興センター	80.2	0.828	0.927
MiLIC (2 日目)		76.3	0.837	0.934
MiLIC (3 日目)		72.3	0.758	0.896

表 2　日齢ベクトルによる影響結果
Table 2　Results of the effects of the seedling age vector

日齢	日齢情報	Accuracy(%)	ROC-AUC	PR-AUC
不一致	無	84.0	0.878	0.900
	有	82.3	0.863	0.888
一致	無	89.9	0.952	0.962
	有	92.2	0.961	0.974

表 3　疑似的日齢情報の検証結果
Table 3　Verification results of the pseudo seedling age information

日齢	入力	Accuracy (%)	ROC-AUC	PR-AUC
不一致	混合 (日齢情報なし)	84.0	0.878	0.901
	1 時系列	81.1	0.853	0.877
	2 時系列	82.4	0.896	0.908
	3 時系列	83.4	0.909	0.932
一致	混合 (日齢情報なし)	89.9	0.952	0.962
	1 時系列	88.7	0.921	0.947
	2 時系列	90.5	0.941	0.964
	3 時系列	91.3	0.959	0.970

3.2 RQ2：日齢ベクトルの導入は分類性能を向上させるか，また，その効果は train/test 間の日齢が一致する場合と一致しない場合のどちらで有効に機能するか

画像特徴量に加えて日齢を数値ベクトルとして入力することで，分類性能の向上を図った．日齢を表すベクトルを日齢ベクトルと呼ぶ．たとえば，撮影期間の 3 日間のうち撮影日が 1 日目に該当する苗には [1, 0, 0] という 3 次元ベクトルを対応させた．図 2 に示すように，ViT-B/16 の先頭クラストークンから得られた 768 次元ベクトルと日齢ベクトルを結合し，全結合層 1 層を経て八重・一重の 2 値分類を行った．

本検証では，train/test 間の日齢が一致しない条件（パターン 1）と一致する条件（パターン 2）の 2 パターンで実験を行った．

パターン 1（日齢不一致）：MiLIC データセットのグループ A（八重：408 枚，一重：291 枚）を train データ，グループ B（八重：396 枚，一重：255 枚）を test データとして用いた．グループ A とグループ B は同一品種であるが日齢が異なるため，train/test 間で日齢が一致しない条件に相当する．

パターン 2（日齢一致）：MiLIC データセットのグループ A を分割し，八重：204 枚，一重：146 枚を train データ，八重：204 枚，一重：147 枚を test データとして用いた．同一グループ内での分割であるため，train/test 間で日齢が一致する条件に相当する．

検証結果を**表 2** に示す．パターン 1（日齢不一致）では日齢ベクトルなしの方が分類性能は高い（正解率 84.0%）一方，パターン 2（日齢一致）では日齢ベクトルありの方が高い（正解率 92.2%）結果となった．ROC-AUC, PR-AUC の各指標についても同様の傾向が見られた．これは，鑑別対象苗の日齢が既知であり，train/test 間の日齢が一致する条件下においてのみ，日齢ベクトルが分類性能の向上に寄与することを示唆している．また，パターン 2 における日齢ベクトルありの結果 (ROC-AUC : 0.961, PR-AUC : 0.974) は，先行研究[1)]に記載のモデル (ROC-AUC : 0.842, PR-AUC : 0.878) と比べて大きな性能向上が見られた．

本研究の RQ2 に対し，日齢ベクトルの導入は，学習データの日齢と同じ日齢の画像が入力された場合において，効果的であるということを示した．学習データの日齢と異なる日齢の画像が入力された場合では，日齢情報の効果が小さくなることが示された．

3.3 RQ3：日齢ベクトルを用いない場合において，時系列画像の導入は日齢ベクトルの代替手段として分類性能の向上に寄与するか

MiLIC データセットには各グループにつき 3 日分の画像が存在することを利用し，複数日の画像を同時に入力することで擬似的な日齢情報をモデルに与えることを試みた．**図 3** に示すように，モデルを並列構成とし，撮影日が 1 日目から 3 日目それぞれの画像を個別の ViT に入力してセットとして処理する方式を採用した．比較のため，1 時系列および 2 時系列についても同様に検証した．本検証では，train/test 間の日齢が一致しない条件（パターン 1）と一致する条件（パターン 2）の 2 パターンで実験を行った．

検証結果を**表 3** に示す．1・2・3 時系列の比較において，3 時系列画像を用いた場合が最も高い分類性能を示した（正解率 91.3%）．train/test 間の日齢が一致しない条件下で行った場合，疑似的日齢情報を考慮しない場合が最も性能が良い結果となった（正解率 84.0%）．その一方で，ROC-AUC や PR-AUC といった二値分類モデルの性能評価指標においては，複数日にわたる時系列画像を活用した場合が最も良い結果となった（ROC-AUC 0.909, PR-AUC 0.932）．train/test 間の日齢が一致する条件下で行った場合では，複数日にわたる時系列画像を活用することで，活用しない場合に比べて分類性能が向上する可能性が示された（正解率 91.3%，ROC-AUC 0.959, PR-AUC 0.970）．

本研究の RQ3 に対し，時系列画像の導入は日齢ベクトルの代替手段として分類性能の向上に寄与することを示した．

4. ま と め

本研究では，八重咲きと一重咲きは同日齢でも苗の大きさが異なるという特性を踏まえ，学習データと鑑別対象苗の日齢のずれが性能低下を引き起こすという仮説を立てた．この仮説に基づき，画像情報に加えて日齢に関する時間的情報を活用することで，生育段階の個体差を考慮したより安定した分類手法の構築を試みた．

実験の結果，train/test 間の日齢差が大きいほど分類性能が低下することが確認され，仮説を支持する結果が得られた（3.1 節）．また，鑑別対象苗の日齢が既知である場合には日齢ベクトルの導入が有効であり，先行研究[1)]のモデルを大きく上回る分類性能が得られた（3.2 節）．さらに，日齢が不明な状況においても，3 時系列画像を活用することで分類性能の向上が期待できることが示された（3.3 節）．

参 考 文 献

1) K. Machida, S. Asukaa, T. Osaka, F. Watanabe, H. Nakajo : “Matthiola Incana Classification using CNN and Class Activation Maps for Model Understanding,” Proc. of 2024 Twelfth International Symposium on Computing and Networking Workshops (CANDARW), pp. 119–124 (2024).

2) 田中 晟雅, 林 角郎 : ストックにおける八重率及び子葉の大きさの品種間差，並びに鑑別後の八重率の推定, 千葉県暖地園芸試験場研究報告 (11), 20-26 (1980).

(2026 年 3 月 1 日 受付)
(2026 年 5 月 31 日 再受付)



堀　越　雄　太
2026 年 兵庫県立大学工学部卒業．在学中はストックの鑑別に関する研究に従事．



奥　田　萌　莉
2024 年 4 月 兵庫県立大学大学院工学研究科助教．2026 年 3 月 兵庫県立大学大学院情報科学研究科博士後期課程修了．2026 年 8 月 理化学研究所生命機能科学研究センター特別研究員．主に機械学習，コンピュータビジョンの研究に従事．博士（情報科学）．



日　浦　慎　作（正会員）
1997 年 大阪大学大学院基礎工学研究科博士後期課程修了．同年京都大学リサーチアソシエイト，1999 年 大阪大学大学院基礎工学研究科助手，2003 年 同助教授．2010 年 広島市立大学大学院情報科学研究科教授，2019 年 兵庫県立大学大学院工学研究科教授．2008-2009 年 マサチューセッツ工科大学メディアラボ客員准教授．三次元空間の画像計測と反射現象・表面質感の解析，コンピュテーショナルフォトグラフィ等の研究に従事．博士（工学）．

IIEEJ ショートペーパー -- ビジュアルコンピューティング特集--

小規模データセットにおけるニューラルネットワークと Hu モーメントを用いた小松菜の日齢推定

奥田萌莉† 日浦慎作†(正会員)

†兵庫県立大学

Age Prediction of Komatsuna Using Hu Moments with Neural Networks for Small Datasets

Moeri OKUDA† , Shinsaku HIURA†(*Member*)

†University of Hyogo

〈あらまし〉 近年，ニューラルネットワーク（NN）などの大規模モデルの利用が増加している．NN を用いる際には，一般的に大量のデータを準備する必要があるが，植物画像を大量に収集することは困難である．そこで本研究では， 少量のデータセットであっても， 幾何学的特徴量と NN を組み合わせることの有効性を示す． 本論文では，コンピュータビジョン分野で従来から用いられてきた幾何学的特徴量である Hu モーメントとニューラルネットワークモデルを組み合わせ， 小松菜の葉齢を予測する手法を提案する． ニューラルネットワークモデルとしては， Vision Transformer (ViT) および ResNet を使用した．実験結果より，Hu モーメントを NN と組み合わせた手法は，データセットサイズに依らず，NN のみを用いた場合と比較して一貫して高い性能を達成することが示された．特に，Hu モーメントと ViT を組み合わせた手法では 96.3%の性能を達成し， 画像特徴のみに依存した ResNet ベースのモデルの 84.2%を上回った．実験結果から，Hu モーメントを NN に組み込むことによる利点は，データセットサイズに関わらず明確に確認できた．

キーワード：小松菜, Hu モーメント, vision transformer

<**Summary**> The use of large-scale models such as neural networks has been increasing in recent years. When using NN, it is generally necessary to prepare a large amount of data. However, collecting a large number of images for plants poses a challenge. Therefore, we demonstrate the effectiveness of combining a geometric feature with NN, even with a small dataset. This paper proposes a method for predicting the age of komatsuna leaves using a combination of a Neural Network model and Hu moments that have traditionally been used as geometric features in computer vision. We used Vision Transformer (ViT) and ResNet as Neural Network models. The results show that Hu moments with NN consistently achieve higher accuracy than only NN across different dataset sizes. Specifically, Hu moments with ViT attained 96.3% accuracy, exceeding the 84.2% of the ResNet-based model relying solely on image features. The experimental results clearly confirm the advantage of incorporating Hu moments into NNs, regardless of the dataset size.

Keywords: komatsuna, Hu-moment, vision transformer

1. はじめに

ディープラーニング，特にニューラルネットワーク（NN）は，コンピュータビジョン分野において標準的な手法となっている．しかし，NN が高い汎化性能を発揮するためには大規模なデータセットが必要であることがよく知られている．先行研究[1)2)]では，学習サンプルが不足している場合，深層モデルは過学習を起こしやすく，有意義なパターンを捉えられないことが示されている．植物科学の分野では，アノテーション作業が労力を要するため，大規模なラベル付きデータセットの取得が困難であり，NN の性能を十分に引き出すことが難しい． この課題に対処するため，本研究では，分野固有の幾何学的特徴量を統合することで，小規模データセットにおいても NN の性能を向上させられるという着想に基づいた手法を提案する．本研究の対象である小松菜は，日本および東アジア諸国で広く栽培されている葉物野菜であり，本研

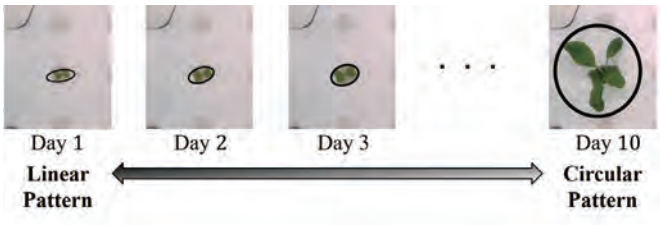


図 1 小松菜成長過程における形状の線形から円形への変化
Fig. 1 Morphological change from a linear to a circular shape during komatsuna growth

究の対象として適している．生育速度が速く，環境耐性が高く，栄養価が高いことが知られている．形態的特性や成長パターンを理解することは，栽培戦略の最適化や収量予測の性能向上において重要である．

Hu モーメントは回転不変性を満たす 7 つの不変量で構成される．これらの不変量を用いることで，対象物の姿勢や撮影条件に依存しない形状特徴を抽出することが可能となる．この性質は，植物画像の取得環境が一定でない場合や，葉の向き・配置が時間とともに変化する場合においても，一貫した特徴量解析を可能にする点で有用である．特に本研究で対象とするような植物の成長過程では，**図 1** に示すように、初期の線形的な形状から，時間経過に伴う放射状・面積的な広がりへの遷移が生じる．このような形状変化は本研究で使用した KOMATSUNA データセット[3]に限らず，多くの葉物植物に共通して観察される一般的な成長特性である．したがって，本研究で扱う問題は小松菜といった特定作物の個別特性の解析というよりも，植物の成長に伴う形状遷移をどのように簡便かつロバストに定量化するかという，より一般的な課題として位置付けられる．この観点から，本研究では，小松菜を具体的な対象として用いながらも，手法自体は他の植物種にも適用可能な汎用的アプローチの構築を目的とする．Hu モーメントは，計算コストが低く実装が容易である一方で，形状の大局的な変化を捉える能力を有しており，3 次元計測[3]のような高コストな手法を用いずに成長遷移を記述する手段として適している．以上より，本研究では Hu モーメントの時間変化に着目し，植物成長に伴う形状遷移を定量的に評価する枠組みを提案する．

さらに，ViT[1]，ResNet152[4]，ViT + Hu，ResNet152 + Hu の 4 つのモデルについて，異なるデータセットサイズにおける分類性能を比較したところ，ViT+Hu の性能がデータセットサイズに関わらずほぼ一貫して最も良い性能であった．データセットサイズが小さい場合の方が，データセットサイズが大きい場合よりも，ViT+Hu の性能が，ViT のみ，ResNet152 のみ，ResNet152+Hu に比べて有効であることが示された．

本研究の貢献は以下の通りである．

- 小松菜画像から葉齢を高性能に推定する手法を提案する．実験結果より，ViT と Hu モーメントを組み合わせた手法は 96.3 %の性能を達成し，画像特徴のみを用いた ResNet152（84.2 %）を大きく上回る性能を示した．
- 小規模データセットにおいて，幾何学的特徴量である Hu モーメントとニューラルネットワークを組み合わせることの有効性を示した．データセットサイズが増加するにつれて分類性能が向上する傾向が確認され，Hu モーメントの導入が性能向上に寄与していることが明らかとなった．
- ViT と Hu モーメントを組み合わせたモデルは， 162 枚の学習データで，ResNet152 が 216 枚のデータを用いた場合と同等の性能を達成した．これは，提案手法が少ないデータ数でも高性能を実現できる高いデータ効率を有していることを示している．

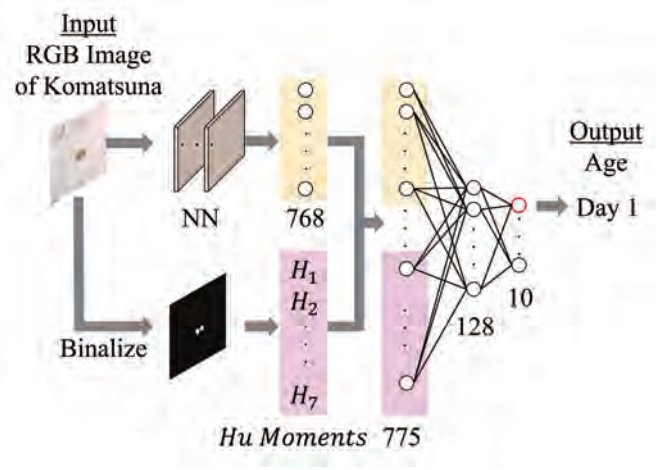


図 2 提案手法の概略図
Fig. 2 Overview of the proposed method

2. 提案手法：Hu モーメントとニューラルネットワークの統合

2.1 問題定義

本研究では，小松菜葉の RGB 画像を入力とし，葉の生育日齢を推定する問題に取り組む．入力画像は高さと幅を持つカラー画像であり，3 つの色チャネル（RGB）から構成される．出力は葉の年齢を表す離散的なラベルで，1 日目から 10 日目までの成長段階に対応する 10 クラスに分類される．**図 2** に本研究で提案する，RGB 画像から小松菜葉の日齢を予測する手法の全体構成を示す．データセットは，画像と対応する日齢ラベルの組から構成されており，全体で N 個のサンプルを含む．本研究の目的は，入力画像から葉の成長段階を正しく分類できるモデルを学習し，分類誤差を最小化することである．モデルは，入力画像がどの成長段階に属するかの確率を推定し，その中で最も確率の高いクラスを予測結果として出力する．

2.2 Hu モーメントを用いたニューラルネットワーク

図 2 に示すように，本手法はニューラルネットワークによる特徴抽出と， Hu モーメントによる形状特徴を組み合わせることで，分類性能の向上を図っている． RGB 画像は

表 1　正解率（Accuracy），適合率（Precision），再現率（Recall），F1 スコア（%）の比較．太字は各指標における最高値を示す．括弧内の数値は訓練データ数．

Table 1　Comparison of accuracy, precision, recall, and F1 score (%). Bold values indicate the highest value for each metric. The number in parentheses indicates the number of training data samples.

Model	Metric	SetA(240)	SetB(216)	SetC(189)	SetD(162)	SetE(135)	SetF(108)	SetG(81)	SetH(54)
ViT	Accuracy	86.6	86.7	73.3	80.0	76.7	73.3	56.7	53.3
	Precision	89.4	87.6	67.1	83.8	82.9	73.9	62.2	54.3
	Recall	88.5	90.0	73.0	80.4	75.6	74.8	62.2	48.3
	F1 Score	85.9	87.2	67.3	80.8	73.1	71.4	57.7	47.5
ResNet152	Accuracy	84.2	82.5	70.8	75.4	72.9	68.3	54.1	49.6
	Precision	86.5	83.1	65.9	78.2	74.5	69.2	58.7	50.2
	Recall	85.9	84.0	72.4	76.0	70.1	68.7	55.6	47.1
	F1 Score	84.7	82.8	66.7	76.9	71.5	67.1	54.9	46.8
ViT + Hu	Accuracy	**96.3**	**93.3**	**86.7**	**83.3**	**83.3**	66.7	**73.7**	**63.3**
	Precision	**97.2**	**92.6**	**86.3**	75.2	85.9	66.6	**71.3**	57.2
	Recall	**97.8**	**95.9**	**89.6**	81.1	89.6	71.4	**71.5**	60.0
	F1 Score	**97.2**	**93.3**	**87.1**	76.1	84.6	61.6	**68.5**	56.8
ResNet152 + Hu	Accuracy	94.8	91.2	84.5	80.7	81.1	**74.2**	69.3	60.4
	Precision	95.3	90.7	83.2	**79.1**	**82.4**	**74.8**	68.1	**59.7**
	Recall	96.1	93.4	86.7	**81.5**	**85.2**	**75.1**	70.2	**61.0**
	F1 Score	95.4	91.6	84.0	**78.3**	**81.7**	**72.8**	67.5	**58.9**

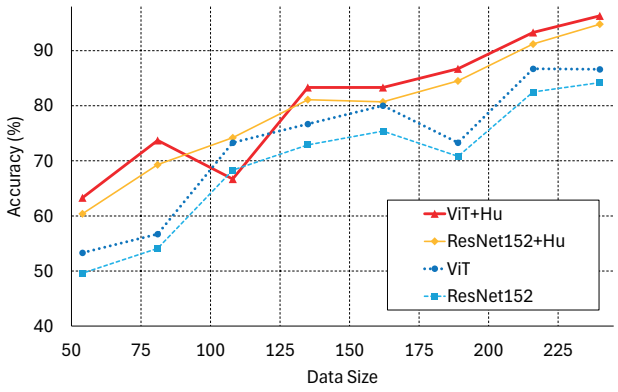


図 3　データサイズに対する各モデルの性能比較
Fig. 3　Comparison of model performance as a function of dataset size

ニューラルネットワークに入力され，画像全体を表現する高次元の特徴ベクトルが抽出される．この特徴ベクトルは，画像の色やテクスチャ，構造的な情報を含む深層特徴を表している．葉の形状に関する特徴を捉えるため，入力画像を二値化処理によって白黒画像に変換する．その後，二値画像から Hu モーメントを計算する．Hu モーメントは，画像の大きさ，回転，位置の違いに対して不変な特徴量であり，葉の形状を効果的に表現することができる．ニューラルネットワークから得られた 768 次元の特徴ベクトルと， 7 次元の Hu モーメントによる形状特徴ベクトルを結合し，最終的な 775 次元の特徴ベクトルを作成する． この特徴ベクトルは全結合層 2 層に入力され，最終的に各成長段階に属する確率が出力される．各全結合層は，ノード数が 128 と 10 である．最も確率の高いクラスが，葉の日齢として予測される． 本手法は， 深層学習による視覚的特徴と， Hu モーメントによる形状特徴の両方を活用することで，葉の日齢推定の性能向上を目指している．

3. 実　　験

本研究では，発芽日から 10 日目までの 30 個体の小松菜を撮影した合計 300 枚の RGB 画像からなるデータセットを使用した．各画像には対応する日齢ラベルが付与されている．

3.1　ニューラルネットワークモデル

本研究では，KOMATSUNA データセットを用いて 2 種類のニューラルネットワークモデルをファインチューニングした．1 つは Vision Transformer（ViT/B/16）であり，もう 1 つは ResNet-152 である．いずれのモデルも，COCO データセットで事前学習された重みを使用している． Vision Transformer は，入力画像をパッチに分割し，自己注意機構を用いて特徴を学習する画像分類モデルである．本研究では，画像全体を表すクラストークンから得られる特徴ベクトルを使用した．一方，ResNet-152 は非常に深い畳み込みニューラルネットワークであり，残差接続を用いることで効率的な学習を可能にしている． これらのモデルから抽出された視覚的特徴は， Hu モーメントと統合され， 小規模データセットにおける分類性能の向上に利用される．

3.2　学習条件および主要パラメータ

実験では，以下のハイパーパラメータを使用した．

最適化手法：Adam 学習率：0.00001 損失関数：クロスエ

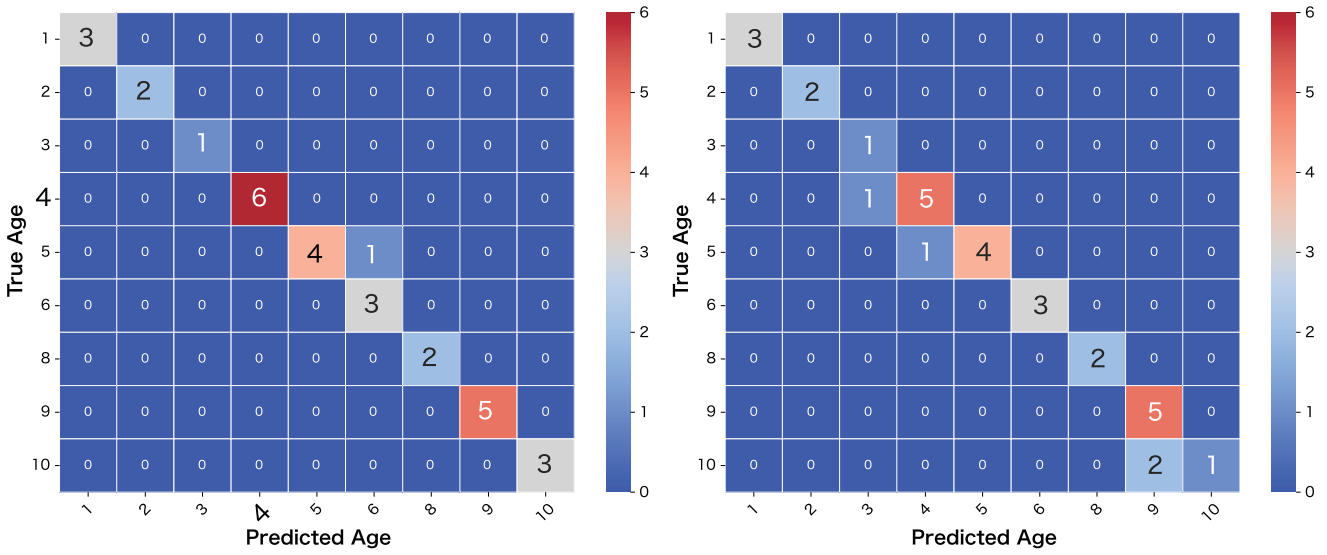


図 4 (左)：ViT と Hu モーメントの組み合わせによる混同行列, (右)：ViT のみによる混同行列

Fig. 4 (Left) Confusion matrix for the combination of ViT and Hu moments; (Right) confusion matrix for ViT only

ントロピー損失関数

3.3 検証方法

本研究では，データセットサイズが限られていることを考慮し，訓練データ数を変化させた複数の実験設定を用いて性能評価を行った．テストデータ（30 枚）および検証データ（30 枚）は固定し，訓練データの画像枚数のみを Set A の 240 枚から Set H の 54 枚まで段階的に減少させた．各 Set の具体的な訓練データ数は**表 1** に示す．

4. 結果と考察

4.1 Hu モーメントがモデル性能に与える影響

表 1 に， Hu モーメントを使用した場合と使用しない場合の各モデルにおける精度，適合率，再現率，F1 スコアを示す．また**図 3** は，この結果のうち精度についてグラフで比較したものである．Hu モーメントを使用しない ViT や ResNet-152 では，比較的大きなデータセットでは一定の性能を示したものの，訓練データ数が減少するにつれて性能が大きく低下した．例えば，ViT 単体では，最大データセットで 86.6%の性能を達成したが，最小データセットでは 53.3%まで低下した．ResNet-152 でも同様の傾向が確認された．一方，Hu モーメントを組み込んだモデルは，すべてのデータセット設定において一貫して高い性能を示した．特に最大データセットでは，ViT と Hu モーメントを組み合わせたモデルが 96.3%という高い性能を達成した．また，最小データセットにおいても，ViT 単体と比較して約 10%の性能向上が確認され，少量データ環境における高いロバスト性が示された．

4.2 訓練データ数が性能に与える影響

訓練データ数とモデル性能の関係を分析した結果，Hu モーメントを使用しないモデルでは，データ数の減少に伴って性能が急激に低下する傾向が見られた． 一方， Hu モーメントを組み込んだモデルでは，訓練データが少ない場合でも比較的安定した性能を維持していた． この傾向は ViT および ResNet-152 の両方で確認されており，形状情報を特徴量として追加することで，モデルの頑健性が向上していることが示唆される．

4.3 誤分類の分析

図 4 左にテストデータにおける ViT+Hu モデルの混同行列を示す．特に 4 日齢の分類では，6 件のサンプルについてすべて正しく分類しており，モデルが特定のクラスに対して非常に高い識別能力を持つことが確認できる．5 日齢の 1 件の誤分類を除き，全てにおいて完璧に分類できていることが確認された．

一方 ViT のみの場合では，図 4 右に示すように，たとえば 4 日齢の分類では，3 日齢に 1 つ誤分類している．それ以外の日齢においても間違えて分類している例があり，これは，視覚的特徴だけでは区別が難しいが，形状変化（Hu モーメント）を加えることでこの曖昧性を低減できたことを示している．

5. 結　　論

本研究では， RGB 画像から小松菜葉の日齢を予測する手法として， Hu モーメントを用いたアプローチの有効性を検証した．その結果，Vision Transformer や ResNet-152 と Hu モーメントを組み合わせた手法は，ニューラルネットワーク単体の手法と比較して，すべてのデータセット設定におい

て高い分類性能を達成した．これらの結果から，複数の特徴表現を統合することで，特に少量データ環境においてモデルの汎化性能と頑健性が向上することが示された．

参考文献

1) A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, N. Houlsby : "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale", arXiv preprint arXiv:2010.11929 (2020)
2) B. Ghorbani, O. Firat, M. Freitag, A. Bapna, M. Krikun, X. Garcia, C. Chelba, C. Cherry : "Scaling Laws for Neural Machine Translation", arXiv preprint arXiv:2109.07740 (2021)
3) H. Uchiyama, S. Sakurai, M. Mishima, D. Arita, T. Okayasu, A. Shimada, R. Taniguchi : "An Easy-to-setup 3d Phenotyping Platform for Komatsuna Dataset", Proc. of the IEEE International Conference on Computer Vision Workshops, pp. 2038–2045 (2017)
4) K. He, X. Zhang, S. Ren, J. Sun : "Deep Residual Learning for Image Recognition", Proc. of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)

(2026 年 1 月 22 日 受付)
(2026 年 5 月 1 日 再受付)



奥 田 萌 莉
2024 年 4 月 兵庫県立大学大学院工学研究科助教．2026 年 3 月 兵庫県立大学大学院情報科学研究科博士後期課程修了．2026 年 8 月 理化学研究所生命機能科学研究センター特別研究員．主に機械学習，コンピュータビジョンの研究に従事．博士（情報科学）．



日 浦 慎 作（正会員）
1997 年 大阪大学大学院基礎工学研究科博士後期課程修了．同年京都大学リサーチアソシエイト，1999 年 大阪大学大学院基礎工学研究科助手，2003 年 同助教授．2010 年 広島市立大学大学院情報科学研究科教授，2019 年 兵庫県立大学大学院工学研究科教授．2008-2009 年 マサチューセッツ工科大学メディアラボ客員准教授．三次元空間の画像計測と反射現象・表面質感の解析，コンピュテーショナルフォトグラフィ等の研究に従事．博士（工学）．

IIEEJ ショートペーパー -- ビジュアルコンピューティング特集--

Gaia 星表を用いた天体写真撮像系の内部標定

松 尾 陽 矢† (正会員)　山 北 彩 未†　前 原 秀 明†† (正会員)

†福岡工業大学大学院工学研究科，††福岡工業大学情報工学部

Intrinsic Calibration of Astrophotography System Using Gaia Catalogue

Haruya MATSUO† (*Member*), Ayami YAMAKITA†, Hideaki MAEHARA†† (*Member*)

† Graduate school of Engineering, Fukuoka Institute of Technology,
†† Faculty of Information Engineering, Fukuoka Institute of Technology

〈**あらまし**〉写真測量分野ではレンズ収差による歪みを補正するために内部パラメータを推定する内部標定が行われ，歪みを補正する．また，天体写真撮像系においても，写真測量用カメラの光学的歪みについて共通する部分が多い．そこで，本研究では，天体写真と Gaia 星表を使用し撮影に用いられた天体写真撮像系の内部標定を試みた．天体写真から各恒星の星像中心を算定したものを星像座標とし，星表データを用いて各恒星の真値とみなす星表座標を作成した．星像座標に内部パラメータを作用させ，星表座標との一致を図り，内部パラメータを推定した．実験結果より，星表座標と星像座標の残差の RMSE が 1 未満の範囲で内部標定ができることが確認できた．

キーワード：天体写真，カメラキャリブレーション，星表

<Summary> In the field of photogrammetry, intrinsic calibration is typically performed to estimate internal parameters, thereby correcting distortions caused by lens aberrations. Since astrophotography systems share similar optical distortion characteristics with photogrammetric cameras, this study attempts the intrinsic calibration of astrophotography systems using an astronomical photograph and Gaia catalogue. We calculated the centroids of stellar images to define image coordinates and generated reference coordinates from the catalog as ground truth. The internal parameters were estimated by transforming the stellar image coordinates to match the star catalog coordinates. Experimental results confirmed that intrinsic calibration was successfully performed, achieving a residual RMSE of less than 1.0 pixel between the star catalog and stellar image coordinates. The remaining residuals are considered to be primarily attributable to the estimation errors in the star centroid calculations.

Keywords: astronomical photograph, camera calibration, star catalogue

1. はじめに

屈折式の天体望遠鏡で得た写真には光学系の歪みが生じるため，精密なコンポジット処理などを行う場合に問題となる．写真測量分野においては，あらかじめ位置決めされているターゲットをカメラで撮影した結果を解析することにより，歪み補正用の内部パラメータを推定する内部標定が行われている．一方，天体写真撮像系と写真測量用のカメラには光学的歪みについて共通する部分が多い．そこで，星表から計算される恒星の座標は歪みの無い天体写真のそれと一致することを利用し，Gaia 星表を用い天体写真撮像系の内部パラメータの推定を試みることとした．

2. 星表を使った望遠鏡の内部標定

2.1 概要

星表とは，天文座標系における恒星の位置や等級、スペクトル型や固有運動などのデータを収録したカタログである [1]．星表から恒星の赤緯・赤経・年周視差などの情報を取得しそれらの情報に基づいて恒星の 3 次元位置を算定し，そこから 2 次元の星表座標を作成する．次に，天体写真において各星像の中心座標を求め，これを星像座標とする．これらの星表座標と星像座標における座標の誤差を最小とするように内部パラメータを決定する．

2.2 星表からの星表座標作成

対象とする天体写真に写っている恒星の赤緯・赤経・年周視差の情報を取得し，各恒星を 3 次元座標系に変換する．なお，恒星までの距離は年周視差に基づいて計算する．次に，投影中心(X_0, Y_0, Z_0)，望遠鏡の姿勢を$(\omega, \varphi, \kappa)$とする条件で共線条件式(1)を適用し，恒星の 2 次元投影座標を作成する [2]．

$$\left.\begin{aligned} x &= -f\frac{m_{11}(X-X_0)+m_{12}(Y-Y_0)+m_{13}(Z-Z_0)}{m_{31}(X-X_0)+m_{32}(Y-Y_0)+m_{33}(Z-Z_0)} \\ y &= -f\frac{m_{21}(X-X_0)+m_{22}(Y-Y_0)+m_{23}(Z-Z_0)}{m_{31}(X-X_0)+m_{32}(Y-Y_0)+m_{33}(Z-Z_0)} \end{aligned}\right\} \quad (1)$$

ここで，(x, y)は画像座標，(X, Y, Z)は恒星の 3 次元座標，$m_{11\sim33}$は$(\omega, \varphi, \kappa)$からなる回転要素，$f$は焦点距離である．

2.3 天体写真から星像座標作成

星像座標とは天体写真に写る各恒星の星像の中心座標であり，星表座標を天体写真上に重ね合わせ，星表座標に対し距離の近い天体写真上の恒星を同一星であるとみなす．天体写真上に重ね合わせた星表座標を**図 1** における切り出し中心とし，星像を含む画像領域を切り出す．切り出した画像とその回転画像間の Zero means Normalized Cross Correlation(ZNCC)を計算し，勾配法によって ZNCC が最大となる切り出し中心座標をサブピクセル単位で算定し，これを星像座標とする．

2.4 内部パラメータの算定方法

星表座標と星像座標を重ね合わせるとレンズ歪みの影響で各座標が一致しない．そこで，星表座標に焦点距離，星像座標にそれ以外の内部パラメータを適用し両座標の一致を図る．

本稿で算定する内部パラメータは，{主点位置$(u_0,\ v_0)$，焦点距離(f)，スケールファクタ$(a,\ c)$，せん断係数(b)，レンズの歪み係数$(K_1,\ K_2,\ K_3,\ P_1,\ P_2)$}である．不一致の評価には Root Mean Square Error(RMSE)を用い， RMSE を最小とすることを内部標定と呼び，内部パラメータの算定方法とする．

式(1)で算定される星表座標はセンサ上の画像座標系における真値と見做す．一方，天体写真から得られる中心推定後の座標は画素座標であり，画像座標に変換する必要がある．デジタルカメラの受光面センサは精密な正方格子であり，せん断歪みを無視できる．しかしながら，ラインジッタの影響が考えられるので，本研究では測量学会で標準的に採用されている式(2)を採用することとした [2)3)]．

$$\left.\begin{aligned} u &= u_0 + ax + by \\ v &= v_0 + cy \end{aligned}\right\} \tag{2}$$

ここで， (u, v)は画素座標，(u_0, v_0)は主点位置である．

星像座標にはレンズ歪みの補正も必要となる．**図 2** は放射方向歪みと接線方向歪みの模式図である [3)]．放射方向歪みとは，レンズの中心からの距離に応じて像が歪む現象である．放射方向歪みには，代表的なものとして糸巻歪みと樽型歪みの 2 種類がある．また，接線方向歪みとは像がレンズの中心から放射線状において接線方向にずれた位置に結像する歪みである．ここで，汎用的に用いられるレンズの歪み係数を使用した放射方向と接線方向の歪みモデルの式(3)を使用し，歪みを補正する [4)]．

$$\left.\begin{aligned} x'' &= x' + \frac{x'}{r}(K_1 r^3 + K_2 r^5 + K_3 r^7) \\ &\quad + P_1(r^2 + 2x'^2) + 2P_2 x' y' \\ y'' &= y' + \frac{y'}{r}(K_1 r^3 + K_2 r^5 + K_3 r^7) \\ &\quad + P_2(r^2 + 2y'^2) + 2P_1 x' y' \end{aligned}\right\} \tag{3}$$

ここで，rはレンズ中心からの距離，$(x'',\ y'')$は歪み補正後の画像座標，$(x',\ y')$は歪み補正前の画像座標，$(K_1,\ K_2,\ K_3)$は放射方向歪み係数，$(P_1,\ P_2)$は接線方向歪み係数である．

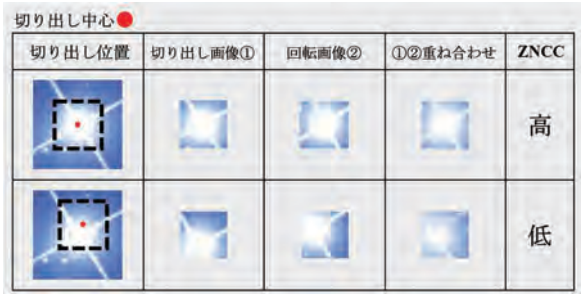
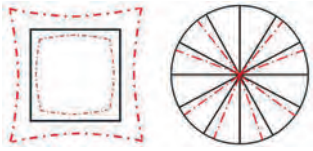


図 1 自己相関に基づく中心座標計算の原理

Fig.1 Autocorrelation based center calculation



(a) Radial distortion (b) Tangential distortion

図 2 レンズの歪みモデル

Fig.2 Lens distortion model



図 3 プレアデス星団(M45)

Fig.3 Pleiades(M45)

3. 実験

実験には，デジタルカメラ Canon EOS 70D に望遠レンズ SIGMA 150-600mm F5-6.3 DG OS HSM Contemporary を装着し，プレアデス星団(M45) (**図 3** 参照）を受光面中央の 1934×1366 (pixel)で記録した天体写真を用いる．星表座標作成のため Gaia Data Release 3(Gaia-DR3)を使用した．Gaia-DR3 は， 欧州宇宙機関が 2013 年に打ち上げた位置天文衛星 Gaia による観測データによって作成された 18 億個以上の天体情報からなる星表である．本実験では，Gaia-DR3 から M45 の恒星 Alcyone の座標を中心として，写真全体に隈なく分布する 8~12 等星の恒星について，120×90arcmin の範囲で約 200 個を取得した．最小化手続きにおいて，内部パラメータの初期値は焦点距離を撮影時のレンズの設定値，レンズの歪み係数について 0，スケールファクタ$(a,\ c)$はカメラの仕様から 243.2，恒星間距離は光年単位であるので $(X, Y, Z) = (0, 0, 0)$とした．計算には勾配法を用い，$(\omega, \varphi, \kappa)$も同時に算定した．

表 1 内部パラメーター一覧

Table 1 Table of intrinsic parameters

パラメータ	初期値	算定値	パラメータ	初期値	算定値
K_1	0.00	-1.45.E-4	$f(mm)$	250	243.00
K_2	0.00	-1.41.E-6	$u_0(pixel)$	0.00	-1.05.E-2
K_3	0.00	-3.88.E-8	$v_0(pixel)$	0.00	1.13.E-2
P_1	0.00	-2.88.E-05	b	0.00	3.73.E-2
P_2	0.00	2.05.E-5			

4. 実験結果と考察

算定した内部パラメータを**表 1** に示す. RMSE は 0.65(pixel)となった. レンズの歪みを含めた諸定数が算定され, RMSE は初期値から 14.97(pixel)減少し 1.00(pixel)未満となったことから，妥当な内部標定が行われたと考える．残差は星像中心の算定誤差によるものと考えられる.

5. むすび

本論文では，Gaia-DR3 を用いた天体撮像系の内部標定のプロセスについて検証した．実験の結果，星表座標と星像座標の RMSE が 1.00(pixel)未満の精度で内部標定が行われたことが確認できた．今後の展開として，再構成型超解像に必要な高精度の画像位置合わせなどへの適応を目的とし，さらなる精度向上を目指す.

参考文献

1) 天文学辞典, 公益社団法人 日本天文学会, 星表, https://astro-dic.jp/star-catalog/ (2026).
2) 日本写真測量学会 編, 三次元画像計測の基礎 -バンドル調整の理論と実践-, 東京電機大学出版局, 近津 博文 監修 (2016).
3) 髙橋 洋二, 近津 博文: “近接デジタル写真測量における民生用デジタルカメラの精度評価”, 写真測量とリモートセンシング, Vol. 49, No. 4, pp.260–268 (2010).
4) D. C. Brown: “Close-Range Camera Calibration”, Photogrammetric Engineering, Vol.37, No. 8, pp.855–866 (1971).

（2026 年 3 月 23 日 受付）
（2026 年 6 月 17 日 再受付）



松 尾 陽 矢 （正会員）
2025 年 福岡工業大学情報工学部システムマネジメント学科卒業. 現在 福岡工業大学大学院工学研究科システムマネジメント専攻修士課程在学中. 主に画像処理について研究. 画像電子学会, 電子情報通信学会学生会員.



山 北 彩 未
2025 年 福岡工業大学情報工学部システムマネジメント学科卒業. 現在 福岡工業大学大学院工学研究科システムマネジメント専攻修士課程在学中. 主に画像処理について研究. 電子情報通信学会 2024 年度成績優秀賞 受賞.



前 原 秀 明 （正会員）
1990 年 宇都宮大学大学院工学研究科修了. 同年 三菱電機㈱入社, 情報技術総合研究所配属. 画像解析に基づく新製品の研究開発に主席研究員として従事. 米国 R&D 100 Awards（2006, 2018）他受賞. 電気学会, 日本写真測量学会会員. 工博. 2020 年より福岡工業大学教授.

IIEEJ 論文

ハイパースペクトル画像を用いた Ki-67 陽性核推定法の開発

石川雅浩[†,††] (正会員)　吉田結真[††]　中野夏澄[††]　黒田真代[††]　茅野秀一[††]　小林直樹[††] (名誉会員)

†近畿大学，††埼玉医科大学

Development of Ki-67 Positive Nuclei Estimation Method Using Hyperspectral Images

Masahiro ISHIKAWA[†] (*Member*), Yuma YOSHIDA[††], Kasumi NAKANO[††], Masayo KURODA[††], Hidekazu KAYANO[††], Naoki KOBAYASHI[††] (*Honorary Member*)

†Kindai University, ††Saitama Medical University

〈**あらまし**〉 病理診断は，組織の状態を正確に把握するために様々な染色を連続切片に施して総合的に診断が行われている．これに対して，HE 染色標本から他の染色を疑似的に作成するデジタル染色が注目されている．特に，HE 染色標本でも推定可能な線維などを特異的に染色する EVG 染色や MT 染色は深層学習を用いることで良好な結果が得られることが報告されている．しかし，免疫染色(Ki-67)のように組織中の抗原抗体反応を染め分ける染色は必ずしもテクスチャに変化が表れないため良好な結果が得られていない．加えて，連続切片を用いて評価するため組織の領域単位の精度評価は可能だが，抗原抗体反応を表す細胞核単位では評価されていない．そこで，本研究では，同一切片から HE 染色標本の HSI と Ki-67 染色標本の RGB 画像を取得し，細胞核単位で陽性核・陰性核の判別が可能かを検討した．ハイパースペクトル画像を用いることで陽性核・陰性核の判別の可能性が示唆されたので報告する．

キーワード：ハイパースペクトル画像, Ki-67，可視光，赤外光，病理標本

<Summary> Pathological diagnosis is performed comprehensively by applying various stains to serial sections in order to accurately diagnose the condition of tissues. In contrast, a digital stain that pseudo-stains other stains from HE-stained specimens has been attracting attention. In particular, EVG staining and MT staining, which specifically stain fibers that can be estimated even in HE-stained specimens, have been reported to yield good results by using deep learning. However, stains that stain for antigen-antibody reactions in tissues, such as the immunohistochemical stain (Ki-67), do not always show changes in texture, and therefore, good results have not been obtained. In addition, although the accuracy of the evaluation can be assessed for each region of the tissue, the evaluation has not been performed for each cell nucleus, which represents the antigen-antibody reaction, because the evaluation is performed using serial sections. In this study, we acquired an HSI of the HE-stained specimen and an RGB image of the Ki-67-stained specimen from the same tissue section and examined whether Ki-67-positive and -negative nuclei could be discriminated on a nucleus-by-nucleus basis. The results suggest that hyperspectral imaging may enable the discrimination of Ki-67-positive and -negative nuclei.

Keywords: hyperspectral imaging, ki-67, visible light, infrared light, pathological slides

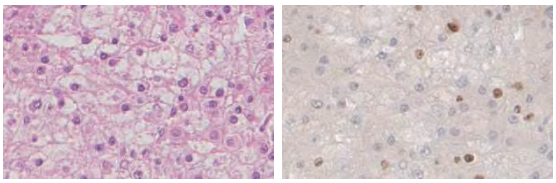
1. はじめに

患者から手術や針生検などで取得した組織を直接観察する病理診断は，確定診断や治療方針の決定など重要な役割を担っている．しかし，臓器によって組織構造が大きく異なるため，臓器ごとの専門医がいるほど病理診断には高度なスキルが求められ，専門的な医学知識に加え病理医の経験に基づいて診断が行われている．これに対して，病理標本全体を撮像可能な仮想顕微鏡と呼ばれる Whole Slide Imaging (WSI)の進展に伴い，コンピュータ診断支援への期待が高まっている [1]．コンピュータ診断支援の初期的検討として，NC 比率（細胞核・細胞質比率）や細胞索の定量化など基本的な組織構造の変化を定量化する方法が提案されている [2),3]．さらに，病理医は診断の際，腫瘍の増殖マーカーとしてがんの悪性度の評価に用いられる Ki-67 などの免疫染色も重要な指標として確認する．Ki-67 は細胞周期の休止期(G0)を除く増殖期で陽性を示す核内抗原であり，これら免疫染色画像などを用いて高度な病理診断を模倣する検討も進められている [4),5]．また，病理診

断の診断効率化の観点からデジタル染色の研究も検討されている．病理診断では，多くの場合最初に HE (Hematoxylin & Eosin)染色と呼ばれる組織構造観察用の染色が選択される．次に，必要に応じて線維を染色する EVG (Elastica van Gieson)染色や MT (Masson trichrome)染色など様々な染色で組織構造や免疫染色を用いて抗原抗体反応などを観察し，複合的に診断が行われる．しかし，これらの追加染色は，特殊な試薬や工程を要するため，コストや時間がかかるだけでなく，貴重な組織検体を消費してしまうという課題がある．デジタル染色は，これら様々な染色法を深層学習などの色変換技術を用いて仮想的に染色する技術である．本研究の目的は，HE 染色標本のハイパースペクトル画像から細胞核単位の抗原抗体反応(Ki-67)の推定を実現することである．検証によりハイパースペクトル画像を用いることで RGB 画像よりも Ki-67 陽性核推定精度が向上するという結果が得られたので報告する．

2. 関連研究

深層学習を用いたデジタル染色に関する研究は近年盛んに行われている．Yang らは C-DNN(cascaded deep neural network)を用いて HE 染色画像から PAS 染色画像を作成する手法を提案している [6]．Haan らは深層学習を用いて HE 染色標本から MT 染色，PAS 染色，銀染色法を提案している [7]．特殊染色は組織構造に合わせて染色液で染めることが目的であるため，HE 染色標本から特殊染色への変換は，組織構造の認識が実現できていれば可能な変換と考えられる．また，CycleGAN などを用いて Ki-67 染色を推定する研究も報告されている [8),9)]．さらに，HE 染色画像から Ki-67 染色標本の特徴を推定し，機械学習の判別に使用することにより，HE 染色標本のみでの判別より精度が向上する手法も提案されている [10)]．しかし，これらの研究は連続切片を用いた外観や間接的な評価であり，本当に抗原抗体反応を推定できているかは明らかではない．また，深層学習に基づく染色変換は組織構造の認識に失敗すると染色に失敗する危険性が常に存在する．これに対して橋本らは [11)]，マルチスペクトル画像を用いて H 単染色標本から HE 染色標本を推定する手法を提案している．マルチスペクトルやハイパースペクトル画像(HSI:Hyper Spectral Imaging)であれば標本を染める染色液の透過光を解析できるため，細胞核や細胞質のテクスチャ情報に頼らずにスペクトルからヘマトキシリンやエオジンの色素量を推定することが可能である [12),13)]．HSI はがん・非がんの判別にも有効なことが提案されている [14)~16)]．また，病理標本の解析に HSI を用いた研究は近年注目されており，国内外で総説論文が報告されている [17)~19)]．本研究の新規性は，Ki-67 染色を推定する際に，HSI を用いることで従来の RGB 画像よりも判別精度が向上することを示した点にある．



(a)HE 染色画像　　(b)Ki-67 染色画像

図 1 同一切片から取得した HE 染色と Ki-67 染色

Fig. 1 HE-stained and Ki-67 stained images obtained from the same .

3. 材料と画像データの取得

3.1 材料

本研究では，研究用病理標本 Tissue Micro Array(TMA)を用いることとした．TMA は異なる患者の組織を円柱状に切り抜いて一定間隔に並べてパラフィンブロックを作成し薄切した標本である．TMA は US Biomax 社のヒト病理標本である T032d(肝臓のがん標本，組織数 25，病変数 12，組織の直径 1.5mm，標本の厚さ 5μm)を用いた．病理標本の撮影には，EBA Japan 社の NH-3(可視光，350～1100nm，752×480 画素)を用いた [20)]．波長分解能は 5nm，データビット数は 10 ビットである．

3.2 病理標本の染色

一般的な病理診断では，連続切片から HE 染色標本と Ki-67 染色標本を作製して相互に参照しながら診断を行う．このため，従来研究の多く [8),9)]は連続切片の Ki-67 から陽性組織領域を推定して精度を算出している．しかし，Ki-67 染色の実現には細胞核単位の推定が必須である．なぜなら，Ki-67 は細胞周期の休止期（G0）以外で陽性を示す核内抗原で細胞の増殖能マーカーとして汎用されるからである．したがって，連続切片による評価では，Ki-67 陽性核を推定できたのか，あるいは単にがん・非がんの判別に成功したのかを判断することが困難である．そこで本研究では，同一切片の HE 染色画像と Ki-67 染色画像を取得することとした．具体的には，HE 染色標本を作製し，HE 染色を脱色せずに Ki-67 染色することとした．得られた結果を**図 1** に示す．脱色せずに染色することで Ki-67 の抗原抗体反応を失うことなく染色に成功していることが確認できる．

3.3 画像データの取得

本研究の具体的なデータ取得は下記の手順で行う．

1. HE 染色標本の WSI 撮影
2. HE 染色標本の HSI を取得
3. Ki-67 染色標本の WSI 撮影

HE 染色標本の WSI 撮影は浜松フォトニクスの Whole Slide Scanner を用いて対物レンズ 20×で撮影し，標本全体の画像サイズは 64,512×26,880 画素，ファイルサイズは 195MB である．画像形式は JPEG で圧縮率は 25.39 である．次に，HE 染色標本の HSI を撮影する．本研究で用いるハイパースペクト

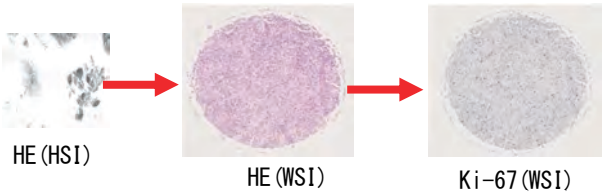


図 2 HE(HSI)と ki-67(WSI)の位置合わせ

Fig. 2 Registration of HE (HSI) and Ki-67 (WSI)

ルカメラは解像度が 752×480 画素であり 1 回の撮影では組織の一部しか撮影できないため XY ステージを制御しながら組織全体を撮影し結合処理を行っている. 最後に標本を Ki-67 染色し WSI 撮影を行う. 標本全体の画像サイズは 66,304×26,880 画素, ファイルサイズは 187MB である. 画像形式は JPEG で圧縮率は 28.5 である.

4. 提案法

4.1 画像データの位置合わせ

本研究では, 深層学習の入力として HE 染色標本のハイパースペクトル画像 (HSI) を, 教師画像として同一切片を Ki-67 染色した標本のホールスライド画像 (WSI) を用いる. そのためには, 異なる顕微鏡 (HS カメラと WS スキャナ) で撮影され, 解像度や色調が全く異なる両画像間の精密な位置合わせが不可欠である. しかし, HE 染色 HSI と Ki-67 染色 WSI を直接対応付けることは極めて困難である. そこで本研究では, 図 2 に示すように 2 段階の位置合わせ手法を提案する. まず, 画像サイズが比較的小さい HE 染色 HSI と, 同じ標本から取得した HE 染色 WSI との位置合わせを行う. これらは同一標本の同一染色像であるため, 正規化相互相関を用いることで高精度な位置合わせが可能である. 次に, この HE 染色 WSI を基準として, 同一切片を Ki-67 染色して取得した WSI との位置合わせを行う. 染色状態が異なるため単純な色情報では位置合わせに失敗するが, 両画像に共通する TMA コアの全体的な形状は不変である. そこで, この工程では画像の回転やスケール変化に頑健な特徴量を抽出する SURF (Speeded Up Robust Features)[21]を用い, 組織の輪郭形状に基づいた特徴量マッチングにより位置合わせを実現する. なお, SURF による特徴量の抽出及びマッチングには, MATLAB の Computer Vision Toolbox を使用した.

4.2 細胞核の抽出と識別

Ki-67 染色標本は, 陽性核を茶色, 陰性核を青色に染色することが重要である. そのため, 細胞質の色合いは病理学的には観察されない情報である. そこで本研究では学習時のノイズを削減するために細胞核を抽出し, 細胞核のみのデジタル染色を検討することとした. 細胞核は Ki-67 染色画像から手動で抽出し細胞核を含むように 31×31 画素で切り出すこととした. HE 染色画像も同じ位置で同じ大きさで切り出すこととした. 次に, 抽出した細胞核の Ki-67 染色画像について, 陽

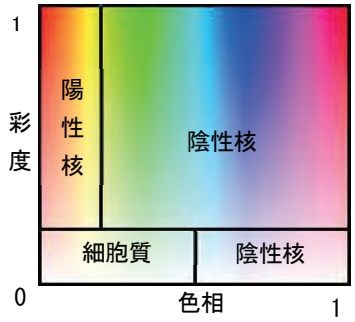
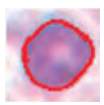


図 3 Ki-67 染色画像における陽性核の決定方法

Fig. 3 Determination method for positive nuclei in Ki-67-stained images

表 1 陰性核と陽性核の HE 染色画像と Ki-67 染色画像

Table 1 HE-stained and Ki-67 stained images of negative and positive nuclei

染色方法	陽性核		陰性核	
HE 染色				
Ki-67 染色				

性核と陰性核の判別を行った. 陽性核は茶色, 陰性核は青色で見えるため HSV 色空間の彩度と色相を用いた分類が適用されることが多い[5]. 本研究でも HSV 色空間による分類法を適用する. HSV 色空間への変換は MATLAB の rgb2hsv 関数を使用し, HSV 色空間は全て[0,1]に正規化されている. 具体的には, H (色相) =0.0 及び 1.0 は赤を示し, 色相は円環構造をもって連続的に扱われる. また, マゼンタ (300° 相当) は H=0.83 付近に位置し, 赤からマゼンタまでの色相も H の連続値として扱われる. HE 染色標本の上に Ki-67 染色を重ねた影響で色合いが若干変化しているため本研究の標本に適した閾値を探索した. HSV 色空間における分類の様子を図 3 に示す.

抽出した細胞核で目視評価し, 閾値を探索したところ, 彩度 0.2 以上で色相 0.2 以下を陽性核画素とした. また, 彩度 0.2 以上で色相 0.21 以上, または彩度 0.2 以下で色相 0.5 以上を陰性核画素とした. いずれの条件にも該当しない画素は, 細胞質などの他の組織とした. 最終的に, 判別した結果から陽性核と判別された画素が多い場合は陽性核, 陰性核と判別された画素が多い場合は陰性核と判別することとした. 切り出して判別した結果の例を表 1 に示す. このとき, HE 染色の最初の陽性核画像において細胞核を赤線で示す. 赤線の内側が細胞核であり, 外側が細胞質である.

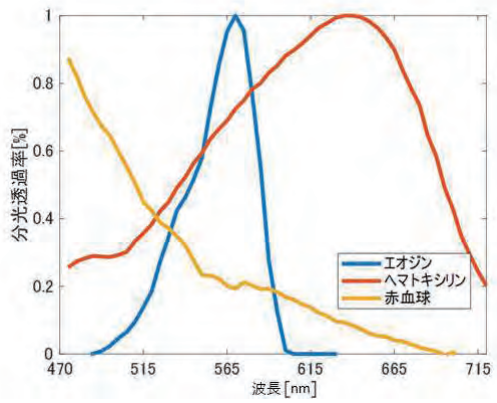


図 4 分光吸収係数
Fig. 4 Spectral absorption coefficient

4.3 細胞核の特徴量算出

本研究では，位置合わせした HE 染色標本の HSI を入力画像として用いる．筆者らは病理標本の染色を考慮した特徴量算出方法を提案しており [15]，本研究でも同様の特徴量算出を行うこととした．提案法は，最初に式(1)のランバートベールの法則から色素量を推定する．

$$\mathbf{a}=\mathbf{XC}+\mathbf{e} \tag{1}$$

ここで観測波長バンド数を N (=51)，色素数を M (=3; ヘマトキシリン，エオジン，赤血球) とすると，$\mathbf{a}$ は各画素の吸光度スペクトルを示す [N×1] のベクトル，$\mathbf{X}$ は各色素の分光吸収係数を列ベクトルとして並べた [N×M] の行列，$\mathbf{C}$ は各色素の量（濃度）を示す [M×1] のベクトル，$\mathbf{e}$ はそれ以外の成分を示す [N×1] のベクトルである．次に，式(2)を用いて分光吸収係数から推定行列 $\mathbf{X}^{+}$を求め，式(3)を用いて色素量を推定する．本研究で用いた分光吸収係数を図 4 に示す．

$$\mathbf{X}^{+}=(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}} \tag{2}$$

$$\dot{\mathbf{C}}=\mathbf{X}^{+}\mathbf{a} \tag{3}$$

最後に推定した色素量から再度吸光度を推定すると染色液以外の成分が抜けた形で吸光度が求められる．本研究では，元の吸光度から再推定した吸光度を減算し差分信号を分離する．この計算で得られた，吸光度(51band)と再推定吸光度(51band)と吸光度と再推定吸光度の差分信号(51band)の合計 153band を特徴量として用いることとした．

4.4 疑似染色方法

本研究では HE 染色分光画像を入力として Ki-67 染色 RGB 画像への疑似染色を行う．色変換については深層学習を用いた様々な手法が提案されており，入力画像に合わせて 2 種の深層学習法を用いることとした．

4.4.1 U-Net

本研究では，ハイパースペクトル画像（HSI）から特定の染色画像（Red, Green, Blue の各バンド）を推定する疑似染色タスクに，深層学習モデルである U-Net を採用した．U-Net は，医用画像のセグメンテーションのために開発された畳み込みニューラルネットワーク（CNN）であり，特徴抽出を行うエンコーダ（収縮パス）と，画素レベルの出力を再構築するデコーダ（拡張パス）から構成される U 字型の対称構造を有する．エンコーダからデコーダへのスキップコネクションにより高解像度の情報を保持し，画素ごとの推定精度を向上させることができるため，本研究に適している．本研究では，U-Net の入力として HSI(153 バンド)を用い，出力として Ki-67 染色 RGB 画像(3 バンド)を用いている．これは，通常の RGB（3 チャネル）の取り扱いと同様に，153 バンドの各画素値をそのまま 153 チャネルの入力として畳み込み層で処理するものである．入力画像は 31×31 ピクセルのパッチとして与えた．推定対象は Red，Green，Blue の各チャネルの連続値であるため，U-Net を回帰タスクに適用するよう構成した．具体的には，MATLAB の unetLayers 関数で生成した標準的な U-Net 構造から分類用の最終層（完全結合層，Softmax 層，セグメンテーション層）を削除し，代わりに出力数が 3 の完全結合層及び回帰層を追加することで，RGB 値を画素単位で推定するように修正した [22]. なお, U-Net の実装には, MATLAB の Deep Learning Toolbox を使用した．

4.4.2 CycleGAN

本研究では，U-Net を用いた直接的な画素回帰に加えて，敵対的生成ネットワーク（Generative Adversarial Network: GAN）の一種である CycleGAN[23]を用いた疑似染色推定も検討した．CycleGAN は，GAN（Generative Adversarial Networks）と呼ばれる敵対的生成ネットワークをベースとした学習手法である．GAN は Generator と Discriminator と呼ばれる 2 つのネットワークで構成される．Generator が Ki-67 染色画像を生成し，Discriminator が Generator の生成した Ki-67 染色画像を評価する．この敵対的学習を繰り返すことで，生成画像の精度が向上することが知られている．これに加えて CycleGAN は，ピクセルの対応関係を学習するのではなく，HE 染色 RGB 画像ドメインと Ki-67 染色 RGB 画像ドメイン間の変換を学習する．本研究では，RGB 画像を入力とした際の比較法として，代表的な手法である CycleGAN を比較のため用いることとした．本手法の U-Net が HSI（153 チャネル）を入力とするのに対し，CycleGAN では HE 染色 WSI（RGB・3 チャネル）を入力とし，Ki-67 染色 WSI（RGB・3 チャネル）を出力として学習する．バンドの取り扱いは，通常の RGB カラー画像と同様である．なお，CycleGAN は J. Zhu らが PyTorch で実装して公開しているプログラムを使用した．

5. 評価実験

5.1 実験環境

本研究において，25 組織の中から陽性核に染まった細胞核の存在する組織を確認したところ 5 症例で確認できた．そこで，陰性核の細胞核が含まれる組織も 5 症例選択して実験を

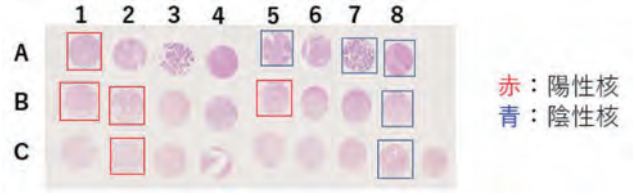


図 5 実験に用いた TMA

Fig. 5 TMA used in the experiment

行うこととした．実際に実験で使用した症例を図 5 に示す．選択した 10 症例の画像から細胞核を抜き出すと，陽性核 283 個，陰性核 1198 個の細胞核画像が得られた．学習時に不均衡データとなるのを避けるため，陰性核画像からランダムに 283 個選択し，陽性核と数を揃えて実験することとした．

5.2 精度評価方法

本研究では，深層学習を用いてデジタル染色した Ki-67 染色画像と対応する正解画像との比較を行う．このため色差を評価する手法として RMSE(Root mean squared error)を用いることとした．加えて，陽性核・陰性核を判別できているか評価するために茶色と青色を閾値処理で判別して，陽性・陰性判別率を算出することとした．算出方法について述べる．

5.2.1 Root mean squared error(RMSE)

本研究では，デジタル染色した画像の精度評価として RMSE を用いることとした．RMSE を式(4)に示す．

$$\mathrm{RMSE} = \sqrt{\frac{1}{NX \times NY}\sum_{x=1}^{NX}\sum_{y=1}^{NY}(D_{(x,y)} - T_{(x,y)})^2} \qquad (4)$$

ここで NX は横方向，NY は縦方向の画素数，D はデジタル染色した画像，T は正解画像である．RMSE の計算は，Red，Green，Blue 画像それぞれで計算する．

5.2.2 陽性核・陰性核判別精度評価

HSI と U-Net を用いたデジタル染色では，茶色や青色に判別できているが，CycleGAN と比べて色差が大きい傾向が見られた．そのため，4.2 節で述べた HSV 色空間での陽性核・陰性核判別を行い，陽性核・陰性核の判別精度を算出することとした．判別精度の算出式を式(5)に示す．

$$\mathrm{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN} \qquad (5)$$

ここで，評価は「判定された核の数」を単位とし，4.2 節で述べた HSV 色空間の分類結果を正解データ（Ground Truth）とする．TP (True Positive; 真陽性) は，正解データにおける「陽性核」 を本手法（U-Net など）が「陽性」と正しく判定した核の数である．TN (True Negative; 真陰性) は，正解データにおける「陰性核」を本手法が「陰性」と正しく判定した核の数，FP (False Positive; 偽陽性) は，正解データにおける「陰性核」を本手法が誤って「陽性」と判定した核の数，そして FN (False Negative; 偽陰性) は，正解データにおける「陽性核」を本手法が誤って「陰性」と判定した核の数を示す．本研究で用いたデータセットは，この正解データに基づき，陽性核 283 個，陰性核 283 個，合計 566 個の細胞核データである．これを学習用（陽性核 255 個，陰性核 255 個）と評価用（陽性 28 個，陰性 28 個）に固定分割して使用した．

表 2 デジタル染色の色差

Table 2 Digital dyeing color differential

手法	Red	Green	Blue	平均色差
RGB(CycleGAN)	21	25	31	26
RGB(U-Net)	28	29	29	29
HSI(U-Net)	27	30	39	32
U-Net,RGB(CycleGAN)	25	31	37	31
U-Net,RGB(CycleGAN) 判別正解画像のみ	21	24	26	23

5.3 実験結果

5.3.1 RMSE によるデジタル染色の精度評価

RGB 画像を用いて U-Net で Ki-67 を推定した結果と RGB 画像を用いて CycleGAN で Ki-67 を推定した結果と HSI を用いて U-Net で Ki-67 を推定した結果画像を作成し色差による評価を行った．結果を表 2 に示す．色差でみると RGB (CycleGAN)が最も色差が小さく平均色差 26 となった．これに対して HIS (U-Net)では平均色差 32 と最も悪い結果となった．これは陰性核・陽性核以外の各周辺の画素も含まれるため背景の推定精度に差が出たものと考えられる．

5.3.2 陽性核・陰性核判別精度

HSV 色空間の閾値処理を行い陰性核・陽性核の判別精度を算出した．結果を表 3 に示す．HSV 色空間で陽性核・陰性核を判別すると CycleGAN では 54%とほとんど判別できていないことが明らかとなった．これに対して HSI を用いた U-Net は，色差は大きいが茶色と青色のどちらに近いかで判別すると 76%の精度で陽性核・陰性核の判別に成功していることがわかった．そこで，HSI から U-Net を用いて作成したデジタル染色結果を HSV 色空間で陽性核または陰性核に判別し，その判別結果に応じて，陽性核用または陰性核用に構築した CycleGAN 染色器を用いて再度デジタル染色することとした．56 枚全ての画像で色差を抽出した結果は，平均色差 31 と RGB (CycleGAN)よりも精度が低下する結果となった．これは，判別精度が低く，染色を誤った際の誤差が非常に大きいためである．加えて，陽性核・陰性核専用の染色器を用意した際の色差を表 4 に示す．陽性核を陽性核，陰性核を陰性核に染色する場合は色差が小さいことがわかる．しかし，陽性核を陰性核，陰性核を陽性核に染色しようとしてしまうと色差が非常に大きくなってしまっている．HSI を用いた陽性核・陰性核の判別が良好に機能し，CycleGAN による染色が成功した

表 3 HSV 色空間での閾値処理による陽性核・陰性核判別精度

Table 3 Accuracy of discriminating positive and negative nuclei by thresholding in HSV color space

手法	精度
RGB(CycleGAN)による判別精度	54%
RGB(U-Net)による判別精度	67%
HSI(U-Net)による判別精度	76%

表 4 陽性核・陰性核用の染色器を用いた色差

Table 4 Color difference using staining models for positive and negative nuclei

		デジタル染色	
		陽性核	陰性核
陽性核	Red	23	27
	Green	25	34
	Blue	28	44
陰性核	Red	21	19
	Green	26	19
	Blue	37	16

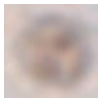
例を**表 5** の成功例①及び②に示す．これらの例が示すように，U-Net による陽性核・陰性核の判別が正しく行われた場合，専用の CycleGAN による染色結果は，正解の Ki-67 染色画像とは色の濃淡が若干異なるものの，細胞核内部の染色される画素と染色されない画素が適切に染め分けられていることが確認できる．一方，陽性核の失敗例①及び陰性核の失敗例①では，事前の判別を誤った場合，正解画像とは大きく異なる色に染色されることが確認できる．

6. 考察

本研究では，ハイパースペクトル画像を用いて細胞核単位の Ki-67 陽性核・陰性核の判別実験を行った．色差による評価では CycleGAN を用いた方が，精度が高かったが，HSV 色空間の閾値処理により茶色に近いか青色に近いかで陽性核・陰性核を判別したところ 76%の精度で陽性核・陰性核が判別できていることがわかった．これに対して，HE 染色標本の RGB 画像を CycleGAN に入力し，Ki-67 染色標本の RGB 画像を推定した場合は判別精度 54%となり，ほとんど判別できていないことがわかる．この結果から，CycleGAN などの敵対的生成ネットワークは，視覚的には非常に精度よく疑似染色するが，細胞核単位で評価すると陽性核・陰性核をほとんど判別できていないことがわかる．これに対して，HSI に U-Net を適用した場合は，CycleGAN と比べて色の模倣度は低いが，陽性核は茶色，陰性核は青色に染めようとしていることが確認できた．このため，HSI は Ki-67 陽性核の推定には重要な情報を含んでいると考えられる．加えて，陽性核・陰性核専用の染色器を用意することで正解画像との色差を小さくできることが明らかとなった．しかし，本研究で得られた 76%の判別精度では陽性核・陰性核の判別を誤ったときの色差が大きくなってしまうため判別精度の向上が必須と考えられる．

表 5 デジタル染色結果

Table 5 Digital staining results

染色方法	HE 染色	Ki-67 染色	デジタル染色
陽性核 成功例①			
陽性核 成功例②			
陽性核 失敗例①			
陰性核 成功例①			
陰性核 成功例②			
陰性核 失敗例①			

7. まとめ

本研究では同一切片に HE 染色と Ki-67 染色を施すことで，1 対 1 関係で HE 染色画像と Ki-67 染色画像を取得可能なことを示した．加えて，今回の検討では，CycleGAN を用いて HE 染色画像から Ki-67 陽性核・陰性核を推定することは困難であることが示唆された．これは全ての抗原抗体反応が細胞核のテクスチャに表れるわけではないためと考えられる．これに対して，HSI を用いることで HE 染色標本画像から Ki-67 染色における陽性核・陰性核を 76%の精度で推定可能なことが明らかとなった．これは，腫瘍の増殖によって生ずる細胞核の変化を吸光度の変化で推定できる可能性があることを示していると考えられる．

謝辞

本研究は JSPS 科研費 JP21K19926 と埼玉医科大学保健医療学部グラント 21-020 の助成を受けたものです．

参考文献

1) 山口雅浩，長橋宏，坂元亨宇，橘口明典，齋藤彰，小林直樹: “定量的病理診断に向けた病理画像解析技術”，電子情報通信学会論文誌 D, Vol. 96, No. 4, pp. 782-790 (2013).

2) M. A. Aziz, H. Kanazawa, Y. Murakami, F. Kimura, M. Yamaguchi, T. Kiyuna, M. Sakamoto: "Enhancing Automatic Classification of Hepatocellular Carcinoma Images through Image Masking, Tissue Changes and Trabecular Features", Journal of Pathology Informatics, Vol. 6, No. 1, p. 26 (2015).
3) M. Ishikawa, Y. Murakami, S. T. Ahi, M. Yamaguchi, N. Kobayashi, T. Kiyuna, M. Sakamoto: "Automatic Quantification of Morphological Features for Hepatic Trabeculae Analysis in Stained Liver Specimens", Journal of Medical Imaging, Vol. 3, No. 2, p. 027502 (2016).
4) F. Negahbani, R. Sabzi, B. Pakniyat Jahromi, D. Firouzabadi, F. Movahedi, M. Kohandel Shirazi, A. Dehghanian: "PathoNet Introduced as A Deep Neural Network Backend for Evaluation of Ki-67 and Tumor-Infiltrating Lymphocytes in Breast Cancer", Scientific Reports, Vol. 11, No. 1, pp. 1–13 (2021).
5) 山下慶子, 喜友名朝春, 山口雅浩: "乳腺 Ki-67 計測定量化のための全自動画像解析システムの開発", Medical Imaging Technology, Vol. 33, No. 3, pp. 148‐155 (2015).
6) X. Yang, B. Bai, Y. Zhang, Y. Li, K. de Haan, T. Liu, A. Ozcan: "Virtual Stain Transfer in Histology via Cascaded Deep Neural Networks", arXiv preprint arXiv:2207.06578 (2022).
7) K. de Haan, Y. Zhang, J. E. Zuckerman, T. Liu, A. E. Sisk, M. F. P. Diaz, K. Jen, A. Nobori, S. Liou, S. Zhang, R. Riahi, Y. Rivenson, W. D. Wallace, A. Ozcan: "Deep Learning-Based Transformation of H&E Stained Tissues into Special Stains", Nature Communications, Vol. 12, p. 4884 (2021).
8) Z. Xu, X. Li, X. Zhu, L. Chen, Y. He, Y. Chen: "Effective Immunohistochemistry Pathology Microscopy Image Generation Using CycleGAN", Frontiers in Molecular Biosciences, Vol. 7, p. 156 (2020).
9) Y. Liu, X. Li, A. Zheng, X. Zhu, S. Liu, M. Hu, Q. Luo, H. Liao, M. Liu, Y. He, Y. Chen: "Predict Ki-67 Positive Cells in H&E-Stained Images Using Deep Learning Independently from IHC-Stained Images", Frontiers in Molecular Biosciences, Vol. 7, p. 166 (2020).
10) 服部英春, 柿下容弓, 坂田晃子, 柳田篤: "CNN を用いた免疫染色特徴量推定による腫瘍組織識別技術", Medical Imaging Technology, Vol. 37, No. 3, pp. 147‐154 (2019).
11) 橋本典明, 村上百合, 山口雅浩, 大山永昭: "マルチスペクトル画像を用いた H 単染色標本からのディジタル H&E 染色とその応用", 電子情報通信学会論文誌 D, Vol. J96-D, No. 4, pp. 791‐801 (2013).
12) T. Abe, Y. Murakami, M. Yamaguchi: "Color Correction of Pathological Images Based on Dye Amount Quantification", Optical Review, Vol. 12, No. 4, pp. 293–300 (2005).
13) 田代充良, 村上百合, 山口雅浩, 小尾高史, 大山永昭, 阿部時也, 八木由香子: "マルチスペクトル病理画像に基づく染色濃度調整処理の実装方法に関する検討", Medical Imaging Technology, Vol. 26, No. 4, pp. 235‐242 (2008).
14) S. Ortega, M. Halicek, H. Fabelo, G. M. Callico, B. Fei: "Hyperspectral and Multispectral Imaging in Digital and Computational Pathology: A Systematic Review [Invited]", Biomedical Optics Express, Vol. 11, No. 6, pp. 3044–3066 (2020).
15) R. Awan, S. Al-Maadeed, R. Al-Saddy: "Using Spectral Imaging for The Analysis of Abnormalities for Colorectal Cancer: When Is It Helpful?", PLOS ONE, Vol. 13, No. 6, p. e0198357 (2018).
16) Y. Takara, N. Manago, H. Saito, Y. Mabuchi, A. Kondoh, T. Fujimori, H. Kuze: "Remote Sensing Applications with NH Hyperspectral Portable Video Camera", Proc. of SPIE – Multispectral, Hyperspectral, and Ultraspectral Remote Sensing Technology, Techniques and Applications IV (SPIE 2012), Vol. 8527, pp. 243–250 (2012).
17) M. Ishikawa, C. Okamoto, K. Shinoda: "Detection of Pancreatic Tumor Cell Nuclei via A Hyperspectral Analysis of Pathological Slides Based on Stain Spectra", Biomedical Optics Express, Vol. 10, No. 11, pp. 5516–5527 (2019).
18) S. Ortega, H. Fabelo, R. Camacho, M. L. Plaza, G. M. Callico, R. Sarmiento: "Detecting Brain Tumor in Pathological Slides Using Hyperspectral Imaging", Biomedical Optics Express, Vol. 9, No. 2, pp. 818–831 (2018).
19) M. Halicek, H. Fabelo, S. Ortega, G. M. Callico, B. Fei: "In-Vivo and Ex-Vivo Tissue Analysis through Hyperspectral Imaging Techniques: Revealing The Invisible Features of Cancer", Cancers, Vol. 11, No. 6, p. 756 (2019).
20) 石川雅浩, 篠田一馬, 小林直樹: "分光画像による病理標本の解析技術", Medical Imaging Technology, Vol. 38, No. 4, pp. 157‐163 (2020).
21) H. Bay, A. Ess, T. Tuytelaars, L. Van Gool: "Speeded-Up Robust Features (SURF)", Computer Vision and Image Understanding, Vol. 110, No. 3, pp. 346–359 (2008).
22) O. Ronneberger, P. Fischer, T. Brox: "U-Net: Convolutional Networks for Biomedical Image Segmentation", Proc. of Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015), pp. 234–241 (2015).
23) J. Zhu, T. Park, P. Isola, A. A. Efros: "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks", Proc. of IEEE International Conference on Computer Vision (ICCV 2017), pp. 2242–2251 (2017).

(2025 年 3 月 28 日 受付)
(2026 年 7 月 23 日 再受付)



石 川　雅 浩　（正会員）

2004 年 新潟大学大学院修士課程修了. 2007 年 同大学大学院博士後期課程修了. 博士（工学）. 東京工業大学産学官連携研究員を経て, 2013 年 埼玉医科大学保健医療学部助教, 2018 年 同講師. 2023 年より近畿大学工学部准教授. 主に医用画像処理, 病理画像認識, コンピュータ支援診断（CAD）に関する研究に従事.



吉 田　結 真

2021 年 埼玉医科大学保健医療学部医用生体工学科卒業. 同年 同大学大学院医学研究科医科学専攻に進学. 現在 同大学院在学中. 画像処理, 病理画像解析, 深層学習に関する研究に従事.



中 野　夏 澄

2021 年 埼玉医科大学保健医療学部臨床工学科卒業. 現在 自治医科大学さいたま医療センター臨床工学部に臨床工学技士として勤務.



黒 田　真 代

1998 年 臨床検査技師免許取得, 同年 埼玉医科大学短期大学に入職後, 主に病理検査学実習を担当し病理組織標本作製法の指導に従事. 2008 年 埼玉医科大学保健医療学部臨床検査学科助教. 2010 年 群馬大学大学院博士前期課程修了. 2013 年 同大学大学院博士後期課程修了. 博士（保健学）. 2014 年 埼玉医科大学保健医療学部講師, 2021 年 同准教授. 現在に至る.



茅 野　秀 一

埼玉医科大学 卒業. 博士・医学
現在 埼玉医科大学保健医療学部学部長, 病理専門医, 細胞診専門医.
日本病理学会功労会員, 日本リンパ腫学会編集委員, 全国医学部国際交流協議会財務理事.



小 林　直 樹　（名誉会員）

1981 年 東京工業大学大学院総合理工学研究科修士課程修了. 2000 年 新潟大学自然科学研究科博士後期課程修了. 博士（工学）. 1981 年 日本電信電話公社（現 NTT）入社, NTT 研究所を経て 2008 年 埼玉医科大学保健医療学部教授. 2024 年 同大学客員教授・名誉教授, 神奈川工科大学客員教授. 現在に至る. 専門分野：医用画像処理, 生体信号処理,画像通信・符号化 ISO/IECJTC1SC29/WG1 国内委員会委員, IEEE, 日本生体医工学会, 日本遠隔医療学会、電子情報通信学会会員.

IIEEJ 論文

改良型 StarGAN を用いたデータ拡張による近代書籍文字認識精度の向上

竹 本 有 紀† 安 田 敏 悠†† 陳 延 偉†† (正会員)

†京都女子大学 ††立命館大学

Character Recognition for Early-Modern Japanese Printed Books with Data Augmentation by Improved StarGAN

Yuki TAKEMOTO†, Toshihisa YASUDA††, Yen-Wei CHEN†† (*Member*)

†Kyoto Women's University, ††Ritsumeikan University

〈**あらまし**〉近代書籍における文字認識では，文字形状の多様性や印刷劣化などの影響により，高精度な認識が困難である．また，高性能な文字認識モデルの学習には大量の学習データが必要であるが，近代書籍文字の学習データの収集には多大なコストを要し，収集できる文字種には限界がある．本論文では，近代書籍文字認識の精度向上を目的として，StarGAN を改良した StarMatchGAN によるデータ拡張手法を提案する．StarGAN とは，複数ドメイン間の画像変換を可能とする生成モデルである．提案手法の StarMatchGAN では，StarGAN におけるサイクル一貫性損失を生成画像と目標画像の L1 損失に置き換えることで，ペア画像を用いた文字骨格変換の精度向上を図る．再構成工程の省略により学習コストを削減し，生成画像を用いたデータ拡張による近代書籍文字認識の精度向上を目指す．実験の結果，提案手法によるデータ拡張が近代書籍文字認識の精度向上に有効であることが確認できる．

キーワード：画像生成, StarGAN, データ拡張, 文字認識, 近代書籍

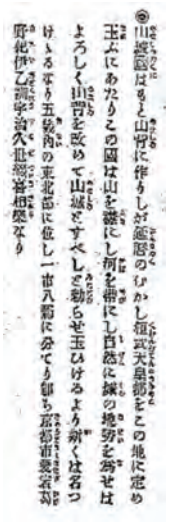
<Summary> Character recognition in early-modern Japanese printed books is challenging due to the diversity of character shapes and degradation in printed materials. In addition, training high-performance recognition models requires a large amount of annotated data; however, collecting training data for early-modern Japanese characters is costly, and the variety of characters that can be collected is limited. In this paper, we propose a data augmentation method based on an improved StarGAN, that called StarMatchGAN, to enhance the accuracy of character recognition for early-modern Japanese printed books. StarGAN is a generative model that enables image transformation across multiple domains. In the proposed method, the cycle-consistency loss in StarGAN is replaced with the L1 loss between the generated image and the target image, which enhances the accuracy of character skeleton transformation using paired images. Furthermore, the reconstruction process is omitted to reduce the training cost. By augmenting the training dataset with generated images, the proposed method aims to improve the accuracy of character recognition for early-modern Japanese printed books. Experimental results confirm that data augmentation using the proposed method is effective in enhancing the accuracy of character recognition for early-modern Japanese printed books.

Keywords: image generation, StarGAN, data augmentation, text recognition, early-modern Japanese printed book

1. はじめに

近年，光学文字認識（Optical Character Recognition，OCR）技術の発達により，画像データに含まれる文字をテキストデータとして抽出することが容易になっている．テキストデータは画像データと比較して文字列検索や音声読み上げなどの機能を利用できるため，データの利便性を向上させることができる．しかしながら，画像データからのテキスト抽出が困難な場合も存在する．その一例が明治から昭和初期に刊行された近代書籍である．

国立国会図書館デジタルコレクション [1]は，国立国会図書館が Web 上で書籍の画像データを公開しているサービスであり，著作権保護期間が満了した書籍を自由に閲覧することができる．さらに，書籍画像から抽出されたテキストデータを閲覧するサービスとして，次世代デジタルライブラリー[2]がある．現代の一般的な書籍は高い精度で文字認識が可能であるが，近代書籍に対する文字認識精度は現代の書籍より低い．



(a)書籍の画像データ

山越國はきと山育に作りしが延暦のひかんね武夫自都をこの地
に定玉ふにあたりこの國は山を發にし河を體にし日然に滅の地
勢を爲せ〓よろしくし音を改めて山城とすべしど動らせ至ひけ
るより斯くはるけよるなり五象内の東北部に位し一市八那に分
てり即ち　嘉締市遂〓〓野紀伊乙訓宇治入世綴喜相樂なり

(b)文字認識結果

図 1　近代書籍の画像データに対する文字認識結果の例

Fig.1 Result of text extraction from an image of an early-modern Japanese printed book

近代書籍の画像データに文字認識を行った結果の例を**図 1** に示す．図 1 (a)は国立国会図書館デジタルコレクションで公開されている書籍 [3]の画像データであり，図 1 (b)は次世代デジタルライブラリーによる文字認識結果である．文字認識結果のうち，正解のテキストデータと異なる文字を赤色で示し，「〓」は文字として認識できなかったことを表す．このように，文字認識結果には誤認識が多く見られる．

近代書籍の文字認識精度が低下する主な原因は，近代書籍文字の特徴が現代の印字文字と大きく異なるためである．近代では，現在使用されていない異字体や旧字体が多く用いられている．また，近代書籍の印刷方法は活版印刷であり，インクのにじみやかすれが生じやすい．さらに，当時は統一されたフォント規格が存在しなかったため，出版者や年代によって近代書籍のフォントが異なる．このような文字形状の多様性や印刷劣化により，現代の印字文字を主な対象として設計された既存の OCR 手法では，近代書籍文字を正確に認識することが困難である．

近代書籍のテキスト化を目的として，近代書籍に特化した文字認識手法が提案されている [4)-6)]．これらの研究では，認識モデルや学習データの改良によって認識精度の向上が試みられているが，実用的な段階には至っていない．その主な原因として，学習データの不足が挙げられる．

近代書籍の印刷に用いられた活字の多くは現存しておらず，学習に用いる文字画像は書籍画像から 1 文字ずつ切り出して収集されている．しかし，書籍における文字の出現頻度は Zipf の法則 [7]に従うため，出現頻度の低い文字の学習データを十分に収集することは困難である [8]．

本論文では，近代書籍文字認識の精度向上を目的として，改良した StarGAN[9]を用いたデータ拡張手法を提案する．StarGAN は複数ドメイン間の画像変換を 1 つのネットワークで学習可能な生成モデルである．近代書籍文字画像と複数の現代印字文字画像から特徴を学習し，現代印字文字を入力として近代書籍文字に類似した特徴を持つ画像を生成する．生成画像を学習データに追加することで，近代書籍文字認識モデルの認識精度の向上を図る．書籍の文字認識ではひらがなおよび漢字が認識対象であるが，本論文では特に構造が複雑で誤認識が生じやすい漢字の認識に注目する．

本論文では，StarGAN に，ペア画像を用いた教師あり学習を導入する．従来の StarGAN はサイクル一貫性損失を用いた教師なし学習により画像変換を学習する．ところが，文字画像はフォントによって文字骨格の形状が異なるため，生成画像の文字骨格が入力の現代印字文字からほとんど変化しないという問題が生じる．そこで，本論文ではサイクル一貫性損失を生成画像と目標画像の L1 損失に置き換えることで，目標となるフォントの文字骨格への変換を学習可能となる．StarGAN の後継モデル（StarGAN v2 ; 後述）に L1 損失を導入した fG-GAN では，生成画像を用いたデータ拡張により近代書籍文字の認識精度向上が確認されている [10]．しかしながら，fG-GAN は 4 つのネットワークで構成されるため計算コストが高い．本論文では 2 つのネットワークで構成され，StarGAN v2 や fG-GAN よりも軽量なモデルである StarGAN に L1 損失を導入することで，計算コストの削減と生成画像の品質向上を両立する．これにより，近代書籍文字認識のさらなる精度向上を目指す．

本研究の主な貢献は以下の通りである．

(1)改良 StarGAN（StarMatchGAN）を用いた近代書籍文字認識のためのデータ拡張手法の提案

(2) StarGAN への教師あり学習の導入による文字骨格変換の精度向上と計算コストの削減

(3) 改良 StarGAN（StarMatchGAN）による生成画像を用いたデータ拡張による近代書籍文字の認識精度向上の実現

本論文の構成は以下の通りである．第 2 章では関連研究について述べる．第 3 章では提案手法について説明する．第 4 章では実験方法および評価結果を示す．第 5 章で本論文のまとめを述べる．

2.　関連研究

2.1　近代書籍文字認識手法

本節では，提案手法に関連する研究として，近代書籍に特化した文字認識手法について述べる．安浪らは畳み込みニューラルネットワーク（Convolutional Neural Network，CNN）を用いた近代書籍文字認識手法を提案している [6]．CNN は文字認識手法として現代印字文字に対して高い認識精度を達成

している．そこで，安浪らはそれぞれ異なるフォントで構成される漢字 2,667 種の近代書籍文字画像 4 セットを学習データとして CNN を学習している．その結果，学習に用いていない漢字 2,667 種の近代書籍文字画像に対する文字認識精度は約 90%である．さらに，安浪らは 21 種類の異なるフォントで構成される漢字 3,011 種の現代印字文字画像を学習データに加えることで，学習に用いていない漢字 2,667 種の近代書籍文字画像に対する認識精度を約 97%まで向上させている．この結果から，認識モデルの性能だけでなく，学習データの量も認識精度に影響を与えることがわかる

しかしながら，近代書籍文字画像の収集は容易ではない．印刷に用いられた活字の多くは現存しておらず，近代書籍文字認識の学習に用いる文字画像は，書籍の画像データから 1 文字ずつ切り出して収集する必要がある．さらに，書籍における文字の出現頻度は Zipf の法則 [7]に従うことが知られている．Zipf の法則とは，出現頻度が k 番目の文字が書籍に出現する回数が全体の k 分の 1 に比例するというものである．実際に，近代書籍における文字の出現頻度を調査した結果，出現頻度は k 分の 1 よりも少なく，k の 2 乗分の 1 に近いことが報告されている [8]．さらに，インクのにじみやかすれが激しい文字画像は学習データに用いることができない．このように，近代書籍文字認識において学習データの不足が大きな課題となっている．

2.2 StarGAN

本節では，本論文で生成モデルとして使用する StarGAN[9]について説明する．従来の生成モデルでは，1 対 1 の画像変換を学習することが一般的である．そのため，複数ドメイン間の画像変換を行う場合には，変換対象となるドメインの数に応じて複数のネットワークを構築する必要がある．これに対し，StarGAN は 1 つのネットワークで複数ドメイン間の画像変換を同時に学習できる生成モデルである．

近代書籍文字認識のためのデータ拡張では，入力となる現代印字文字のフォントを複数用意することで，より多様な特徴を持つ文字画像を学習データに加えることが可能となる．このような複数フォント間の変換を効率的に学習できるため，本論文では StarGAN をデータ拡張のための生成モデルとして採用する．

StarGAN の概要を**図 2** に示す．StarGAN は一般的な GAN と同様に，生成器Gと識別器Dの 2 つのネットワークで構成される．生成器Gは入力画像xとクラスラベルcを入力とし，クラスラベルに対応する特徴を持つ画像$G(x,c)$を生成する．一方，識別器Dは入力画像が本物の画像か偽物の画像かを判断する$D_{starSrc}(x)$とともに，入力画像が属するクラスの分類結果$D_{starCls}(x)$を出力する．このクラスラベルを用いることで，StarGAN の生成器は複数ドメイン間の画像変換を同時に学習することができる．

識別器Dから与えられる分類結果のうち，本物画像に対するドメイン分類損失$\mathcal{L}_{cls}^{r}$は識別器D，生成画像のドメイン分類損失$\mathcal{L}_{cls}^{f}$は生成器Gの学習に用いられる．ドメイン分類損失は入力画像のドメイン分類誤差を表す損失関数であり，それぞれ以下の式(1)，(2)で表される．

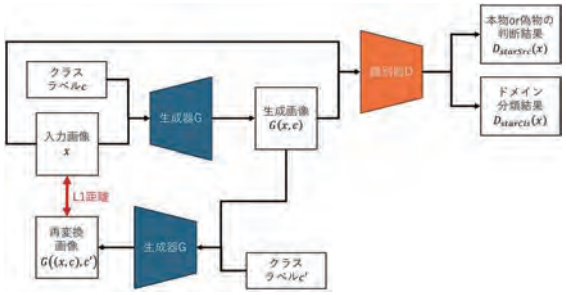


図 2 StarGAN の概要

Fig.2 Outline of StarGAN

$$\mathcal{L}_{starCls}^{r} = \mathbb{E}_{x,c'}[-\log D_{starCls}(c'|x)] \tag{1}$$

$$\mathcal{L}_{starCls}^{f} = \mathbb{E}_{x,c}\left[-\log D_{starCls}\left(c\middle|G(x,c)\right)\right] \tag{2}$$

StarGAN は入力画像と目標画像のペアを用いない教師なし学習を実現するため，入力画像xとクラスラベルcから生成された画像$G(x,c)$を元のドメインに再変換し，入力画像と再変換画像$G\left((x,c),c'\right)$が一致するように学習を行う．このとき用いる損失関数は，以下の式(3)で表されるサイクル一貫性損失である．

$$L_{cyc} = \mathbb{E}_{x,c,c'}\left[||x - G\left((x,c),c'\right)||_1\right] \tag{3}$$

さらに，生成画像を本物の画像に近づけるため，敵対的損失を導入する．この損失関数は以下の式(4)で表される．

$$\begin{aligned}\mathcal{L}_{starAdv} = {} & \mathbb{E}_x[\log D_{starSrc}(x)] \\ & + \mathbb{E}_{x,c}\left[\log\left(1 - D_{starSrc}\left(G(x,c)\right)\right)\right]\end{aligned} \tag{4}$$

StarGAN の識別器と生成器における目的関数はそれぞれ以下の式(5)，(6)で表される．

$$\mathcal{L}_{starD} = -\mathcal{L}_{starAdv} + \lambda_{starCls}\mathcal{L}_{starCls}^{r} \tag{5}$$

$$\mathcal{L}_{starG} = \mathcal{L}_{starAdv} + \lambda_{starCls}\mathcal{L}_{starCls}^{f} + \lambda_{cyc}\mathcal{L}_{cyc} \tag{6}$$

ここで，$\lambda_{starCls}$とλ_{cyc}はそれぞれクラス分類損失とサイクル一貫性損失の重みを決定するハイパーパラメータである．

2.3 fG-GAN を用いた近代書籍文字画像の拡張手法

StarGAN を拡張したモデルとして StarGAN v2[11]が提案されている．StarGAN v2 はスタイルコードの導入により，画像変換における特徴表現の多様性を向上させた画像生成モデルである．fG-GAN[10]は StarGAN v2 に改良を加えることでフォン



図 3　提案手法の概要

Fig.3 Outline of proposal method

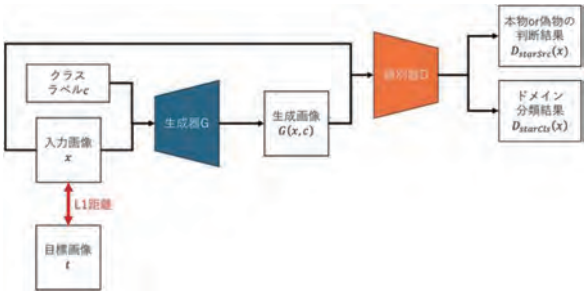


図 4　生成モデルの概要

Fig.4 Outline of generative model

ト変換の性能を向上させている．

fG-GAN における StarGAN v2 からの主な改良点は，ペア画像を用いた教師あり学習の導入である．StarGAN v2 では，2.2 節で述べた StarGAN と同様に，入力画像から生成された画像を再び入力画像のドメインへ再変換し，再変換された画像が入力画像と一致するように学習を行う．この学習により，StarGAN v2 は入力画像と目標画像の対応付けを必要としない教師なし学習による画像変換を実現している．しかしながら，文字画像はフォントによって文字骨格の形状が異なる．そのため，StarGAN v2 にフォント変換を学習させた場合，生成画像の文字骨格が入力画像の文字骨格からほとんど変化しないという問題が生じる．

この問題を解決するため，fG-GAN ではサイクル一貫性損失を生成画像と目標画像の L1 損失に置き換えることで，ペア画像による教師あり学習を導入している．これにより，目標となるフォントの文字骨格への変換を学習することが可能となる．実際に， fG-GAN による生成画像でデータ拡張を行ったところ，近代書籍文字の認識精度が向上することが確認されている．しかしながら，fG-GAN は 4 つのネットワークで構成されており，計算コストが高いという課題がある．そこで，本論文では 2 つのネットワークで構成され比較的軽量な StarGAN に生成画像と目標画像の L1 損失を導入することで，計算コストの削減と生成画像の精度向上の両立を図る．これにより，近代書籍文字認識のさらなる精度向上を目指す．

3.　提案手法

3.1　手法の概要

本論文では，近代書籍文字認識の精度向上を目的として，改良した StarGAN を用いたデータ拡張手法を提案する．以降，本モデルを StarMatchGAN と呼ぶ．提案手法の概要を図 3 に示す．まず，現代印字文字画像を入力として，StarMatchGAN により近代書籍文字の特徴を再現した文字画像を生成する．次に，近代書籍文字認識モデルの学習データである近代書籍文字画像に生成画像を追加することで，学習データを拡張する．その結果として，近代書籍文字認識の精度向上を図る．

3 章の構成は以下の通りである．3.2 節では，近代書籍風文字画像を生成する方法について説明する．3.3 節では，拡張された学習データを用いた近代書籍文字認識の方法を示す．

3.2　近代書籍風文字画像の生成

本節では，データ拡張に用いる近代書籍風文字画像の生成方法について説明する．近代書籍風文字画像の生成モデルの概要を図 4 に示す．本生成モデルは，現代印字文字画像を入力として，近代書籍文字の特徴を再現した文字画像を生成する．モデルは StarGAN と同じネットワーク構造を持つ識別器と生成器によって構成される．

生成モデルにおけるドメイン分類損失は以下の式(7)と(8)で表される．

$$\mathcal{L}^{r}_{starCls} = \mathbb{E}_{x,c'}[-\log D_{cls}(c'|x)] \tag{7}$$

$$\mathcal{L}^{f}_{cls} = \mathbb{E}_{x,c}\left[-\log D_{cls}\left(c\middle|G(x,c)\right)\right] \tag{8}$$

StarMatchGAN では，文字骨格の変換を学習させるため，StarGAN にペア画像を用いた教師あり学習を導入する．具体的には，StarGAN におけるサイクル一貫性損失を，生成画像 $G(x,c)$ と目標画像 t の L1 損失で置き換える．生成モデルにおける L1 損失は以下の式(9)で表される．

$$L_{L1} = \mathbb{E}_{x,c,t}[||G(x,c) - t||] \tag{9}$$

さらに，生成モデルにおける敵対的損失は以下の式(10)で表される．

$$\begin{aligned}\mathcal{L}_{adv} &= \mathbb{E}_{x}[\log D_{src}(x)] \\ &+\mathbb{E}_{x,c}\left[\log\left(1 - D_{src}\left(G(x,c)\right)\right)\right]\end{aligned} \tag{10}$$

生成モデルの識別器と生成器における目的関数は，それぞれ以下の式(11)と(12)で定義される．

$$\mathcal{L}_D = -\mathcal{L}_{adv} + \lambda_{cls}\mathcal{L}_{cls}^{r} \tag{11}$$

$$\mathcal{L}_G = \mathcal{L}_{adv} + \lambda_{cls}\mathcal{L}_{cls}^{f} + \lambda_{L1}\mathcal{L}_{L1} \tag{12}$$

ここで，λ_{cls}とλ_{L1}はそれぞれドメイン分類損失と L1 損失の重みを決定するハイパーパラメータである．

3.3 拡張された学習データを用いた近代書籍文字認識

本節では，3.2 節で生成した近代書籍風文字画像を用いて学習データを拡張し，近代書籍文字認識を行う方法について説明する．

文字認識モデルの学習には，近代書籍文字画像に加えて，3.2 節で生成した近代書籍風文字画像を学習データとして用いる．これにより，学習データの多様性を向上させる．

認識モデルは文字画像を入力とし，入力画像がどの文字クラスに属するかを出力する．提案手法では，認識モデルとして安浪らの手法に基づく CNN [6]を採用する．

近代書籍文字認識では，印刷劣化の影響やフォントの多様性により学習データが不足している．生成された近代書籍風文字画像は，近代書籍文字に見られる文字骨格や印刷劣化の特徴を再現しており，既存の学習データでは不足している文字形状の多様性を補完することができる．これらの特徴を持つ学習データを増加させることができる．これにより，近代書籍文字認識モデルの精度向上が期待される．

4. 評価実験

本章では，提案する StarMatchGAN を用いたデータ拡張手法の近代書籍文字認識に対する有用性を検証するための実験を行う．実験は，近代書籍風文字画像の生成実験と，データ拡張を用いた近代書籍文字認識実験の 2 つに分けて実施する．4.1 節でそれぞれの実験概要を説明し，4.2 節および 4.3 節で生成実験と認識実験の結果および考察を述べる．

4.1 実験概要

本節では，生成実験と認識実験の概要を示す．まず，生成実験の概要について説明する．生成実験では，近代書籍文字画像と複数の現代印字文字画像を学習データとして生成モデルを学習し，近代書籍風文字画像を生成する．生成された文字画像を比較することで，提案手法による生成画像の性能を評価する．比較対象となる生成モデルは StarGAN および fG-GAN である．3 つのモデルで学習時間，生成時間，平均画素一致率，平均ピーク信号対雑音比（Peak Signal-to-Noise Ratio : PSNR），平均構造的類似度（Structural Similarity Index Measure : SSIM）を比較する．

生成画像の品質を評価するため，本実験では PSNR および SSIM を用いる．PSNR は本物の画像と生成画像の平均二乗誤差に基づく指標であり，画素値の再現性を評価する．PSNR の計算式は以下の式(13)で表される．

$$PSNR = 10\log_{10}\frac{L^2}{MSE} \tag{13}$$

Lは最大振幅，MSE（Mean Square Error）は平均二乗誤差であり，以下の式(14)で表される．

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(g_i - t_i)^2 \tag{14}$$

ここで，t_iは本物画像，g_iは生成画像，Nはデータ数を示す．

SSIM は本物画像と生成画像の輝度，コントラスト，構造的類似性に基づいて画像の品質を評価する．本物画像tと生成画像gに対する SSIM の計算式は以下の式(15)で表される．

$$SSIM(t,g) = \frac{(2\mu_t\mu_g + C_1)(2\sigma_{tg} + C_2)}{(\mu_t^2 + \mu_g^2 + C_1)(\sigma_t^2 + \sigma_g^2 + C_2)} \tag{15}$$

ここで，μ_t，μ_gはそれぞれ本物画像tと生成画像gの平均輝度であり，σ_t^2，σ_g^2はそれぞれ本物画像tと生成画像gの分散，σ_{tg}は本物画像tと生成画像gの共分散を表す．C_1とC_2は分母が微小になることによる不安定化を防ぐための定数であり，以下の式(16)，(17)で表される．

$$C_1 = (K_1 L)^2 \tag{16}$$

$$C_2 = (K_2 L)^2 \tag{17}$$

このとき，K_1，K_2は安定化のための小さな定数である．

次に，認識実験の概要について説明する．認識実験では，近代書籍文字画像のみを認識モデルに学習させた場合と，StarMatchGAN による生成画像を含む様々な文字画像によってデータ拡張を行った学習データを学習させた場合で近代書籍文字の認識精度を比較する．データ拡張に用いる文字画像は，現代印字文字画像，StarGAN による生成画像，fG-GAN による生成画像，および StarMatchGAN による生成画像である．

まず，近代書籍文字画像のみを学習データに用いる場合とデータ拡張を行った学習データを用いる場合を比較し，データ拡張の有用性を検証する．さらに，現代印字文字画像と StarMatchGAN による生成画像でデータ拡張する場合を比較することで，生成モデルによる生成画像の有用性を検証する．加えて，StarGAN と fG-GAN による生成画像を用いた場合と StarMatchGAN を比較し，StarMatchGAN の有用性を検証する．

次に，2 つの実験に用いるデータセットについて述べる．本実験では，文字画像は 3 冊の近代書籍と 2 冊の漢和辞典から収集した近代書籍文字画像と，現代印字文字画像を使用する．実験に用いる文字画像の例を**図 5** に示す．

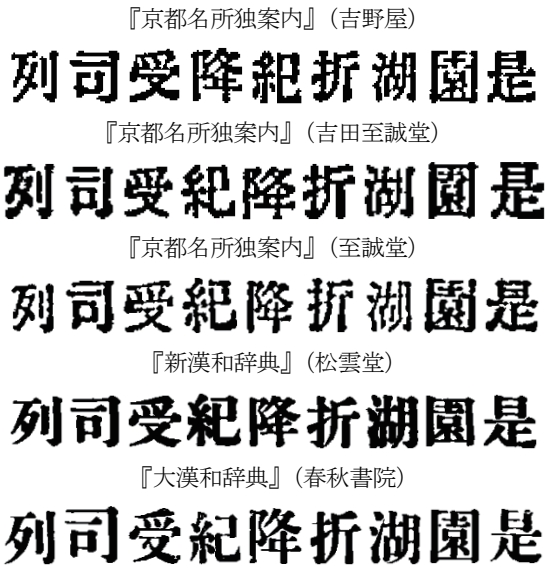


図 5　実験に用いる近代書籍文字画像

Fig.5 Character images of early-modern Japanese printed books used in experiments

表 1　実験に使用する現代印字文字のフォント 20 種

Table 1 20 fonts of modern printed character used in experiments

A1 ゴシック	ヒラギノ明朝 W8
ヒラギノ角ゴシック オールド W9	黎ミン Y40U
HGS ゴシック M	BIZ UD 明朝 Medium
梅ゴシック S5	秀英明朝 M
HGS ゴシック E	HGS 秀英プレゼンス EB
ヒラギノ角ゴシック オールド W6	黎ミン Y40B
ヒラギノ角ゴシック Std	リュウミン U-KL
ヒラギノ丸ゴシック ProN	凸版見出し明朝 EB
小ぶりなゴシック W6	秀英横太明朝 M
UDD デジタル 教科書体 NP-B	UDD デジタル 教科書体 NP-R

3 冊の書籍はいずれも『京都名所独案内』という同一内容の書籍であり，出版者のみが異なる．具体的には，吉野屋，吉田至誠堂，至誠堂の 3 つの出版者による書籍 [12)-14)]を用いる．『京都名所独案内』は京都の観光名所を紹介する書籍であり地名や観光名所などの固有名詞に様々な文字が使用されるため，多様な文字画像を収集することができる．収集の結果，吉野屋版から 1,372 種，吉田至誠堂版から 1,167 種，至誠堂版から 1,055 種の文字を収集している．このうち，吉野屋版の文字画像を生成実験の学習データ，吉田至誠堂版の文字画像を認識実験において学習後に最適なモデルを決定するための検証データ，至誠堂版の文字画像を認識実験のテストデータに用いる．

表 2　実験環境

Table 2 Environment of experiments

CPU	11th Gen Intel® Core (TM)i7-11700 2.50GHz
GPU	NVIDIA GeForce RTX 3080Ti
メモリ	32GB
OS	Windows 10 Pro
使用フレームワーク	Python ver.3.8.10

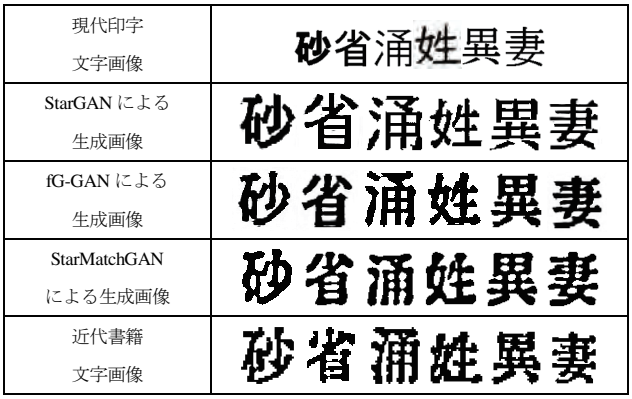


図 6　3 つの生成モデルによって生成された文字画像の例

Fig.6 Samples of character images generated by 3 generative models

認識実験の学習データとして 2 冊の漢和辞典を用いる．具体的には，松雲堂から出版された『新漢和辞典』[15)]と春秋書院から出版された『大漢和辞典』[16)]である．それぞれから 2,547 種および 2.513 種の文字画像を収集している．漢和辞典は漢字を網羅的に収録しているため，1 冊から収集できる文字種は『京都名所独案内』よりも多い．これらの文字画像を学習データとして用いることで，認識モデルは多様な文字形状の特徴を学習することができる．

現代印字文字は生成実験の学習データおよび認識実験のデータ拡張に用いる．使用する現代印字文字のフォントを**表 1** に示す．現代印字文字のフォントは 20 種類用意し，各フォントに対して 2,945 種の文字画像を用いる．これにより，各生成モデルから 2,945 種類の文字画像が 20 通り生成される．近代書籍文字は現代の明朝体に類似した特徴を持つが，本実験では明朝体に加えてゴシック体や教科書体などの複数フォントを用いる．これにより多様な特徴を持つ文字画像を生成する．

本章における実験に用いる計算機環境を**表 2** に示す．また，生成実験における学習時のバッチサイズは 16，学習回数は 30,000 回である．認識実験における学習時のバッチサイズは 256，学習回数は 200 回である．両実験に使用する画像サイズは 64px×64px に統一する．

表 3 生成モデルによる文字画像生成性能の比較

Table 3 Comparison of character image generation performance by deference in generative model

生成方法	StarGAN	fG-GAN	StarMatchGAN（提案手法）
学習時間↓	21h31m	14h20m	**7h30m**
生成時間↓	**43m**	56m	**43m**
平均画素一致率↑	77.23%	80.74%	**81.12%**
平均 PSNR↑	6.5289	7.2802	**7.3678**
平均 SSIM↑	0.3124	0.3773	**0.3841**

表 4 学習データの違いによる認識率の比較

Table 4 Comparison of recognition rate by difference in learning data

学習データ	認識率
データ拡張なし	86.98%
現代印字文字画像を追加	94.69%
StarGAN による生成画像を追加	94.78%
fG-GAN による生成画像を追加	95.83%
StarMatchGAN（提案手法）による生成画像を追加	**97.16%**

4.2 近代書籍風文字画像の生成実験

本節では，近代書籍文字画像および複数の現代印字文字画像を学習データとして学習した生成モデルにより，近代書籍文字画像を生成した結果について述べる．まず，3 つの生成モデルによって生成された文字画像の例を**図 6** に示す．6 種類の現代印字文字画像を入力として，StarGAN，fG-GAN，StarMatchGAN によって生成された文字画像と近代書籍文字画像を比較する．

StarGAN による生成画像では，文字骨格の形状が現代印字文字に近く，近代書籍文字が持つ特徴を再現できていない．一方，fG-GAN および StarMatchGAN による生成画像では，文字骨格の形状が入力画像から変化し，近代書籍文字に類似した特徴が再現されている．特に，「砂」の第六画におけるはねの形状や「省」の文字線の強弱においては，StarMatchGAN の方が近代書籍文字の特徴をより正確に再現している．また，「桶」「姓」では，fG-GAN による生成画像ではインクのにじみが強く，文字の特徴が一部損なわれている．さらに，「異」の第 8，9 および 11 画，「妻」の第 1，3，6 および 8 画における止めの部分では，StarMatchGAN は近代書籍文字に見られるインクのにじみ方を再現しているのに対し，fG-GAN では現代印字文字の特徴が残っている．このように，StarMatchGAN による生成画像は fG-GAN よりも近代書籍文字が持つ特徴をより正確に再現していることが確認できる．

学習データ	データ拡張なし	現代印字文字画像を追加	StarGAN による生成画像を追加	fG-GAN による生成画像を追加	StarMatchGAN（提案手法）による生成画像を追加
文字画像	箇 事	我 殊	東 銀	峨 栖	瓜 鵜
認識結果	隔 町	孜 公	丈 尿	蛾 栢	爪 鶏

図 7 学習データの違いによる誤認識文字の比較

Fig.7 Comparison of recognition error by difference in learning data

次に，3 つの生成モデルにおける学習時間，生成時間，平均画素一致率，平均 PSNR，平均 SSIM の結果を**表 3** に示す．学習時間は StarGAN が 21h31m と最も長く，StarMatchGAN は最も短い上に，StarGAN と比較して約 3 分の 1 まで削減されている．これは，サイクル一貫性損失の計算に必要な再変換画像の生成過程が不要であるためであると考えられる．また，fG-GAN と比較しても，StarMatchGAN の学習時間は約半分である．これは，ネットワーク構造の軽量性による影響と考えられる．さらに，StarGAN と StarMatchGAN の生成時間が fG-GAN よりも短いことから，ネットワーク構造の軽量性は生成時間にも寄与していると考えられる．平均画素一致率，平均 PSNR，平均 SSIM のいずれも StarMatchGAN が最も高い値を示しており，fG-GAN がそれに続き，StarGAN が最も低い結果となっている．以上の結果から，StarMatchGAN は StarGAN に生成画像と目標画像との L1 損失を導入することで，計算コストを大幅に削減しながら fG-GAN と同等以上の生成性能を達成していることが確認できる．

4.3 データ拡張を用いた近代書籍文字の認識実験

本節では，近代書籍文字画像に加えて様々な文字画像を用いて学習データを構成し，認識モデルを学習させた場合の近代書籍文字認識結果について述べる．学習データの違いによる近代書籍文字の認識率の違いを**表 4** に示す．

まず，データ拡張を行わず近代書籍文字画像のみを学習した場合の認識率は 86.98%である．これに対し，現代印字文字画像を用いてデータ拡張を行った場合の認識率は 94.69%となり，データ拡張により認識率が大きく向上することが確認できる．さらに，StarGAN，fG-GAN，StarMatchGAN による生成画像を用いてデータ拡張を行った場合の認識率は，それぞれ 94.78%，95.83%，97.16%である．いずれのデータ拡張手法においても，データ拡張を行わない場合と比較して認識率が向上している．一方，StarGAN による生成画像を用いた場合は，現代印字文字画像を用いた場合と比較して認識率の差が小さい．これは，StarGAN による生成画像では文字骨格の形状が現代印字文字から大きく変化していないことが原因と考えられる．また，近代書籍文字の文字骨格に類似した特徴を再現している fG-GAN と StarMatchGAN を比較すると，

StarMatchGAN の方が高い認識率を示している．これは，前節で述べたように, StarMatchGAN が文字骨格だけでなく文字線の強弱などの特徴も再現しつつ，特徴の欠損が少ない文字画像を生成できているためと考えられる．

次に，各学習データで学習した認識モデルにおける誤認識文字について分析する．学習データの違いによる誤認識文字の比較を**図 7** に示す．データ拡張を行わない場合および現代印字文字を用いたデータ拡張では，文字画像と認識結果の文字が大きく異なる誤認識が多く見られる．両者の文字形状に明確な類似性は見られない．StarGAN による生成画像を用いた場合には，「東」を「丈」と誤認識している．これらの文字は 1 画目の横方向の文字線と左右のはらいが共通している．一方，「銀」を「尿」と誤認識している例では，両者の文字形状に明確な類似性は見られない．これに対し，fG-GAN および StarMatchGAN による生成画像を用いた場合には，部首や文字の概形が近い文字に誤認識する傾向が見られる．この結果から，近代書籍文字に近い特徴を持つ文字画像を用いてデータ拡張を行うことで，文字の特徴をより詳細に学習できることが示唆される．以上の結果より，StarMatchGAN による生成画像を用いたデータ拡張は，近代書籍文字の認識精度向上に有用であることが示される．

5. まとめ

本論文では，近代書籍文字認識の精度向上を目的として，改良した StarGAN に基づく生成モデル StarMatchGAN によるデータ拡張手法を提案した．提案手法では，StarGAN におけるサイクル一貫性損失を生成画像と目標画像との L1 損失に置き換えたことで，ペア画像を用いた教師あり学習を導入している．これにより，近代書籍文字の特徴をより正確に再現した文字画像の生成を可能としている．

提案手法の有用性を検証するため, 2 つの実験を実施した．1 つは近代書籍文字画像と複数の現代印字文字画像を用いて 3 つの生成モデルにより近代書籍文字に類似した特徴を持つ画像を生成し，その性能を比較する生成実験である．もう 1 つは，様々な文字画像を用いたデータ拡張により学習された認識モデルの近代書籍文字に対する認識精度を比較する認識実験である．

生成実験の結果, StarMatchGAN は既存手法である StarGAN および fG-GAN と比較して，学習時間を大幅に削減しながら平均画素一致率，PSNR，SSIM のいずれにおいても優れた生成性能を示した．また，認識実験では，StarMatchGAN による生成画像を用いた場合に最も高い認識率を達成した．さらに，誤認識文字の分析から，近代書籍文字に近い特徴を持つ生成画像を用いることで，文字構造の特徴をより適切に学習できることが示された．

以上の結果より，StarMatchGAN による生成画像を用いたデータ拡張は，近代書籍文字認識の精度向上に有効であることが確認される．今後は，StarMatchGAN のハイパーパラメータを様々に変化させて生成した画像を組み合わせることで，より多様な生成画像を用いたデータ拡張を行い，近代書籍文字認識のさらなる精度向上を目指す．

参考文献

1) “国立国会図書館デジタルコレクション”, http://dl.ndl.go.jp/, [参照: 2026-3-1].
2) “次世代デジタルライブラリ―”, https://lab.ndl.go.jp/service/tsugidigi/, [参照: 2026-3-1].
3) 的場 麗水，京都名所独案内,至誠堂, https://lab.ndl.go.jp/dl/book/904016?keyword= 京 都 名 所 独 案 内 &page=18, [参照: 2026-3-1].
4) K. Kosaka, T. Awazu, Y. Ishikawa, M. Takata, K. Joe: “An Effective and Interactive Training Data Collection Method for Early-Modern Japanese Printed Character Recognition, Proceedings of International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA’15) Vol.II, pp.263– 269 (2015).
5) K. Fujimoto, Y. Ishikawa, M. Takata, K. Joe: “Early-Modern Printed Character Recognition using Ensemble Learning“, Processing of the 2017 International Conference on Parallel and Distributed Processing Technologies and Applications (PDPTA’17), Vol. I, pp.288–294 (2017).
6) S. Yasunami, Y. Takemoto, Y. Ishikawa, M. Takata, K. Joe: “Applying CNNs to Early-Modern Printed Japanese Character Recognition”, Proceedings of the 2019 International Conference on Parallel and Distributed Processing Technologies and Applications (PDPTA’19), pp.189–195 (2019).
7) G. K. Zipf: Human Behavior & The Principle of Least Effort, An Introduction to Human Ecology, Addison-Wesley Press Inc. (1949)
8) 藤田 未希, 竹本 有紀, 石川 由羽, 髙田 雅美, 城 和貴: “近代書籍における低出現頻度文字種の獲得“, 情報処理学会研究報告, Vol.2019-MPS-126, No.6, pp.1–6 (2019).
9) Y. Choi, M. Choi, M. Kim, J.-W. Ha, S. Kim, J. Choo: “StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation”, Computer Vision and Pattern Recognition (CVPR), pp.8789–8797 (2018).
10) 竹本 有紀, 石川 由羽, 髙田 雅美, 城 和貴: “特定の近代書籍出版者における低出現頻度文字種の獲得方法”, 情報処理学会論文誌数理モデル化と応用(TOM), Vol.15, No.3, pp.71–89 (2022).
11) Y. Choi, Y. Uh, J. Yoo, J.-W. Ha: “StarGAN v2 Diverse Image Synthesis for Multiple Domains”, Computer Vision and Pattern Recognition (CVPR), pp.8188– 8197 (2020).
12) 的場麗水，京都名所独案内，吉野屋(1895), https://dl.ndl.go.jp/pid/765652. [参照: 2026-3-1]
13) 的場麗水，京都名所独案内，吉田至誠堂(1903), https://dl.ndl.go.jp/pid/765624. [参照: 2026-3-1]
14) 的場麗水，京都名所独案内,至誠堂(1903), https://dl.ndl.go.jp/pid/904016. [参照: 2026-3-1]
15) 新漢和辞典，松雲堂(1927): https://dl.ndl.go.jp/info:ndljp/pid/1127099. [参照: 2026-3-1]

16) 大漢和辞典, 春秋書院(1925),
https://dl.ndl.go.jp/info:ndljp/pid/1018382. [参照: 2026-3-1]

(2026 年 4 月 6 日 受付)
(2026 年 8 月 6 日 再受付)



竹 本 有 紀
2022 年 奈良女子大学大学院人間文化総合科学研究科博士後期課程複合現象科学専攻修了. 博士 (情報科学) を同大学より取得. 2022 年同大学院人間文化総合科学研究科博士研究員. 2023 年 立命館大学情報理工学部特任助教. 2026 年 京都女子大学データサイエンス学部講師, 現在に至る. 画像処理, 自然言語処理, 深層学習に関する研究に従事.



安 田 敏 悠
2024 年 立命館大学情報理工学部卒業. 2026 年 立命館大学情報理工学研究科情報理工学専攻博士前期課程修了. 現在 京都府公営企業管理事務所に勤務. 画像処理に関する研究に従事.



陳 延 偉 (正会員)
1990 年 3 月 大阪大学大学院工学研究科博士後期課程修了. 工学博士. (財) レーザー技術総合研究所研究員, 琉球大学工学部講師, 助教授, 教授を経て, 2024 年 3 月 立命館大学情報理工学部教授, 現在に至る. コンピュータビジョン, 医用画像解析, 知能計算に関する研究に従事. 電子情報通信学会, 医用画像工学会, IEEE などの会員.

IIEEJ 報告

第 10 回アレキサンダー・ベイン賞 授賞報告

The Report of the Tenth Presentation of Alexander Bain Award

ベイン賞選定委員会

Alexander Bain Award Committee

1. まえがき

『アレキサンダー・ベイン賞』は今から 13 年前に画像電子学会に創設された本学会で最も権威のある表彰である．その創設の経緯については本学会誌通巻 225 号（Vol.42, No.4, pp.509～511）の報告記事をご参照いただきたい．

2. 本年度の授賞

本年度は選定委員会において，第 10 回受賞者として大野邦夫フェローを選出し，令和 8 年 6 月 23 日の本学会総会で発表した．表彰式は，つくば市筑波大学で開催された第 54 回画像電子学会年次大会二日目の令和 8 年 8 月 25 日を予定していたが残念ながら受賞者のご出席が叶わず，年次大会が対面形式のみでオンライン形式を選べなかったこともあり，贈呈者より受賞者のお名前と表彰状を読み上げ，会場参加者から万雷の拍手を得る形となった．今回の受賞者の功績内容（サイテーション）は『電子化文書管理システムの先駆的研究活動・標準化と学会活動推進』である．

写真 1 に授与されたメダルを着用された大野氏のお写真を掲げる．また，以下に大野氏のご略歴の抜粋（**表 1**）と受賞ご挨拶を掲げる．

写真 1 メダルを授与された大野邦夫 氏



表 1 大野邦夫氏ご略歴

1970 年	東京工業大学大学院機械工学専攻修士課程修了
同年～	日本電信電話公社電気通信研究所で交換機部品・通信端末の実用化，リスプマシン ELIS の開発に従事
1985 年～	NTT 新規事業開発室で ELIS の商品化後 NTT-IT に出向し営業 SE を担当
1990 年～	NTT グループ事業推進本部で米国の Interleaf 社との JV 検討後，OMG の CORBA を調査
1993 年～	NTT ヒューマンインタフェース研究所で公衆電話機保守支援システムを実用化
1995 年～	INS エンジニアリング（株）に転籍し，文書ソリューション事業を担当
2000 年～	INS エンジニアリング（株）はドコモシステムズ（株）となり，モバイルソリューション事業を担当
2004 年～	（株）ジャストシステムに転籍し，XML 関連の標準化を担当
2007 年～	職業能力開発総合大学校に転籍し，通信システム工学科教授として研究・教育・職業訓練を担当
2011 年	同大学校 定年退職
同年～	（株）安土 取締役
2015 年～	（株）モナビ IT コンサルティング執行役員
2025 年	同 執行役員 退任
現在	フリーランス

●学会・委員会活動歴：

1994 年～1997 年 OMG Japan SIG 委員
1994 年～2008 年 日本規格協会 オブジェクト指向標準化検討委員会，エージェント技術検討委員会，XML 標準化検討委員会、将来型文書システム（AIDOS）検討委員会 委員
1996 年～2002 年 情報処理学会デジタルドキュメント研究会運営委員； 2002 年～2006 年 同 主査
2004 年～2006 年 W3C Advisory Committee 委員
2007 年～2010 年 画像電子学会メタデータ SG 委員長
2010 年～2020 年 画像電子学会デジタルサイネージとインタラクション研究会 委員長
2020 年～現在 画像電子学会デジタルミュージアム・人文学研究会 副委員長

●受賞歴：

1995 年 INS エンジニアリング社長表彰
2000 年 情報処理学会 山下賞
2001 年 ドコモステムズ社長表彰
2021 年 画像電子学会 フェロー
2022 年 画像電子学会 年次大会貢献賞

大野邦夫氏　受賞ご挨拶

この度は,「電子化文書管理システムの先駆的研究活動・標準化と学会活動推進」という趣旨でアレキサンダー・ベイン賞を賜り，非常に光栄に思います.

電子化文書への興味の発端は, 1980 年代初頭の Xerox PARC におけるダイナブックとそのコンセプトを実現するための Smalltalk でした. 背景のオブジェクト指向技術に関しては, NTT のリスプマシン ELIS の開発を通じて習得しました.

1990 年に，米国 Interleaf 社から NTT グループに画像, 図形, 帳票, 数式等の高度な編集機能を包含する複合文書作成管理システムの日本語版の共同開発が提案され, 私はその技術検討責任者になりました. このシステムは，Lisp 言語でのカスタマイズが可能で，ELIS で習得した技術がそのまま応用できる興味深い製品でした.

その後，INS エンジニアリング（株）が Interleaf 社の代理店になり私も転籍して文書ソリューション事業を担当しました.

その後，複合文書は，オブジェクト指向の標準化団体である OMG で標準化が検討され，W3C により XML をコンポーネントとする Web となり，さらにセマンティック Web を目指してメタデータ技術，オントロジ技術へと発展しました.

画像電子学会でも 2007 年にメタデータ研究グループ（SG）を立ち上げ，XML 複合文書となった高機能な Web のデータ型とアルゴリズム，意味的な基本技術の研究，ビジネスやサービスへの活用を検討しましたが, 直接企業が関わるビジネス分野は難しいことから, 基礎技術と公的・学際的な分野にフォーカスしました.

そのような経緯から本学会の「デジタルサイネージとインタラクション研究会」,「デジタルミュージアム・人文学研究会」を担当させていただき，画像電子学会の基礎技術と公的・学際的な応用分野に少しは貢献できたと思っています.

最後に私のこれまでの様々な活動を支えてくださった皆様に深謝申し上げると共に, 画像電子学会の益々のご発展を祈念致します.

3. むすび

これまでの本表彰式の実施形態を振り返ると，**表 2** に示すように第 1 回から 4 回までは通常の理事会の終了後に行っていた．第 5 回はコロナ禍の下での初の実施となり，オンライン開催となった年次大会に総会及び本賞の表彰式が組み込まれ, 受賞者 2 名には総会の会場にご臨席いただき，オンラインでの年次大会参加者にセレモニーを中継するという形式で行われた．第 6 回から総会と年次大会とが切り離され，本賞の表彰式は年次大会で行うことになり，現在に至っている.

年次大会と総会の分離は定着したと想定されており，今後も多くの参加者が見込める年次大会の場を授賞式に選ぶことになるであろう．受賞者に負担のない範囲で，より盛り上がる実施形式の検討が望まれるともいえる.

表 2　これまでの表彰形式と受賞者一覧

	表彰式	併催	受賞者（敬称略）
第 1 回	2013 年 5 月 17 日	理事会	窪田啓次郎 小林一雄 梶 光雄
第 2 回	2015 年 5 月 15 日	理事会	安田靖彦
第 3 回	2016 年 5 月 13 日	理事会	尾上守夫 南 敏
第 4 回	2018 年 7 月 20 日	理事会	釜江尚彦 富永英義
第 5 回	2021 年 6 月 25 日	オンライン総会・年次大会	西田友是 小宮一三
第 6 回	2022 年 9 月 1 日	ハイブリッド年次大会	小野文孝 松本充司
第 7 回	2023 年 8 月 29 日	ハイブリッド年次大会	羽鳥好律 近藤邦雄
第 8 回	2024 年 8 月 27 日	ハイブリッド年次大会	松木　眞
第 9 回	2025 年 8 月 26 日	年次大会	髙橋 時市郎 河村 尚登
第 10 回	2026 年 8 月 25 日	年次大会	大野 邦夫

IIEEJ 報告

2026 年度 第 54 回 画像電子学会年次大会報告

藤澤　誠（年次大会 実行委員長）

筑波大学

The Report of IIEEJ Media Computing Conference 2026

Makoto FUJISAWA (Chair, IIEEJ Annual Conference)

University of Tsukuba

1.　はじめに

2026 年度第 54 回画像電子学会年次大会は，2026 年 8 月 24 日(月)から 26 日(水)に筑波大学筑波キャンパス春日エリア情報メディアユニオンで開催された（**写真 1**）．本大会は本学会が主催，日本画像学会，日本写真学会，日本印刷学会の協賛によって開催された．また，大会スポンサーとして株式会社リズのご協力を得た. 本大会の参加者は 79 名と昨年度より減少したものの，盛況であった．　その内訳を**表 1** に示す．以下では本年次大会について報告する．

2.　大会概要

2.1 年次大会会場および実施概要

本年度の年次大会の会場となった筑波大学筑波キャンパス春日エリアは，つくばエクスプレスつくば駅から徒歩で約 10 分であり，都内から 1 時間程度という良好なアクセスの地で開催できた．講演発表は主に情報メディアユニオンという建物の 2 つの会議室（メディアホールと共同研究会議室 1）を用いて行い，ポスターセッションはメディアホール前ホワイエで行った．本大会は，現地参加を基本とし，やむを得ない講演者のみオンラインでの発表とした．

本大会では昨年の年次大会 [1]を踏襲し，Conference Track と Journal Track を設定した．Journal Track は大会原稿と共にジャーナル論文を学会誌に投稿いただき，通常と同様の査読を経て採録された論文を画像電子学会誌の特集号に掲載するという企画である．

本大会は例年通り特別講演，通常講演のほか表彰式，アレキサンダー・ベイン賞授賞式から構成された．また，懇親会は 2 日目の午後 6 時より春日エリア内にある「東京バル」で実施した．なお，参加登録と論文投稿は従来通り，イベント管理サービス Peatix と EasyChair をそれぞれ利用した．



写真 1　年次大会会場の情報メディアユニオン

2.2 実行委員会の構成

本大会の実行委員会の体制を**表 2** に示す．本表の他，企画委員会, セミナー委員会のメンバーにもご協力をいただいた．

表 1　参加者の内訳

		人数
参加者数		79
内訳	会員	29
	非会員	10
	会員学生	5
	非会員学生	17
	参与	1
	無料参加	17

表 2　実行委員会体制

実行委員長	藤澤 誠（筑波大学）
副実行委員長	高村 誠之（法政大学）
	北本 英里子（立命館大学）
	山本 奏（NTT 研究所）
プログラム委員長	駒形 英樹（東京情報大学）
プログラム副委員長	内田 理（東海大学）
	竹島 由里子（東京工科大学）
企画セッション委員長	平山 亮（大阪工業大学）
	河合 紀彦（大阪工業大学）
ポスターセッション委員長	佐藤 周平（法政大学）
オンライン運営委員長	藤堂 英樹（拓殖大学）

3. 開催状況

3.1 会議の構成

会議は特別講演 2 件と通常講演からなり，通常講演のセッションは，一般セッション，学生セッション，企画セッション，ポスターセッション，企業セッションから構成される．大会のタイムテーブルを**表 3** に示す．

3.2 通常講演

通常講演のセッション数と発表件数を以下に示す．

[1] 一般セッション　2 セッション　9 件
[2] 学生セッション　4 セッション　18 件
[3] ポスターセッション　1 セッション　5 件
[4] 企画セッション　5 セッション　17 件
(ア) 建築と画像電子技術　7 件
(イ) デジタル人文学と AI　招待 2 件
(ウ) 画像符号化と標準化　2 件
(エ) アクセシブルな電子コミック　招待 2 件
(オ) ドローンが作る映像のいろいろ　4 件
[5] 企業セッション　1 件

3.3 特別講演

特別講演は，2 日目の午後に 2 件実施した．

特別講演 ①では，折り紙の研究で著名な筑波大学の三谷純氏に「曲線折り紙設計手法」と題して，これまでの折り紙に関わる研究についてご講演いただくと共に，会場の参加者も実際に曲線折り紙を折ってみるという体験型の企画も実施していただいた（**写真 2**）．

また，特別講演 ②では，カワムラ・テクノラボの河村尚登氏から「グリーンノイズの画像工学への応用」と題して，河村氏が永年取り組まれているグリーンノイズの原理とその様々な応用に関する研究成果をご紹介いただいた（**写真 3**）．

表 3　タイムテーブル

日付	開始	終了	メディアホール		開始	終了	共同研究会議室 1
8/24（月）	13:10	13:20	オープニング				
	13:25	14:25	[S1] 学生(CG1)		13:25	15:05	[P1] 企画(建築と画像電子技術)
	14:35	15:35	[S2] 学生(画像認識)				
	15:45	17:00	[S3] 学生(画像一般)		15:20	16:45	[P2] 企画(デジタル人文学と AI)
	開始	終了	メディアホール前ホワイエ				
	17:00	18:00	[PO] ポスターセッション				
	開始	終了	メディアホール		開始	終了	共同研究会議室 1
8/25（火）	10:30	11:30	[G1] 一般(CG/VR)		10:50	11:30	[P3] 企画(画像符号化と標準化)
	11:40	12:00	[C1] 企業セッション				
	昼食						
	13:00	14:00	[SP1] 特別講演 1(三谷純 氏)				
	14:10	15:25	[G2] 一般(画像一般)		14:10	15:30	[P4] 企画(アクセシブルな電子コミック)
	15:35	16:35	[SP2] 特別講演 2(河村尚登 氏)				
	16:40	17:50	表彰式				
	18:00	20:00	懇親会				
	開始	終了	メディアホール		開始	終了	共同研究会議室 1
8/26（水）	10:20	11:35	[S4] 学生(CG2)		10:20	11:40	[P5] 企画(ドローンが作る映像のいろいろ)
	11:45	11:55	クロージング				



写真 2　特別講演①の様子



写真 3　特別講演②の様子

3.4 表彰式・ベイン賞授賞式

画像電子学会においてこの 1 年で認められた顕著な業績を表彰する表彰式が，恒例により年次大会の中で挙行された．内容はフェロー称号授与，最優秀論文賞・優秀論文賞，2025 年度研究奨励賞・優秀研究賞，研究会若手奨励賞の表彰であり，その対象者は本年 7 月号の画像電子学会誌をご参照頂きたい．なお，画像電子技術賞は今年度該当無しであった．

また，残念ながら当日のご出席が叶わず，表彰状・メダルの授与や記念のご講演は行われなかったが，アレキサンダー・ベイン賞の授賞式も実施し，受賞された大野邦夫氏と，その表彰題目「電子化文書管理システムの先駆的研究活動・標準化と学会活動推進」をご紹介した．

なお，学生の研究発表のさらなる活性化のために 2024 年度大会で実施された学生発表奨励賞を本大会でも企画し，その表彰をクロージングの中で実施した．学生発表奨励賞の受賞者一覧は本大会 Web ページをご覧いただきたい．

4.　まとめ

第 54 回画像電子学会年次大会に関する実施報告を簡単にとりまとめた．今年度も昨年度に引き続き首都圏での開催となり，アクセス面はよく，一般発表や一般会員の参加者は昨年度並みであったのに対して，他の大会との時期的な関係のせいか，学生発表や学生参加者は昨年度から半減し，予定した集客には至らなかった．ただ，各セッションでは非常に活発な議論がなされ，各発表者には今後の研究の発展につながる成果が得られる良い大会になったと考えている．

最後に，大会実行委員の皆様，事務局の皆様には本大会実施のために企画から運営まで多大なるご協力をいただき，無事に本大会を終えることができました．厚く御礼申し上げます．また，スポンサー企業をはじめとして，ご支援を賜った関係者の皆様にも，この場を借りて厚く御礼申し上げます．

参考文献

1) 岩 崎　慶：“2025 年度 第 53 回 画像電子学会年次大会報告”，画像電子学会誌, Vol.54, No.4, pp.438–440 (2025).



藤澤　誠　（正会員）

2008 年 3 月 静岡大学大学院理工学研究科博士課程修了．2 021 年 2 月より筑波大学図書館情報メディア系准教授．現在に至る．2025 年から 2027 年まで画像電子学会副会長を務める．コンピュータグラフィックス，物理シミュレーションに関する研究に従事．

理事会だより

第 358 回 理事会

1．日時：2026 年 9 月 30 日(水)17:00-18:00

2．場所：オンライン（ZOOM）開催

3．出席者：理事 22 名

総出席者 27 名

【会長】土橋

【副会長】内田、藤澤、山本

【編集】今給黎、董、佐野

【企画】坂東、盧、菅野、澤口、馬場、河合

【財務】前原、茂木学、西村

【総務】藤堂、茂木龍、木全

【技術専門】早瀬、草野、小玉

【監事】金盛 、竹島

＜事務局＞ 大嶽

4．議題・資料

	担当	付番
前回議事録確認	＜内田副会長＞	358-000-01
議 題		
（1）議題 1：第 5 4 回年次大会速報	＜藤澤実行委員長＞	358-001-01
（2）議題 2：第 5 5 回年次大会実行委員長選定について	＜土橋会長＞	358-002-01
（3）議題 3：学会ブランドの見直しに関する準備委員会設置について	＜土橋会長＞	358-003-01
（4）議題 4：高臨場 DF の共催見直しについて	＜坂東企画委員長＞	358-004-01
報 告		
（5）報告 1：編集委員会報告	＜内田編集委員長＞	358-005-01
（6）報告 2：企画委員会報告	＜坂東企画委員長＞	358-006-01
（7）報告 3：画像関連学会連合会　第 12 回合同秋季大会について	＜事務局＞	355-007-01
（8）報告 4：フェロー候補、アレキサンダー・ベイン賞、技術賞の募集について	＜事務局＞	355-008-01
（9）報告 5：法人登記進捗について	＜事務局＞	無し

5. 議事 司会：内田副会長、議事録：藤堂総務理事

＜お知らせ＞

会員各位

画像電子学会

「会員用 Web」への登録実施のお願い

当学会では、現在「会員用 Web（シクミネット）」を開設しており、2026 度より会費の精算が会員用 Web に統一しております。

つきましては、以下の会員用 Web にアクセスいただき、会員登録の作業を実施頂きますよう、よろしくお願い致します。

https://iieej.shikuminet.jp/login/

会員ホームページへのログインの際に、ログイン ID、パスワードがご不明の場合は、下記のお問い合わせ先までメールにてご連絡ください。

《お問い合わせ先》　gazo_iieej@iieej.org

一般社団法人　画像電子学会
〒116－0002
東京都荒川区荒川 3-35-6 ライオンズマンション三河島第二 101 号
TEL:03-5615-2893　FAX:03-5615-2894
E-mail: gazo_iieej@iieej.org

2026年度画像電子学会研究会等予定

研究会等名	開催日	場所	テーマ	締切	記事
ビジュアルコンピューティングワークショップ2026	12月17日(木) - 18日(金)	まるは食堂旅館 (愛知県南知多)	講演/参加登録締切:11/24(火)		
第319回研究会	2月24日（水） （予備日として＋25日（木））	沖縄セルラー	画像一般		
第12回DMH研究会	2月（予定）				
第55回VMA研究会/第21回 視覚・聴覚支援システム研究会	3月（予定）				
第13回DMH研究会	3月（予定）				
第320回研究会-映像表現/芸術科学フォーラム2027	3月（予定）				共催
ドローン研究会	3月（予定）				

*研究会の場合，いずれも「画像一般」を含む　*空欄部は未定

*上記の予定は変更になる場合があります

問合わせ先：〒116-0002 東京都荒川区荒川3-35-4-101
一般社団法人 画像電子学会
TEL:03-5615-2893 FAX:03-5615-2894
E-mail:kikaku@iieej.org
https://www.iieej.org/
https://www.facebook.com/IIEEJ

画像電子学会 第 319 回研究会 in 沖縄 発表募集（2/24（水）-2/25（木）開催）

下記により、画像電子学会 第 319 回研究会の発表論文を募集いたします。ご関係の各位には奮ってご応募ください。今回は現地（沖縄セルラー電話 会議室）とオンラインによるハイブリッド形式で実施します。画像電子学会の第一種研究会では「コメント・サービス」に加え、優れた研究発表を行った若手研究者に対して「研究会若手奨励賞」を授与する予定です。

－記－

■テーマ：

画像一般（色再現、立体映像、CG、VR、アニメーション技術、画像処理、医療画像関係、画像符号化など幅広く扱います。修論、卒論の発表も歓迎します。）

■開催日：

2027 年 2 月 24 日（水）13:00～17:30（予定）、2 月 25 日（木）9:00～13:00（予定）

（*第 2 日目は発表件数によりキャンセルされる可能性があります）

■会場・発表形式：

現地とオンラインのハイブリッド開催、質疑応答含み 1 件当たり 30 分前後

（実施方法や発表時間は、今後の状況により変更になる場合があります。）

■現地での発表・聴講形式：

沖縄セルラー電話株式会社 会議室

沖縄県那覇市松山 1 丁目 2 番 1 号

https://corp.okinawa-cellular.jp/corporate/profile/

発表者は各自の PC にてビデオ会議システム Zoom の画面共有機能により画面を共有しつつ発表を行っていただきます。発表前日までに発表者に Zoom への接続情報を電子メール等にてお送りします。（接続トラブルなどを想定し、念のため、USB メモリ等にて PowerPoint ファイル等を別途ご持参下さい。）

■オンラインでの発表・聴講形式：

発表者は各自の PC にてビデオ会議システム Zoom の画面共有機能により画面を共有しつつ発表と質疑応答を行っていただきます。各自マイク、カメラ、スピーカーを備えた PC をご用意ください。発表前日までに発表者と参加者に Zoom への接続情報を電子メール等にてお送りします。

■申込締切日：2026 年 12 月 14 日（月）

参加申込フォーム: https://forms.gle/RHFBRMRG3gfvQBKFA

よりお申し込みください。申込受付後、原稿の執筆要項と提出方法をお送り致します。

■原稿締切日：2027 年 1 月 25 日（月）
原稿の目安：A4 サイズ　2 段組　2〜4 頁

■参加費（予稿集付）：2,000 円（発表数により一日開催の場合は 1,000 円とします）
参加申込、および、参加費の支払い方法は別途ご案内いたします。

■コメント・サービスについて：
研究会（本学会主催/共催の第一種研究会）では、発表内容がその後学会誌論文としての投稿に発展することを促進するため、発表者の希望に基づいて、当学会が指名した研究者が事前に発表者の最終原稿を読んだ上で、当日発表時に質問やコメントをし、さらに後日、コメントを発表者に送付する「コメント・サービス」を実施しています。
このサービスを活用すると、発表に関してより踏み込んだ議論が期待できますので、今回の研究会での発表内容を将来的に学会誌論文として投稿することを検討されている発表者は、申込締切後に画像電子学会より送付されるアンケートにて「コメント・サービス希望」とご回答ください。なお、本サービスの対象発表数には限りがありますので、最終原稿に基づき選抜させていただくことがあります。あらかじめご了承ください。

■研究会若手奨励賞について：
画像電子学会第一種研究会において優れた研究発表を行った若手研究者に対して「研究会若手奨励賞」を毎年若干名授与します。受賞資格は以下のとおりです。研究会奨励賞の対象者（研究会における発表登壇時において 35 歳以下）は発表申込み時に生年月日をご記入ください。
– 本奨励賞の受賞時において本学会会員（学生会員含む）であること。
– 研究会における発表登壇時において 35 歳以下であること。
– 過去に同賞を受賞したことがないこと。

■お問い合せ先：一般社団法人 画像電子学会
〒116-0002 東京都荒川区荒川 3-35-4　ライオンズマンション三河島第二 101 号室
TEL: 03-5615-2893　FAX: 03-5615-2894　E-mail: kikaku@iieej.org

詳細については学会ホームページで順次掲載いたします。
画像電子学会ホームページ　https://www.iieej.org/

■その他：
当学会では研究会予稿を電子化しており、予稿（電子ファイル）の事前配布を可能とするため、予稿の公開日は研究会開催日の 1 週間前となる予定です。

以上

ビジュアルコンピューティング論文特集号　論文募集

画像電子学会編集委員会

ビジョンとグラフィックスの統合を目指したビジュアルコンピューティング研究委員会（現ビジュアルコンピューティング研究会）は 1993 年に発足し，シンポジウムおよびワークショップを通して活発な活動を行ってきております．そこで本年度も，表記の論文特集号を下記の要領により企画いたしました．採録論文は本学会論文賞，および，西田賞の対象にもなりますので，奮って御応募下さい．

1. キーワード
   - グラフィックスの基礎（レンダリング，モデリング，形状モデル，シミュレーション）
   - グラフィックスシステム（入出力デバイス，GPU，Web／モバイルシステム）
   - グラフィックスの応用（マルチメディア，アニメーション，可視化，医療応用，地理情報処理）
   - デジタルファブリケーション（3D プリンタ，建築応用，デジタルエンジニアリング）
   - エンタテインメント（ゲーム，エージェント・人工知能応用，インタラクション応用)
   - VR 基礎・応用（複合／拡張現実，五感インタフェース，視覚心理・認知モデル）
   - 画像処理（画像の表示・計測・認識等，ビジョンとグラフィックスの融合技術）
   - コンテンツ処理（形状・素材データの検索／編集，異種媒体での表示）

これらに少しでも関連をもつ，より新しいまたより広い分野からの研究成果の投稿を期待しています．

2. 論文の種類と取扱い

「論文」「ショートペーパー」を受け付けます．他の種目については，編集幹事が内容を確認した上で，受け付けるかどうかを決定します．

取り扱いは通常の論文投稿と同一です．なお採録決定が特集号に間に合わなかった場合には，通常の投稿として扱います．

3. 特集号発行　　画像電子学会誌 2027 年 10 月号予定（電子版）

4. 論文投稿締切日　　2027 年 3 月 23 日（火） 必着

5. 論文の投稿方法

論文投稿システムよりご投稿ください．
http://www.editorialmanager.com/iieej/
論文投稿にあたっては，下記の論文投稿の手引きをご一読ください．
https://www.iieej.org/call-for-papers/guide-paper-submission/

6. 問い合わせ先

画像電子学会事務局
〒116-0002 東京都荒川区荒川 3-35-6 ライオンズマンション三河島第二 101 号
TEL: 03-5615-2893　　FAX: 03-615-2894
事務局　福島　hensyu@iieej.org
http://www.iieej.org/（電子投稿）

ゲスト編集委員長　　土橋 宜典（北海道大）
ゲスト編集幹事　　謝 浩然（北陸先端大/早稲田大）　藤堂 英樹（拓殖大）
楽 詠灝（青山学院大）

2017 年 1 月 6 日改訂
2018 年 7 月 6 日改訂
2020 年 10 月 2 日改訂
2024 年 11 月 6 日改訂
2024 年 12 月 19 日改訂

論文投稿の手引き

1. 論文投稿

(1) 投稿前

- 著者は「論文投稿の手引き」,「スタイルフォーマット」を学会ホームページの「学会誌」,「英文論文誌」欄からダウンロードして，投稿論文の準備をする。
- 「スタイルフォーマット」は TeX 版，Word 版の 2 種類が準備されており，出版コストおよび出版に至る労力削減のために，TeX 版の利用を推奨している。
- 原稿には以下の 6 種類がある。
  - ○ 論文：画像電子技術に関する独創的な研究成果，開発結果，検討結果を学術論文（フルペーパー）としてまとめたもの。新しいアイディアの提案を目的とする通常論文であり，新規性，有用性，信頼性，了解性に基づいて評価する。原則として刷り上がり 8 ページ以内とする。
  - ○ ショートペーパー：研究速報的な性格を持ち，部分的にある程度良好な結果，新しい知見が得られたことを研究報告としてまとめたもの。原則として刷り上がり 4 ページ以内とする。（論文・システム開発論文のいずれとも組み合わせが可能とする）
  - ○ システム開発論文：通常論文の新規性，有用性とは別に，既存技術の組み合わせ，或いはそれ自身に新規性があり，開発結果が従来法より優れているか，或いは他システムへ応用可能であり，新しい知見が示されているもの。原則として刷り上がり 8 ページ以内とする。
  - ○ 資料論文：試験，運用，製造などの新しい成果をまとめたもの。あるいは現場の新技術紹介や提案などで，必ずしも独創的でなくてもよい。原則として刷り上がり 8 ページ以内とする。
  - ○ サーベイ論文：新しい研究の発表を目的とせず，既存の研究・開発をある視点の下に体系付け，比較してその位置づけを明確にし，時代的な変遷を明らかにすること等を目的とする論文。参考文献の網羅，全体像の呈示，客観的評価，特定機関の宣伝に与しないことなどが要求される。また，以下の要件を満たすことが望ましい。その分野に関する状況・課題・及びそれらが，研究者や関係者に与える影響を理解していること。その分野への参入，開発・応用等への進展を促しさらには他分野への展開や新たな研究を触発する可能性を有すること。原則として刷り上がり 8 ページ以内とする。必要に応じて複数パート構成とすること。
  - ○ 実践論文：事例※やフィールドワークに関する発見，工夫，既存技術の新たな組み合わせ等で，事例設計の客観性・相互主観性を有し，個別の事例（フィールドワークも含む）から，他の事例（より一般的な事例も含む）への発展可能性（一般法則化），利活用・応用可能性などを示せるもの。原則として刷り上がり 6 ページ以内とする。
    ※事例とは，画像関連技術の開発・利活用に関する，フィールドワーク，システム開発，システム導入，ソフトウェア実装，アプリケーション開発，表示・提示方法，入出力インタフェース開発，試験・評価方法，サービス，データベース作成等の実例を示す。
- 原稿の投稿および掲載の条件として，論文，ショートペーパー，資料論文，実践論文のいずれも，著者のうち必ず 1 名以上は本学会の正会員または学生会員であることを必要とする。
- 原稿の二重投稿を禁止する。論文，ショートペーパー，システム開発論文，資料論文，実践論文

のいずれも，その採録決定以前に投稿原稿と同一内容のものが，投稿原稿と同一の著者もしくは少なくともその中の 1 名を含む著者によって他の公開出版物に掲載または投稿中の場合には，原則として採録としない。公開出版物とは，内外の書籍，雑誌および官公庁，学校，会社等の機関誌，紀要などをいう。ただし，本学会または他学会（国外の学会を含む）等の年次大会，研究会，シンポジウム，コンファレンス，講演会などの予稿，特許公開公報，当学会編集委員会で認めたものなどについてはこの限りではない。また，本会誌にショートペーパーとして採録になったものについて，成果を追加し内容を充実させ，論文として再び投稿したものについては二重投稿とみなさない。

(2) 投稿

- 投稿原稿では，全ての著者情報を削除すること。ただし，参考文献はこの限りではない。
- 著者は以下の論文投稿のページよりユーザ登録を行った上で，ログインして必要事項の記入と「スタイルフォーマット」を使用し作成した論文原稿のアップロードを行う。論文原稿のファイルは pdf 形式（TeX の dvi を変換したもの）または Word 形式とする。「スタイルフォーマット」を使用するため，原則として図表等は論文原稿内に組み込む。（査読の参考となる別のファイルを同時にアップロードすることは可能）

  https://www.editorialmanager.com/iieej/

- 投稿に関して質問がある場合には，学会事務局編集担当者に相談のこと。（連絡先：〒116-0002 東京都荒川区荒川 3-35-4 ライオンズマンション三河島第二 101 号 TEL: 03-5615-2893 FAX: 03-5615-2894 E-mail: hensyu@iieej.org 画像電子学会事務局編集担当）

2. 論文査読と対応

(1) 査読

- 論文原稿は該当専門分野に詳しい査読者によって審査され，「採録」，「条件付採録」または「返戻」のいずれかに判定される。査読結果は，学会事務局より連絡著者宛にメールで通知される。なお，最終的な採録判定についてはショートペーパーを除き，いずれのカテゴリーの論文であっても 2 名以上の査読者の採録を必要とする。
- 評価方法
  - ○ 通常論文は，次の基準で評価する。
  - ✓ 新規性：投稿内容に著者の新規性があること。
  - ✓ 有用性：投稿内容が学術や産業の発展に役立つものであること。
  - ✓ 信頼性：投稿内容が読者から見て信用できるものであること。
  - ✓ 了解性：投稿内容が明確に記述されており，読者が誤解なく理解できるものであること。
  - ✓ 文献調査：投稿内容に関連する文献調査が行われており，必要な比較が行われていること。
  - ○ ショートペーパーは，通常論文の新規性，有用性とは別に，研究内容に速報性があり，斬新性の尺度からある程度良好な結果，部分的にでも新しい知見が得られたことを評価する。
  - ○ システム開発論文は，通常論文の新規性，有用性の代わりに次の基準で評価する。
  - ✓ システム開発の新規性：既存技術の統合であっても，組み合わせの新しさ，システムの新しさ，開発したシステムで得られた知見の新しさなどをシステムの新規性として認める。
  - ✓ システム開発の有用性：類似システムと比較し，総合的或いは部分的に優れていること。システムとして先駆的な新しい応用コンセプトが示されていること。組み合わせが実用化のために合目的最適性を有すること。性能限界や実利用システムの性能事例が示されていること。
  - ○ 資料論文は，通常論文における新規性要件については，試験，運用，製造などの新しい成果，

現場の新技術紹介や提案など，必ずしも独創的でなくても優先性があれば新規性として認める。また，有用性要件については新しい成果が既存技術と比較し優れており，学術や産業の発展に役立つものであればそれを評価する。

○ サーベイ論文は，通常論文における新規性に代えて網羅性，俯瞰性，客観性について評価する。信頼性，了解性，及び文献調査においては，基本的に通常論文と共通の観点により評価する。

○ 実践論文は，通常論文の新規性，有用性に代えて合目的性，独自性，適応性の基準で評価する。また，信頼性，了解性，及び文献調査においては，基本的に通常論文と共通の観点により評価する。(なお，5 段階評価において通常論文の評価項目を流用する場合は通常論文の新規性を「合目的性」に，有用性を「独自性」に，信頼性を「適応性及び信頼性の総合」に読み替えて記入するものとする。)

✓ 合目的性：投稿内容に合目的性があること。

✓ 独自性：投稿内容に著者の独自性や拡張性があること。

✓ 適応性：投稿内容が先行事例より優れていること。また，当該事例の合目的性に対する要求条件及び適用範囲において，当該事例が同一条件下で先行事例と比べ，複数の視点で高い客観性・相互主観性を有し，サービスの向上に貢献するものであること。

これらの具体的な例を以下に示す。

✓ 合目的性：当該事例への提案内容の導入が一定の目的にかなった方法であることが示されていること。また，分析結果の適用範囲において高い合目的性を有すること。

✓ 独自性：当該事例において先駆的な独自応用コンセプトが示されていること。先行事例の組み合わせであっても組み合わせの独自性，導入事例の適用範囲の独自性，事例構成の独自性，他の事例（システム，サービス）への拡張性，現状の画像電子技術に係るサービスの有効性，得られた知見の独自性などを事例の独自性として認める。

✓ 適応性：提案事例のもつ適応性を判断するために充分な導入分析結果及び考察が示されていること。提案事例と先行事例との適応性に関する比較結果が，共通条件下で示されていること。提案事例が先行事例と比較し，総合的或いは複数の部分で性能が優れていること，或いは課題，合目的性に対する要求条件及び適用範囲において，より高い客観性・相互主観性を有していること。

(2) 査読結果受領後の対応

・ 「採録」の場合，著者は最終原稿の準備（3. に記載）に進む。

・ 「採録（参考意見有り）」の場合，著者は査読者からの参考意見を考慮して論文を修正し，最終原稿の準備（3. に記載）に進む。

・ 「条件付採録」の場合，著者は査読者から示された「採録の条件」に基づき，指定された期日まで（60 日以内）に論文を修正し，照会内容への対応を明記した回答書とともに再提出し，再度審査を受ける。修正論文には，修正箇所がわかるように，色を付けるか，下線を引くなどすること。また，回答書は必ず添付し，採録条件全てに対し，修正点，主張点，今後の課題などを丁寧に説明すること。

・ 「返戻」の場合，著者は以降のプロセスに進むことができない。査読者から示された「返戻の理由」を吟味して論文内容を再検討し，あらためて投稿することができる。

(3) 修正原稿査読

・ 「条件付採録」で再提出する場合は，回答書・修正清書原稿・修正箇所マーク原稿の修正原稿一

式を投稿ページより再提出する。但し，「条件付採録」で指定された期日を過ぎて提出された場合は，新規の初回投稿論文として取り扱われるので注意すること。
- 再提出された修正原稿は原則として同一の査読者によって再審査され，「採録（参考意見付記を含む）」または「返戻」のいずれかに判定される。
- 以降の対応は（2）と同様である。

3. 最終原稿提出から出版まで

(1) 最終原稿提出

- 「採録」通知受理後は，事務局から最終原稿作成に関するメールが送付されるので，著者はその指示に従い最終原稿一式（電子データ）を準備し，指定の期日までに，事務局へ送付する。
- 最終原稿一式とは，「スタイルフォーマット」を使用して作成した論文原稿のソースファイル（TeX 版または Word 版），その pdf ファイル，すべての図表の eps ファイル(bmp, jpg, png なども含む)，著者紹介用写真（縦横比 3:2，胸より上，300dpi 以上の eps ファイル，或いは，jpg ファイル）であり，これらを zip 等の圧縮ファイルに格納して提出する。
- 最終論文原稿には，著者名・組織名，著者紹介，必要により謝辞を記入する。（スタイルファイル内のマクロのフラグを解除する）
- 採録が決定した著者は，当該論文の掲載料を支払わなければならない。なお，別刷りの購入は任意。（別紙１「別刷価格表」参照）

(2) ゲラ刷りの校正

- 論文掲載号の発行のおよそ１ヶ月前に論文原稿のゲラ刷り（ハードコピー）の確認が依頼される。著者は指定の期日まで（約１週間）に校正を行い，修正した原稿をスキャンした pdf ファイルを作成し，事務局へメールにて送付する。この際，「別刷り購入票」および「コピーライトフォーム」を記入し，同様にスキャンした pdf ファイルをアップロードする。
- 本会誌に掲載されるすべての記事については，電子的形態による利用も含め，著作権は原則として本会に帰属する。
- 「別刷申込書」および「コピーライトフォーム」は学会ホームページの「学会誌」欄からダウンロードする。（別紙２「別刷申込書」，別紙３「コピーライトフォーム」参照）

(3) 出版

- 最終校正された論文原稿は学会誌，或いは，英文論文誌（電子ファイル形態）としてオンライン出版されるとともに，学会ホームページに掲載される。
- 著者は、グリーンオープンアクセス（セルフアーカイブ）として，本誌からの出版ののちすぐに，自身の著者最終稿を機関リポジトリなどの公的なオンラインリポジトリから公開することができる。

以上

アレキサンダー・ベイン賞候補推薦募集

2026年10月1日

一般社団法人　画像電子学会

会　長　土橋　宜典

本学会では，画像工学の学術の進展に対して卓越した業績のあった会員，または産業界の発展に特別の功労がありその功績が顕著であった会員へのアレキサンダ ー・ベイン賞推薦候補を募集します.

別紙の候補者推薦書にご記入の上，2027年2月12日（金）必着で学会事務局までにメール添付で送付ください（推薦書の電子ファイルは上記学会ホームページ（http://www.iieej.org/）からダウンロードできます).

アレキサンダー・ベイン賞規程（抜粋）

第１条　本会定款第４条に基づく画像工学に関する学術に関し卓越した業績のあった者，または産業界や本学会関連事業に関し特別の功績があった者に対し，本規程による表彰を行う.

第８条　本賞は，10 年を超えて本学会会員である者を受賞対象とし，過去に本賞を受賞していない者のうちから，原則として毎年若干名以内を選定し贈呈する.

第９条　各年度に適切な該当者がいない場合は，該当者なしとする.

第１０条　本賞として，賞状および記念メダルを贈呈する.

アレキサンダー・ベイン賞選定手続き規程（抜粋）

第２章　選定委員会選定細則

第５条 アレキサンダー・ベイン賞の一般推薦の推薦状フォーム並びに推薦締め切り期日については期日前にフェロー会員に周知する.

２．推薦状にはフェロー会員２名の推薦を必要とする.

３．推薦状フォームは本規程別紙に掲げるものを使用する

［贈呈式］

2027年の総会もしくは年次大会で行う予定です.

上記，アレキサンダー・ベイン賞規程等の詳細は当学会ホームページ（https://www.iieej.org/admission/arexande/）を参照してください.

以上

アレキサンダー・ベイン賞　　候補者推薦書

<table>
<tr><td>受賞候補者名</td><td>氏名
会員番号,
在籍年数所属,
住所,
電話番号,
メールアドレス</td></tr>
<tr><td rowspan="3">受賞候補者経歴</td><td>学歴，職歴：</td></tr>
<tr><td>本学会での活動歴：</td></tr>
<tr><td>学術的業績，産業的貢献，社会的貢献：</td></tr>
<tr><td colspan="2">推薦理由（サイテーション）</td></tr>
<tr><td colspan="2">主たる功績（いずれか，または複数に〇印）
（1）学術的業績　（2）産業的貢献　　（3）社会的貢献（4）学会活動
推薦理由</td></tr>
<tr><td>受賞候補者の推薦者
氏名（フェロー認定年）</td><td>（1）　　（　　）
（2）　　（　　）</td></tr>
</table>

画像電子技術賞候補の推薦のお願い

画像電子学会では，画像電子に関する極めて優れた製品，システム，デバイスを開発した個人またはグループを毎年若干件選出して画像電子技術賞として表彰します．つきましては，本技術賞に相応しいと思われる候補（個人もしくはグループ）を下記要領に従って，推薦下さるようお願い致します．なお，選定は画像電子技術賞選定委員会で行います．

記

・推薦者の資格：本学会正会員であること（一人１件に限る）．
・候補者の資格：表彰時において，本学会会員であること．
・推 薦 方 法 ：次頁の推薦用紙またはこれをコピーしたものに，
（１）推薦する個人もしくはグループ全員の氏名，所属（連絡先）
（２）推薦する製品名，システム名，デバイス名
（３）推薦理由
（４）推薦者の氏名，所属，住所，電話／ＦＡＸ番号，e-mailを記入の上，学会あてＦＡＸ，郵便またはe-mailでお送り下さい．
尚，推薦理由を裏付ける資料があれば，その案内（論文の場合であれば，題目，発表機関，ＶＯＬ．ＮＯなど）を推薦理由欄に付記して下さい．

・推薦の締切：2027年2月12日（金）
・送 付 先：〒116-0002 東京都荒川区荒川三丁目35番4 ライオンズマンション三河島第二 101号
一般社団法人　画像電子学会　事務局
・表彰内容：楯及び表彰状を贈呈します．
・表彰方法：通常総会もしくは年次大会の席で表彰するとともに，学会誌に解説を掲載します．
・技術展示：受賞技術は，2027年の年次大会にて技術展示をお願いします。

■ 画像電子技術賞授賞リスト（最近の例）

受賞年度	受賞技術名もしくは製品・システム名	受賞者所属先
2023年	波長依存PSFメタレンズを用いた圧縮スペクトルイメージング	曽我部 陽光（日本電信電話株式会社）他
2021年	３次元データ空間の歪み検出・補正およびスケール付与システム	青木 勝彦（株式会社AB.do）他
2020年	非エピポーラ光伝搬の計測と解析に基づく皮下血管のリアルタイムイメージング	久保尋之（東海大学）他
2019年	非接触式血中酸素飽和度測定装置	長尾嘉満（早稲田大学）
2018年（2件）	羽虫の群れの動き生成システム	北海道大学大学院情報科学研究科
	展示会見学体験向上を目的とするユニバーサルオブジェクト認識技術を用いたMICEアプリの提供	NTTサービスエボリューション研究所
2017年（２件）	アクティブ照明と多視点カメラ入力による実時間インテグラル立体表示	名古屋大学大学院工学研究科/日本放送協会
	視覚の知覚メカニズムを活用した視点移動対応裸眼3D映像スクリーン技術	日本電信電話株式会社/東北大学
2016年	24時間連続稼動全天球ライブストリーミングカメラ RICOH R Development Kit	㈱リコー 技術経営センター

画像電子技術賞表彰規定　　2022年3月18日改訂

第1条（目的）
画像電子技術賞は，画像電子に関する極めて顕著な新しい技術，製品，システム を開発した個人またはグループを毎年若干件選出して表彰するものである．

第2条（選定）
画像電子技術賞は，別途定める「画像電子技術賞候補選定手続規定」に従って 候補を選出し，理事会で決定するものとする．理事会の承認後受賞者には内定通知を出し、学会ホームページに受賞内定者を掲載する。

第3条（表彰）
画像電子技術賞は，楯（1 件に 1 個）と賞状（各自）とし，翌年度の年次大会、或いは、翌年度の本学会総会で表彰するとともに，学会誌上で紹介する．なお，年次大会にて技術展示を依頼する

画像電子技術賞候補推薦用紙

＜技術・製品または，システム名＞		
＜受賞候補者＞（複数の場合は全員ご記入ください）		
氏　名	所　属	TEL／FAX／ e-mail
＜推薦者＞		
氏名：		
所属：		
TEL：		
FAX：		

以下の各項目にご記入ください。

・画像電子分野との関連性

・新規性

・有効性

・販売，採用などの実績

・参考となる資料

2026年度フェロー候補推薦募集

2026年10月1日

一般社団法人画像電子学会

会 長 土橋　宜典

本学会では，学術上の業績，学会活動の活性化，教育・社会活動への取り組み，標準化活動などにおいて，特に貢献のあった会員に対しフェローの称号を授与する制度を設けています．フェロー称号認定につきましては以下の通り実施致しますので，フェロー候補の推薦をお願いしま す．なお，本情報は学会ホームページからも参照できます．

［推薦募集要領］

別紙のフェロー候補推薦書にご記入の上，2027年2月12日（金）必着で学会事務局までにメール添付で送付ください（推薦書の電子ファイルは学会ホームページからダウンロードできます）．選定委員会で審査の上，理事会で決定されます．

なお，条件，資格等は学会ホームページに記載されていますので参照して下さい．

（フェロー制度規程.pdf）

また，候補者 1 名に対し推薦者は 2 名必要ですが，その内少なくとも一人は候補者と同一の機関に属されない方といたします．

［称号授与］

フェロー称号授与式は2027年の総会もしくは年次大会で行う予定です．

以上

画像電子学会フェロー候補推薦書

フェロー 候補者	氏名 会員番号 在籍年数 所属 住所 電話番号 メールアドレス
候補者略歴	学歴，職歴：
	本学会での活動歴：
	業績；（特に，本学会分野での業績）
功績内容（サイテーション）（２２文字以内）	
主たる功績（いずれか，または複数に〇印） （１）工学的・科学的先駆者　（２）学会活動推進者　（３）技術開発指導者 （４）教育者　（５）標準化・社会活動 推薦理由	
フェロー候補の推薦者 氏名（会員番号）	（１）　　（　　）㊞ （２）　　（　　）㊞

（注）　記入スペースが不足する場合はそのまま行数を増やして書いてください．複数ページでも結構です．

代議員候補推薦のお願い

一般社団法人 画像電子学会
代議員選挙管理委員長

画像電子学会では，2027年度の代議員改選を行います.
下記の要領で代議員のご推薦をお願いします.

1)代議員とは

毎年１回6月に開催される総会に出席し，予算，決算，役員，事業計画など学会の重要事項につき審議する役割の人です．以前の総会には正会員なら誰でも出席し審議に参加出来たのですが，代議員制では正会員及び名誉会員を代表する形で代議員が総会に出席することになります.

総会は，学会の最高議決機関ですのでそれに正会員及び名誉会員を代表して出席する重要な役割を担うことになります.

2)代議員の人数と任期は

30人ぐらいで，任期は2年で（2027年選出の方は，2027.6.総会日～2029.6.総会前日），2期程度とします.

3)選出する方法は

立候補及び学会の正会員，名誉会員などからの推薦により候補を定め，正会員，名誉会員の選挙によって選出します.

4)どのような人が代議員になれるのか

代議員は正会員，名誉会員から選出されます．但し役員（理事，監事）に就いている人は代議員になれません.

5）推薦決定及び選挙方法は

ご推薦頂いた方を代議員選挙管理委員会にて資格審査し信任方法で選挙を行います.

学会の発展のためにお尽力頂ける意欲的な方をご推薦下さい.
勿論**自薦でもかまいません．よろしくお願いします.**

・推薦及び送付方法：添付の推薦書に必要事項を記入しメール又はFAXで
事務局 大嶽宛にお送り下さい.
・送り先：FAX；03-5615-2894　E-mail；hyoujun@iieej.org
・締め切り日：2027.2.12(金)
・推薦の場合は，被推薦者の了解を得ておいてください．（非会員の場合は推薦後**正会員にご入会頂きます）**

氏　名	
所属・役職	
住所	
TEL・FAX	
e-mail	

推薦者＿＿＿＿＿＿

＜参考＞ 代議員選挙規程（必要な方は，事務局へお申し込み下さい．メールでお送りします）

日本画像学会誌　　第 65 巻　第 4 号　（通巻 282 号）　2026 年

目　　次

解　　説
多次元顕微分光分析のためのインフォマティクス技術活用……永村直佳，五十嵐康彦…318（ 2 ）
複写機遺産特集
Introduction……加藤知之，宮本栄一…326（ 10 ）
第 5 回複写機遺産概要……複写機遺産委員会…327（ 11 ）
第 5 回複写機遺産認定機種の解説
第 15 号富士ゼロックス 813……富士フイルムビジネスイノベーション株式会社…337（ 21 ）
第 16 号キヤノン NP-8500……キヤノン株式会社…342（ 26 ）
第 17 号ミノルタ EP450Z……コニカミノルタ株式会社…345（ 29 ）
第 18 号リコー リコピー FT4060……株式会社リコー…349（ 33 ）
Imaging Today
「4DFF2025 特集」
Introduction……朝武　敦，名越応昇，宮本栄一…352（ 36 ）
2025 年日本国際博覧会（大阪・関西万博）から考える 4D 造形の展望……田中浩也…353（ 37 ）
平面板による空間的パズルを用いた立体造形と創作支援……呂　亜輝，敖　俊偉，邵　　帥…358（ 42 ）
イシガキフグの棘および表皮の立体構造
……禾丈一郎，下田悠太，天童智也，森田哲朗，斉藤一哉，舘　知宏…363（ 47 ）
ボクセルファイル形式 FAV を利用した 3D モデリング用ファイル生成の基礎技術
……山崎　淳，水野　修…372（ 56 ）
材料押出法におけるカテナリー構造を有する柔軟構造体の設計手法……中原稜将，木内俊克…382（ 66 ）
大阪関西万博トイレ「島の蜃気楼」における，3D プリント外装材の設計・
製造プロセスと廃材リサイクル……田住　梓，鷲見　良，池本祥子，田川直樹…388（ 72 ）
Imaging Highlight
シャープミュージアム……内田隆史…395（ 79 ）
教育講座
画像技術を支える人間工学の基礎（IV）
―人間工学に基づいた高齢者・障がい者に配慮した機器の設計指針に関する検討―
西口宏美…399（ 83 ）
研究室訪問
東海大学イメージング研究センター……410（ 94 ）
2025 年度日本画像学会表彰……413（ 97 ）
一般社団法人　日本画像学会第 69 回定時総会資料……426（110）
会報……480（164）
会告……483（167）
投稿案内……485（169）
日本写真学会誌の目次……487（171）
日本印刷学会誌の目次……488（172）
画像電子学会誌の目次……489（173）
Journal of Imaging Science and Technology の目次……490（174）
画像閑話

令和八年編集委員

編集委員長	宮本栄一（京セラドキュメントソリューションズ）	編集幹事	石塚一輝（コニカミノルタ）
編集副委員長	名越応昇（三菱製紙）	編集副幹事	中村一希（千葉大学）
	山口大地（リコー）		

編集委員

新井啓之（日本工業大学）	次田将大（神奈川県立産業技術総合研究所）	名越応昇（三菱製紙）
石塚一輝（コニカミノルタ）	常安翔太（東京工芸大学）	前田秀一（東海大学）
岩田　基（京都橘大学）	椿　頼尚（シャープ）	宮本栄一（京セラドキュメントソリューションズ）
梅津信二郎（早稲田大学）	朝武　敦（ATI-C コンサルティング）	村上和歌子（リコー）
加藤知之（東芝テック）	内藤裕義（大阪公立大学）	山口大地（リコー）
岸　由美子（リコー）	中村一希（千葉大学）	山﨑　弘（元コニカミノルタ）
齋藤将史（京セラドキュメントソリューションズ）	中山信行（東京工芸大学）	横山優樹（富士フイルムビジネスイノベーション）

日本画像学会誌　　第65巻　第5号（通巻283号）　2026年

目　　次

論　文

フィルム成膜における温度不均一性に起因したすじ形成要因
……西岡柊哉，畑中悠希，中田星空，岡　稔，日野原耀，砂見雄太，岩森　暁…494（2）

セラミックス造粒粉を用いたバインダージェッティング技術における造形中の空隙形成挙動の観察
……小西鷹介，鴨田紀一，關根暢秀…505（13）

ロール形状に応じたシート状メディア変形の発生形態……木内　豊，前田秀一…512（20）

高抵抗メディアの静電気課題を改善する接触・非接触型のハイブリッド除電技術
……湯浅宏一郎，木内　豊…519（27）

速　報

遮蔽・隔離環境下におけるハンディ型LiDARを用いた3Dモデル構築に関する検討
……熊坂憲広，岩田悠希，藏部光一郎，内田孝幸…529（37）

Imaging Today

伝承から飛躍へ～画像技術を支える材料技術
……山﨑　弘，朝武　敦，名越応昇，齋藤将史，椿　頼尚…534（42）

トナーの進化を支えた材料技術動向―総説―……トナー技術部会…535（43）

トナー用樹脂……本夛　将，前田真一…543（51）

トナー用ワックスの基礎と最新動向……西村佳真…550（58）

トナー用外添剤の進化：基礎から見る技術の変遷……山田謙太，内藤直弘…555（63）

トナー用酸化チタン系外添剤……田中貴康，古賀俊之…564（72）

電子写真トナー向け顔料……原田大輔，森本早貴…570（78）

レーザープリンター用トナー向けカーボンブラック特性と選定に関する推奨事項
……ドミトリー・フォミチェフ…576（84）

Imaging Highlight

「写真」からイメージングの次の100年へ―日本写真学会創立100周年記念式典から―
……常安翔太…581（89）

教育講座

交通システムにおける画像技術と人間工学―ADAS，自動運転，交通インフラにおける人とシステムの協調―……平岡敏洋…585（93）

研究室訪問

大阪公立大学・森ノ宮キャンパス……594（102）

会報……598（106）

会告……604（112）

訂正……605（113）

投稿案内……606（114）

日本写真学会誌の目次……607（115）

日本印刷学会誌の目次……608（116）

画像電子学会誌の目次……609（117）

Journal of Imaging Science and Technology の目次……610（118）

画像閑話

令和八年編集委員

編集委員長　宮本栄一（京セラドキュメントソリューションズ）
編集副委員長　名越応昇（三菱製紙）
　　山口大地（リコー）
編集幹事　石塚一輝（コニカミノルタ）
編集副幹事　中村一希（千葉大学）

編集委員

新井啓之（日本工業大学）　次田将大（神奈川県立産業技術総合研究所）　名越応昇（三菱製紙）
石塚一輝（コニカミノルタ）　常安翔太（東京工芸大学）　前田秀一（東海大学）
岩田　基（京都橘大学）　椿　頼尚（シャープ）　宮本栄一（京セラドキュメントソリューションズ）
梅津信二郎（早稲田大学）　朝武　敦（ATI-Cコンサルティング）　村上和歌子（リコー）
加藤知之（東芝テック）　内藤裕義（大阪公立大学）　山口大地（リコー）
岸　由美子（リコー）　中村一希（千葉大学）　山﨑　弘（元コニカミノルタ）
齋藤将史（京セラドキュメントソリューションズ）　中山信行（東京工芸大学）　横山優樹（富士フイルムビジネスイノベーション）

日本写真学会誌
89巻Ⅲ号（2026年/令和8年）

口絵・口絵解説

141 「画像からくり」
第74回　LED光源を移動して作り出す空中ディスプレイ
桑山 哲郎

143 会 告 VISION100未来サポート基金・魅力工学研究会シンポジウム・スマホ技術セミナー・カメラ技術セミナー・画像保存セミナー
他学会誌目次 日本画像学会誌・画像電子学会誌・IIEEJ Transactions on Image Electronics and Visual Computing・日本印刷学会誌

151 日本写真学会　2026年度　学会賞受賞者

157 特集 2025年の写真の進歩

特集 2025年のカメラ

203 解説 ラージフォーマットレンズ一体型カメラ FUJIFILM GFX100RFの開発
田村一紀

207 解説 M.ZUIKO DIGITAL ED 50-200 mm F2.8 IS PROの開発
横山恵太郎

211 解説 EOS R1の開発
荒川世一

一般論文

216 狭隘環境におけるハンディ型LiDARスキャン利用の検討：洞窟環境の保全に資する西表島クーラ洞窟の三次元記録
内田孝幸，熊坂憲広，岩田悠希，武井琉奈，安田洋司，渡辺　信

226 2026年度（一社）日本写真学会年次大会　講演要旨

279 一般社団法人日本写真学会　第15回 定時社員総会 議事録

280 会 報 理事会抄録，入会のおすすめ

編集委員：山田 勝実，久下 謙一，稲垣 由夫，岩崎　仁，桑山 哲郎，横山 恵太郎，水口　淳，西垣 仁美，宮本 成悟，常安 翔太、白石 幸英

日 本 印 刷 学 会 誌

目 次

第 63 巻　第 4 号（2026 年 8 月）

■ 巻頭言
共進する印刷とデジタル技術 . . . . . 藤井里絵　199
■ 総説　特集「印刷資材のトレンドウォッチ」
UV 硬化型インキの技術動向 . . . . . 花田朋広　200
印刷品質と生産性を支えるテープ技術
―フレキソ印刷用クッションテープと高速印刷・加工工程における
スプライステープの技術動向― . . . . . 安斎敏一　204
積層造形における粉体材料ハンドリングの最新動向 . . . . . 尾形公一郎　208
■ 印象記
2025 年度日本印刷学会 西部支部見学会
「『大阪シーリング印刷株式会社 イノベーションセンター』
および『ラベル史料館 – サロン OSP–』見学会」 . . . . . 舟越一郎　214
2026 年度スクリーン印刷技術研究会セミナー
「スクリーン印刷基礎講座『入門・基本編』」 . . . . . 吉田旦人　216
■ Abstract 217
■ 資料 218
■ 学会だより 222
■ 日本印刷学会賛助会員名簿 229

日本印刷学会誌編集委員

編集委員長　矢口博之（東京電機大学）
委　　員　東　吉彦（東京工芸大学）　杉山　徹（大日本印刷（株））　小泉真一（共同印刷（株））
光本知由（富士フイルム（株））　前田満穂（(独）国立印刷局）
矢島久夫（東洋インキ（株））　伊藤　学（TOPPAN ホールディングス（株））
次田将大（KISTEC）　大場久恵（東京電機大学）　常安翔太（東京工芸大学）
幹　　事　小関健一
表紙デザイン　益田宏樹

日 本 印 刷 学 会 誌

目 次

第 63 巻　第 5 号（2026 年 10 月）

■ 巻頭言
持続可能な社会実現に果たす紙の役割 . . . . . 仲山伸二　237
■ 総説　特集「印刷と立体デザイン」
折り紙食器「beak」と今後の展望. . . . . 山田秀生　238
工学としての折り紙：機械的メタマテリアルから宇宙構造まで. . . . . 安田博実　242
硬い電子素子を用いた伸縮ディスプレイ . . . . . 佐藤　峻・岩﨑　渉・岩瀬英治　246
印刷技術で実現する紙の自律構造形成と折紙デバイス . . . . . 本山　聖・重宗宏毅　252
形状プログラマブル材料から三次元追従型フレキシブルエレクトロニクスまで
. . . . . 荒木徹平・阿部岳晃　259
■ 印象記
2026 年日本印刷学会夏期セミナー「印刷技術の革新と持続可能な未来」. . . . . 千葉英輔　263
グラビア研究会　第 12 回グラビア基礎講座に参加して . . . . . 笠原太一　264
2026 年度技術委員会フレキソ研究会第 23 回研究例会 . . . . . 横井郷平　266
■ Abstract 268
■ 学会だより 270

日本印刷学会誌編集委員

編集委員長　矢口博之（東京電機大学）
委　　員　東　吉彦（東京工芸大学）　杉山　徹（大日本印刷（株））　小泉真一（共同印刷（株））
光本知由（富士フイルム（株））　前田満穂（(独）国立印刷局）
矢島久夫（東洋インキ（株））　伊藤　学（TOPPAN ホールディングス（株））
次田将大（KISTEC）　大場久恵（東京電機大学）　常安翔太（東京工芸大学）
幹　　事　小関健一
表紙デザイン　益田宏樹

会 報

○新入会員紹介（2026 年 9 月 30 日）

正会員
澤山 正貴（北海道大学）
大島 裕明（兵庫県立大学）
王 旭（北海道大学）
金子 悟士（(株) リコー）
内田 拓人（ＪＲ東日本情報システム）
岩崎 真己（ＪＲ東日本情報システム）
髙橋 ひとみ（ＪＲ東日本情報システム）
方志 卓朗（ＪＲ東日本情報システム）
伊藤 和敬（ＪＲ東日本情報システム）
日浦 慎作（兵庫県立大学）
山口 直木（高松大学）
大平 直也（京都大学）

学生会員
野坂 亮介（北海道大学）
稲垣 壮冴（北海道大学）
須佐 千風（北海道大学）
中尾 温人（北海道大学）
五十嵐 京（北海道大学）
福原 将真（北海道大学）
中村 薫（北海道大学）
青木 ひとみ（北海道大学）
吉村 斗来（北海道大学）
栗田 千輝（北海道大学）
鈴木 周平（北海道大学）
梅内 飛翔（北海道大学）
Nguyen Hoang Anh（北海道大学）
本間 蒼空（北海道大学）
山田 昂和（北海道大学）
宮原 毅仁（北海道大学）
北ノ原 健太（北海道大学）
中島 悠碧（北海道大学）
浅利 かなた（宇都宮大学）
平岡 良和（徳島大学）
冨田 暁翔（大阪工業大学）

.○会員現況

名誉会員	18 名
正会員	419 名
学生会員	57 名
賛助会員	14 社 19 口
特殊会員	23 件

編集後記

令和 8 年 10 月号をお届けします．本号は，コンピュータビジョン，コンピュータグラフィクス，及び関連の画像処理に関する研究を取り上げるビジュアルコンピューティング（VC）特集号の 32 回目となります．特集論文としては「Degrad2vec: 劣化埋め込みベクトルを用いた対話的画像復元」，「かぎ針編みにおける動作センシングと技能習得支援」，「ストックの八重鑑別における日齢の影響検証」，「小規模データセットにおけるニューラルネットワークと Hu モーメントを用いた小松菜の日齢推定」，「Gaia 星表を用いた天体写真撮像系の内部標定」の 5 編を掲載しました．画像復元から手仕事の支援，農業，天体観測まで，VC 技術の応用の広がりを感じていただける内容となりました．また，本特集で採択された質感探索に関する英語論文 1 編は英文誌 12 月号への掲載を予定しております．併せてご一読いただければ幸いです．

次に，一般論文としては「ハイパースペクトル画像を用いた ki-67 陽性核推定法の開発」，「改良型 StarGAN を用いたデータ拡張による近代書籍文字認識精度の向上」の 2 編を収録しています．また，坂東幸浩企画委員長による随想「研究会の楽しみ方」では，研究会のもつ成果の発表の場という位置付けに加え，さらに広くその成果を育む交流の場としての活用法についても示唆深い考察を頂戴しました．特に若手の研究者には是非目を通していただきたく存じます．

さて，当学会最大の年次行事である本年の年次大会は，8 月 24 日から 26 日まで筑波大学にて無事開催されました．通常講演 50 件に加え，三谷純氏による「曲線折り紙設計手法」，河村尚登氏による「グリーンノイズの画像工学への応用」という 2 件の特別講演が行われ，各セッションでは活発な議論が交わされました．本号にはこの年次大会報告と，大野邦夫氏が受賞された第 10 回アレキサンダー・ベイン賞の授賞報告も掲載しています．大会の様子や受賞者のご功績につきましては，ぜひこれらの記事をご参照ください．

また，本学会 VC 研究会などが共催する Visual Computing 2026 は，9 月 8 日から 10 日まで北海道大学学術交流会館にて開催され，本年も論文発表やポスター発表，招待講演，企業展示などを通じて，研究者・技術者・クリエイターが交流する貴重な機会とできました．ご参加いただいた皆様，ならびに企画・運営にご尽力いただいた関係者の皆様に，この場を借りて厚く御礼申し上げます．

本誌に掲載される研究論文は，扱う対象も用いる手法も様々ですが，身近なテーマやそれに伴う実際の作業・観測における画像・視覚情報の課題を丁寧に捉え，技術的な解決へとつなげています．たとえ異なる分野であってもその工夫に触れることは，自らの研究を見つめ直すヒントとなるものと思われます．また，巻頭の随想にも取り上げられているように，研究発表での質疑や発表後の会話から生まれる議論も，新たな研究の方向へ導く大切な一歩になるでしょう．本学会を，そうした出会いや研究成果の発展の場として活用いただけましたら幸いです．最後になりましたが，本誌の論文，記事をご執筆いただいた方々，査読・編集にご協力いただいた皆様に改めて心より御礼申し上げます．

（ゲスト編集幹事　謝　浩然 記）

画像電子学会誌
第 55 巻 第 4 号（通巻 278 号）
令和 8 年 10 月 30 日発行（年 4 回発行）
©2026 画像電子学会
E-mail：hensyu@iieej.org
https://www.iieej.org/

発行所 一般社団法人 画像電子学会
〒116-0002 東京都荒川区荒川 3-35-4 ライオンズマンション三河島第二 101 号
TEL（03）5615-2893 FAX（03）5615-2894（振替 00180-3-166232）
編集・発行者 内田 理

画像電子学会誌
Journal of the IIEEJ 278
2026 Vol.55 No.4

ビジュアルコンピューティング論文特集号

第55巻 第4号