令和 8 年 8 月 20 日

国立大学法人福井大学
University of British Columbia
Vector Institute

少ないデータ・少ない計算で賢く学ぶ画像認識 AI を実現
――生成 AI と基盤モデルを活用した 3 つの新技術を開発――

本研究成果のポイント

◆医療現場や製造現場では AI 導入を阻む大きな課題があったが、今回の3つの新技術の開発により、不良品のデータが乏しい製造ラインの自動検品、専門医の手作業といった負担が大きい医療画像診断の支援など、従来では困難だった現場への適用が期待される。

①少データでも AI が賢く学ぶ「unCLIP-KD」

画像生成 AI（注1）を用いて、わずかな画像データから本質は変えずに見た目のバリエーションだけを自動作成し、少ないデータでも効率よく学習させられる技術。

② 復元不要で超高速な異常検知（注2）「InvAD」

画像をノイズにした状態から直接傷や病変を見つけることで、圧倒的なスピードと手軽さを実現する技術。

③人間の見方に学ぶ自律学習「DSeq-JEPA」

人間のように「写真の重要なポイント」を判断し、優先順位をつけて分析するデータラベル不要の自己教師あり学習（注3）技術。

概要

福井大学学術研究院工学系部門情報・メディア工学講座の長谷川達人准教授、顧淳祉講師らの研究グループは、カナダ・ブリティッシュコロンビア大学などとの国際共同研究により、少ないデータと計算資源で高性能な画像認識 AI を実現する3つの新技術を開発しました。近年の AI は大規模化によって精度が向上する一方、膨大な学習データと計算コストを必要とするため、工場の検査装置やスマートフォンといった現場の小型機器で活用することが困難でした。研究グループは、(1) 画像生成 AI を活用し、ごく少数のデータから大規模 AI の知識を軽量な AI へ受け渡す「知識蒸留（注4)」技術 unCLIP-KD、(2) 生成 AI の計算を「逆向き」にたどるという発想の転換により、従来の2倍以上の速度で製品の傷や病変を発見する異常検知（注5）技術 InvAD、(3) 人間の視覚のように重要な部分から順に注目して自ら学ぶ自己教師あり学習技術 DSeq-JEPA を開発し、いずれも国際的な標準ベンチマークで既存手法を上回る性能を確認しました。これらの成果は、データや計算資源が限られた製造現場や医療現場への AI 導入を加速すると期待されます。

〈研究の背景と経緯〉

深層学習による画像認識は、モデルと学習データの大規模化によって飛躍的に進歩し、近年では CLIP（注 5）に代表される基盤モデル（注 6）が、さまざまな認識タスクに高い精度で対応できるようになりました。しかしその一方で、(a) 巨大なモデルは訓練に大規模なデータセットを必要とし、誰もが構築できるものではない、(b) 医療画像や製造業の不良品画像のように、そもそもデータの収集が困難な分野では学習データが不足する、(c) 大量のデータに人手でラベルを付けるコストが膨大である、という 3 つの課題が実用化の壁となっていました。

従来の知識蒸留は大規模な学習データを前提としており、少数データでは画像の回転や切り抜きといった単純なデータ拡張に頼らざるを得ませんでした。また、生成 AI の中核技術である拡散モデルを用いた異常検知は「画像を一度壊して復元し、元画像と見比べる」方式が主流で、ノイズ強度の細かな調整が必要なうえ、多段階の復元計算のため処理速度が遅いという本質的な課題がありました。さらに、ラベルなしで学習する自己教師あり学習の代表的手法は、画像内のすべての領域を一様に扱っており、人間のように「重要な場所から順に注目する」仕組みを持っていませんでした。

〈研究の内容〉

長谷川達人准教授を研究代表者とする研究グループは、これらの課題に対して以下の 3 つの技術を開発しました。

（1）unCLIP-KD：画像生成 AI を活用した少数データからの知識蒸留

実画像から基盤モデル CLIP で抽出した特徴ベクトルを「意味アンカー」（画像の意味の中心点）として固定し、これを条件に画像生成 AI（unCLIP）で見た目の異なる多様な画像を多数生成します。軽量な生徒モデルには、生成された画像群から元のアンカーを言い当てるよう学習させることで、見た目の変化に惑わされない本質的な特徴の抽出能力を転移させます。教師モデルの計算は事前の一度だけで済むため学習も高速です。学習データが 1 割しかない少数データ設定において、5 つの標準ベンチマークすべてで既存手法を上回り、特に STL-10 データセット（注 7）では従来比約 7.5%の精度向上を達成しました。

（2）InvAD：「復元しない」高速・高精度な異常検知

従来の拡散モデル型異常検知の発想を逆転させ、画像を復元する代わりに、入力画像を学習済みモデルの決定論的な計算経路に沿って潜在空間（注 8）上のノイズへと「逆変換」し、正常データの分布からのズレを直接スコア化する枠組みを提案しました。復元が不要なためノイズ強度の調整が原理的に不要となり、毎秒約 88 枚の画像を処理できます。これは従来の拡散モデル型手法の 2 倍以上の速度であり、産業製品の外観検査ベンチマーク（MVTecAD など）と医療画像ベンチマーク（BMAD）の双方で世界最高水準の検出精度を達成しました。既存の異常検知モデルに後付けで組み込んで高速化できる汎用性も備えています。本成果はブリティッシュコロンビア大学・Vector Institute との国際共同研究によるものです。

（3）DSeq-JEPA：人間の視覚に学ぶ自己教師あり学習

人間は風景を見るとき、最も情報量の多い部分（例：鳥の頭部の模様）から順に視線を移して理解を深めます。この知見に着想を得て、代表的な自己教師あり学習手法 I-JEPA（注 9）を拡張し、AI が自ら生成する注意マップから「識別に重要な領域」を抽出し、

重要度の高い順に次の領域の特徴を逐次予測させる新しい学習方式を提案しました。「どこを」「どの順番で」見るかを学習に組み込むことで、ImageNet での評価で従来手法を上回るとともに、鳥の種類や車種の識別といった細かな分類、物体検出や領域分割など幅広いタスクで一貫した性能向上を確認しました。本成果はブリティッシュコロンビア大学・ストラスブール大学との国際共同研究によるものです。

〈今後の展開〉

3 つの成果に共通するのは、生成 AI や基盤モデルの持つ膨大な知識を「AI の効率化」のために活用するという視点です。これにより、不良品データがほとんど手に入らない製造ラインの外観検査、専門医のラベル付けコストが大きい医療画像診断の支援、通信環境に依存しないエッジ機器上でのリアルタイム AI など、これまで AI 導入が難しかった現場への応用が期待されます。今後は、異なる撮影環境・分野へまたがる知識転移への拡張、さらなる推論の高速化（1 ステップ化）、視覚と言語を統合した基盤モデルへの発展などに取り組む予定です。なお、本研究は JSPS 科研費（26K02879、25H01110、23K11164）等の支援を受けて実施されました。

〈参考図〉

図 1：unCLIP-KD の概要

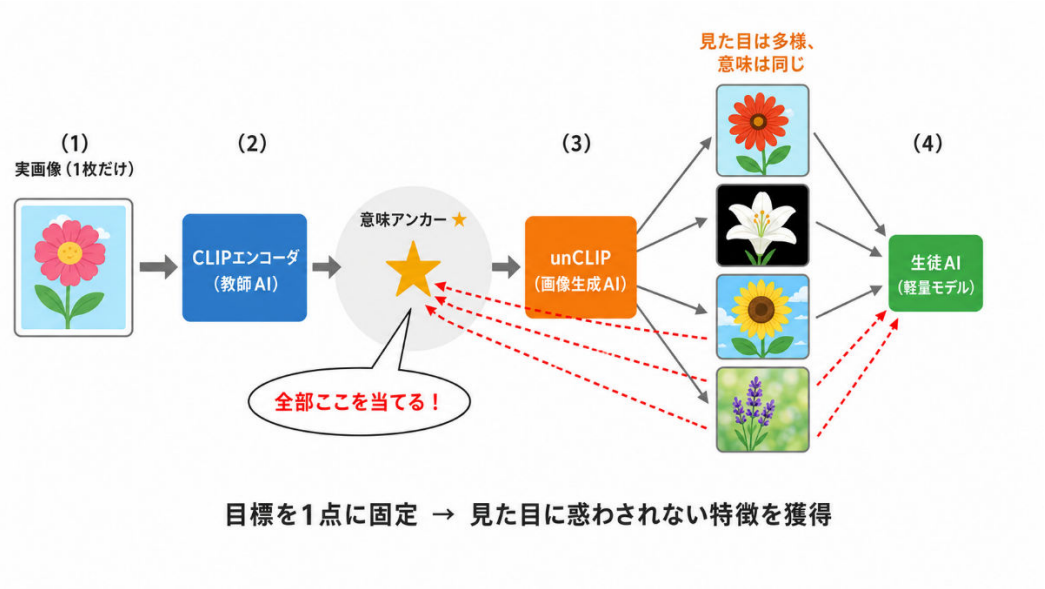


図２：InvAD の概要

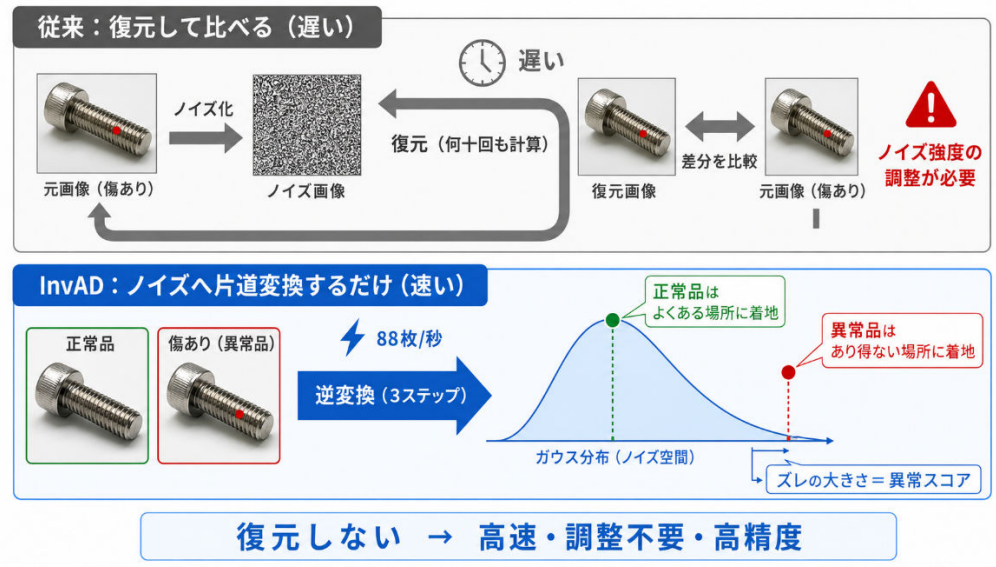
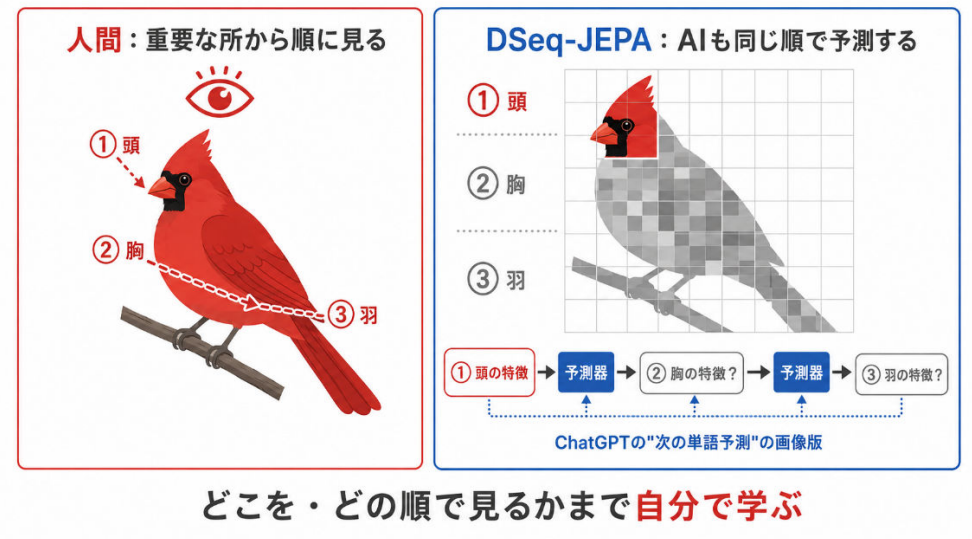


図 3：DSeq-JEPA の概要



〈用語解説〉

（注 1）画像生成 AI（拡散モデル）

テキストや画像などの条件から新しい画像を作り出す AI。本研究では、画像にノイズを加える過程を逆にたどり、ノイズから段階的に画像を生成する「拡散モデル」と呼ばれる方式（Stable Diffusion などで採用）を利用している。

（注 2）異常検知

正常なデータのパターンだけを学習し、そこから外れたデータ（製品の傷・欠陥や医療画像の病変など）を自動的に見つけ出す技術。異常データの収集が難しいため、正常データのみで学習する方式が主流となっている。

（注 3）自己教師あり学習

人手によるラベル付けを行わず、データ自身から学習の手がかり（例：画像の一部を隠し、残りの部分から隠された内容を予測させる）を作り出して学習する方法。ラベル付けコストを大幅に削減できる。

（注 4）知識蒸留

大規模で高性能な AI（教師モデル）が獲得した知識を、小規模で軽量な AI（生徒モデル）に受け渡す技術。教師の出力や内部の特徴を生徒に模倣させることで、軽量なまま高い性能を実現する。

（注 5）CLIP

米 OpenAI 社が開発した、画像とテキストを共通の特徴空間で結びつける代表的な基盤モデル。大量の画像とテキストの組で事前学習されており、幅広い画像認識タスクに応用できる。

（注 6）基盤モデル

膨大なデータで事前学習され、多様なタスクに汎用的に応用できる大規模 AI モデルの総称。

（注 7）STL-10 データセット

画像認識や機械学習の研究で用いられる公開画像データセット。飛行機や鳥などの 10 カテゴリーからなる 96×96 ピクセルのカラー画像を収録しており、学習用・評価用のラベル付き画像に加え、10 万枚のラベルなし画像を含む。少量のラベル付きデータと大量のラベルなしデータを活用する学習手法の評価などに利用される。

（注 8）潜在空間

AI がデータを内部的に表現するための特徴の空間。画像の見た目（画素値）そのものではなく、その意味や構造を圧縮して表現した空間を指す。

（注 9）I-JEPA

画像の見えている部分の特徴から、隠された部分の特徴を予測させる自己教師あり学習手法。Meta 社のヤン・ルカン氏らが提唱した JEPA（結合埋め込み予測アーキテクチャ）の画像版として広く用いられている。

論文 1（unCLIP-KD）
〈論文タイトル〉
英語タイトル "unCLIP-KD: A Generative Knowledge Distillation Framework Conditioned on CLIP Representations"（日本語タイトル：「unCLIP-KD：CLIP 表現を条件とした生成型知識蒸留フレームワーク」）

〈著者〉
英語表記
Tatsuya Sasaki,
Shunsuke Sakai,
Tatsuhito Hasegawa

日本語表記（所属先・職名）
佐々木達也（福井大学 大学院工学研究科 博士前期課程 2 年）
坂井俊介（福井大学 大学院工学研究科 博士後期課程 1 年）
長谷川達人（福井大学 学術研究院工学系部門 情報・メディア工学講座 准教授）

〈発表雑誌〉
雑誌名「Neurocomputing」（カタカナ表記：ニューロコンピューティング、Elsevier 社刊）（2026 年 11 月に掲載）巻数：701(7)巻、掲載ページ：pp. 1-15 ※採録決定・掲載後に追記
アブストラクト URL：
https://www.sciencedirect.com/science/article/pii/S0925231226018990
DOI 番号：https://doi.org/10.1016/j.neucom.2026.134501

論文 2（InvAD）
〈論文タイトル〉
英語タイトル "InvAD: Inversion-based Reconstruction-Free Anomaly Detection with Diffusion Models"
（日本語タイトル：「InvAD：拡散モデルの逆変換に基づく復元不要な異常検知」）

〈著者〉
英語表記
Shunsuke Sakai,
Xiangteng He,
Chunzhi Gu,
Leonid Sigal,
Tatsuhito Hasegawa

日本語表記（所属先・職名）
坂井俊介（福井大学 大学院工学研究科 博士後期課程 1 年）
Xiangteng He（ブリティッシュコロンビア大学／Vector Institute 研究員）
顧淳祉（福井大学 学術研究院工学系部門 情報・メディア工学講座 講師）
Leonid Sigal（ブリティッシュコロンビア大学 教授／Vector Institute）
長谷川達人（福井大学 学術研究院工学系部門 情報・メディア工学講座 准教授）

〈発表雑誌〉
会議名「IEEE/CVF Conference on Computer Vision and Pattern Recognition 2026（CVPR 2026)」（カタカナ表記：シーブイピーアール 2026。コンピュータビジョン分野における世界最高峰の国際会議）（2026 年 6 月に米国・デンバーで開催。国際会議のため巻数・掲載ページなし）
アブストラクト URL：https://arxiv.org/abs/2504.05662

論文 3（DSeq-JEPA）
〈論文タイトル〉
英語タイトル "DSeq-JEPA: Discriminative Sequential Joint-Embedding Predictive Architecture"
（日本語タイトル：「DSeq-JEPA：識別的逐次型結合埋め込み予測アーキテクチャ」）

〈著者〉
英語表記
Xiangteng He,
Shunsuke Sakai,
Kun Yuan,
Nicolas Padoy,
Tatsuhito Hasegawa,
Leonid Sigal （※Xiangteng He 氏と Shunsuke Sakai 氏は共同筆頭著者）

日本語表記（所属先・職名）
Xiangteng He（ブリティッシュコロンビア大学／Vector Institute 研究員）
坂井俊介（福井大学 大学院工学研究科 博士後期課程 1 年）
Kun Yuan（ストラスブール大学 研究員）
Nicolas・Padoy（ストラスブール大学 教授）
長谷川達人（福井大学 学術研究院工学系部門 情報・メディア工学講座 准教授）
Leonid Sigal（ブリティッシュコロンビア大学 教授／Vector Institute）

〈発表雑誌〉
会議名「European Conference on Computer Vision 2026（ECCV 2026)」
（カタカナ表記：イーシーシーブイ 2026。コンピュータビジョン分野における欧州最大級の国際会議）
（2026 年 9 月に開催予定。国際会議のため巻数・掲載ページなし）
アブストラクト URL：https://arxiv.org/abs/2511.17354

〈研究支援〉
JSPS 科学研究費助成事業　26K02879、25H01110、23K11164

〈お問い合わせ先〉
（研究に関すること）
長谷川達人（はせがわ たつひと）
国立大学法人福井大学学術研究院工学系部門情報・メディア工学講座
〒910-8507 福井市文京 3 丁目 9 番 1 号
TEL：0776-27-8929　E-mail：t-hase@u-fukui.ac.jp

（報道担当）
国立大学法人福井大学 広報センター
〒910-8507 福井市文京３丁目９番１号
TEL：0776-27-9733　E-mail：sskoho-k@ad.u-fukui.ac.jp