

2026 年 5 月 1 日

報道機関 各位

国立大学法人東北大学

**通信不要で賢く判断する AI へ
少量データで動く省電力技術を開発
—Few-shot 学習と知識蒸留で精度向上と消費電力削減を両立—**

【発表のポイント】

- わずかなデータでも高精度に認識可能な AI 技術を開発しました。
- 精度を最大 14%向上させつつ、大幅なモデルの軽量化を実現しました。
- 消費電力を約 37%削減し、現場で使える AI を実現しました。
- 通信不要で動作可能なエッジ AI^(注1)の実用化に貢献します。
- AI の課題である「大量データ依存」の解決につながる成果です。

【概要】

AI は近年急速に高性能化していますが、その一方で、大量の学習データや高い計算能力、さらには大きな電力消費を必要とすることが課題となっています。特に、通信環境が制限される現場やリアルタイム処理が求められる状況では、クラウドに依存しない軽量かつ省電力な AI 技術の実現が求められています。

東北大学 電気通信研究所の鬼沢直哉准教授、羽生貴弘教授の研究グループと、フランス IMT Atlantique 大学の Vincent Gripon 教授の研究グループは、こうした課題を解決するため、少量のデータから効率的に学習できる Few-shot 学習^(注2)と、大規模 AI モデルの知識を小型モデルに転移する知識蒸留^(注3)を組み合わせた新しい AI 技術を開発しました。

本手法では、軽量の AI モデルでありながら高い認識性能を実現し、画像認識の評価実験において、従来手法と比較して最大約 14%の精度向上を達成しました。また、実機評価では消費電力を約 37%低減しつつ、高速処理(2.6 ミリ秒)を実現しました。これらの成果により、本技術は通信環境に依存せず現場で動作可能な「エッジ AI」の実用化を前進させるものです。医療、製造、インフラ監視など幅広い分野での応用が期待されます。

本研究成果は、2026 年 5 月 4 日から 8 日に開催される信号処理に関する分野の国際会議 ICASSP 2026 において発表され、5 月 5 日に発表予定です。

【詳細な説明】

研究の背景

近年、人工知能(AI)は画像認識や自然言語処理などの分野で飛躍的な性能向上を遂げており、医療、製造、インフラ監視など幅広い分野での活用が進んでいます。しかしその一方で、こうした高性能なAIの多くは、大量の学習データと高い計算能力を前提としており、特にエッジ環境における実装には大きな課題が存在しています。エッジ環境とは、クラウドではなく端末側で処理を行う環境を指し、通信遅延の低減やプライバシー保護といった利点がある一方で、計算資源や電力に厳しい制約があります。このため、大規模なAIモデルをそのまま適用することは難しく、軽量で効率的なAI技術の開発が求められています。

さらに、多くの実世界の応用においては、大量のラベル付きデータを収集すること自体が困難であるという問題も存在します。例えば医療分野では、希少疾患のデータが限られているため、十分な学習データを確保することが難しい場合があります。このような状況において有効とされているのが「Few-shot 学習」であり、少量のデータから学習可能な技術として注目されています(図 1)。しかし、従来の Few-shot 学習では、軽量なモデルを用いた場合に精度が十分でないことが課題とされてきました。また、高精度なモデルは計算量が大きくなり、エッジ環境への適用が難しくなるというトレードオフも存在しています。

このように、エッジ環境での実用性と高精度を両立しつつ、少量データでも学習可能なAIの実現は、現在のAI研究における重要な課題の一つとなっています。本研究は、この課題に対して新たなアプローチを提案するものです。

今回の取り組み

本研究では、少量のデータから効率的に学習可能な Few-shot 学習(図 2)と、大規模なAIモデルの知識を小型モデルに転移する「知識蒸留」と呼ばれる技術を組み合わせることで、軽量かつ高精度なAIモデルの実現を目指しました。具体的には、従来広く用いられてきた ResNet^(注4)系のモデルに代えて、軽量でありながら高い表現能力を持つ MobileViT^(注4)と呼ばれる構造を採用し、その事前学習に知識蒸留を適用しました。この手法により、大規模モデルが持つ一般化能力を小型モデルに効率的に取り込むことが可能となります。

評価実験では、画像認識タスクにおいて本手法の有効性を検証しました。その結果、従来手法と比較して、少量データ環境における分類精度を大きく向上させることに成功しました。特に、1枚のサンプルのみで学習する設定において最大約14%の精度向上を達成しており、限られたデータ条件下でも高い性能を発揮できることが示されました。また、モデルのパラメータ数や計算量についても大幅な削減が実現されており、軽量化と高精度化を同時に達成しています。

さらに、本手法の実用性を検証するため、Jetson Orin Nano と呼ばれるエッジ向けデバイス上で動作検証を行いました。その結果、従来モデルと比較して消費電力を

約 37%削減するとともに、推論処理の高速化も確認されました(図3)。これにより、本手法が単なる理論的な改善にとどまらず、実際のハードウェア上でも有効であることが明らかとなりました。これらの成果は、エッジ AI における性能と効率の両立に向けた重要な一歩といえます。

今後の展開

本研究で開発した技術は、少量データ環境においても高精度な認識を実現しつつ、低消費電力で動作可能であることから、さまざまな実社会への応用が期待されます。例えば、医療分野においては、限られた症例データしか存在しない疾患の診断支援への応用が考えられます。また、製造業においては、不良品検知や設備異常の早期発見など、現場でリアルタイムに判断を行うシステムへの応用が期待されます。さらに、インフラ監視や災害対応など、通信環境が不安定な状況においても自律的に動作する AI としての活用も有望です。

今後は、本技術のさらなる高性能化とともに、FPGA などの専用ハードウェアへの実装を進めることで、より一層の省電力化および高速化を図る予定です。また、画像認識以外の分野への応用や、異なるデータ形式への拡張についても検討を進めていきます。さらに、実環境における長期的な運用を見据えた評価や、ユーザビリティの向上に関する研究も重要な課題となります。

本研究は、エッジ AI の実用化を加速させるとともに、AI 技術の適用範囲を広げる基盤技術として、今後の社会実装に大きく貢献することが期待されます。

従来のAIと本研究の違い

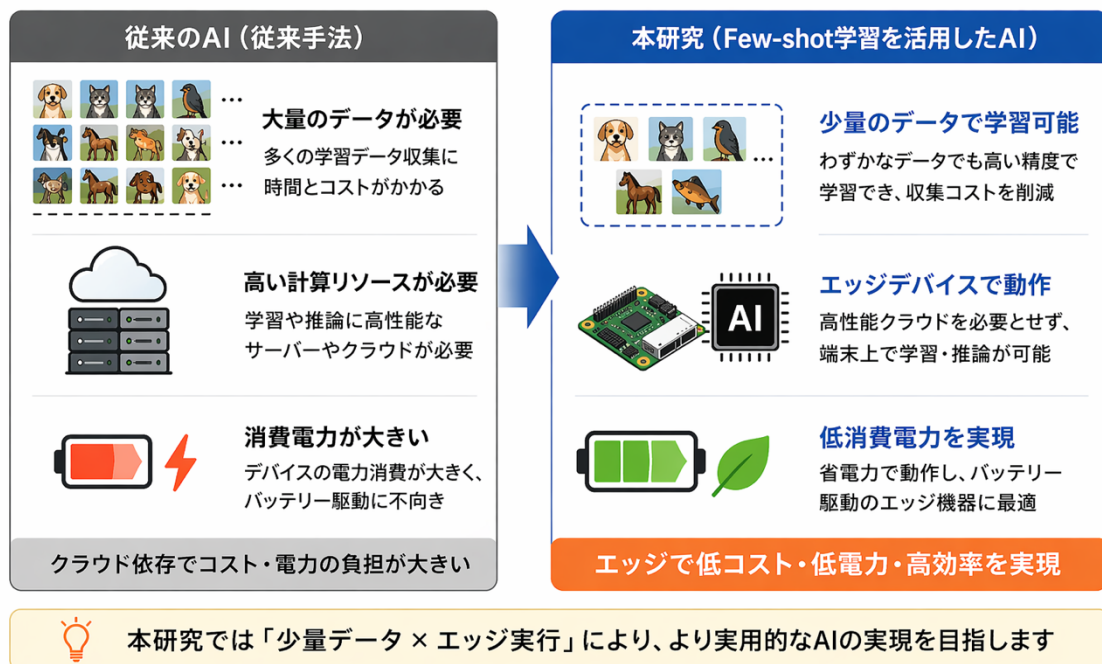


図 1. 従来 AI と本研究手法の比較

従来は大量データと高い計算資源を必要としたのに対し、本研究では少量データで学習可能かつ省電力で動作するエッジ AI を実現。

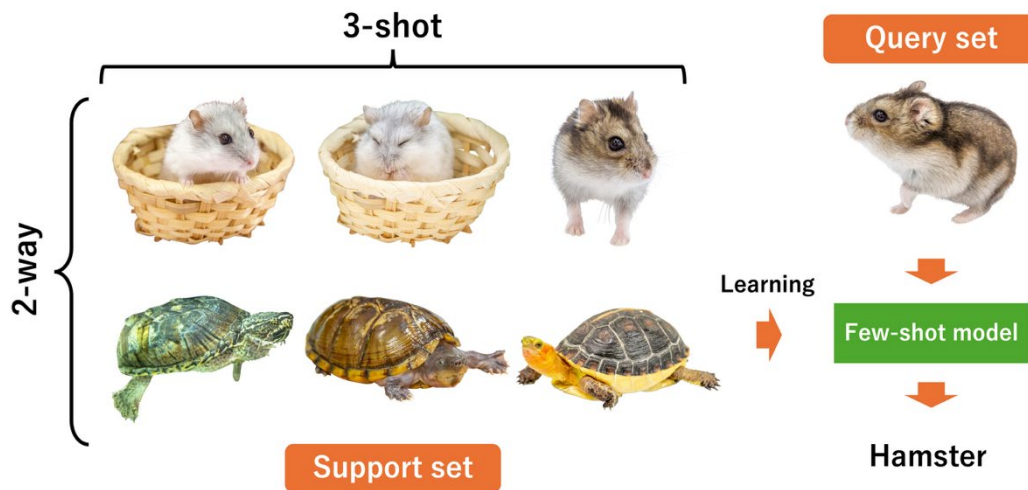


図 2. 少量データから学習する Few-shot 学習の概念図

少数のサンプル(サポートセット)から特徴を学習し、新しいデータ(クエリ)を分類する仕組みを示す。本研究では、この Few-shot 学習に知識蒸留を組み合わせることで、少量データでも高精度な認識を実現している。

本研究の成果（従来手法との比較）

Few-shot学習と知識蒸留を組み合わせた軽量AIモデルにより、精度向上と省電力化を実現

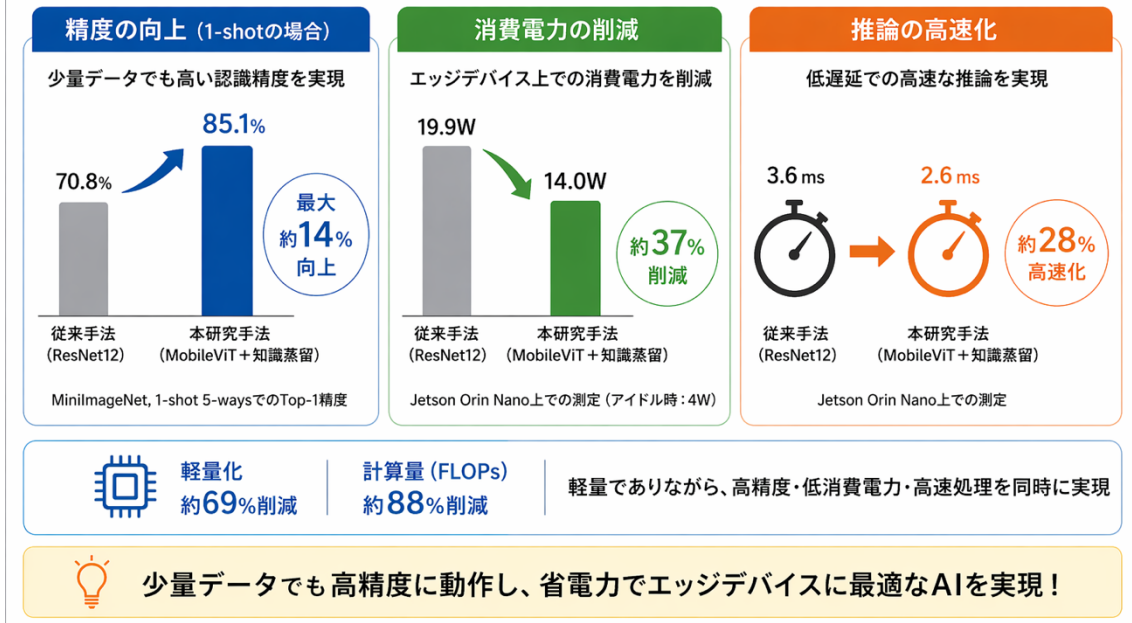


図 3. 本研究手法の性能向上および省電力化の効果

従来手法と比較して、精度向上（最大約 14%）と消費電力削減（約 37%）を達成し、高速処理も実現。

【謝辞】

本研究の一部は、公益財団法人 村田学術振興・教育財団の助成による研究として実施されました。

【用語説明】

注 1. エッジ AI

クラウドではなく、端末（エッジ）上で AI 処理を行う技術。

注 2. Few-shot 学習

少量のデータ（数枚～数十枚程度）からでも学習できる AI 技術。

注 3. 知識蒸留

大規模な AI モデルの知識を小型モデルに引き継ぐ技術。

注 4. ニューラルネットワークモデル (ResNet・MobileViT)

画像認識などに用いられる AI の基本構造。ResNet は従来広く使われてきたモデルであり、MobileViT は軽量で高性能な新しいモデル。

【論文情報】

タイトル : Efficient Few-Shot Learning for Edge AI via Knowledge Distillation on MobileViT

著者 : Shuhei Tsuyuki, Reda Bensaid, Jérémy Morlier, Mathieu Léonardon, Naoya Onizawa*, Vincent Gripon, Takahiro Hanyu

*責任著者 : 東北大学電気通信研究所 准教授 鬼沢直哉

掲載誌 : ICASSP 2026 (IEEE International Conference on Acoustics, Speech and Signal Processing)

URL: <https://2026.ieeeicassp.org/>

【問い合わせ先】

(研究に関すること)

東北大学電気通信研究所

准教授 鬼沢直哉

TEL: 022-217-5546

Email: naoya.onizawa.a7@tohoku.ac.jp

(報道に関すること)

東北大学電気通信研究所 広報室

TEL: 022-217-5427

Email: riec-kohoshitsu@grp.tohoku.ac.jp