AIエッジを強化するディープラーニング技術

玉井　秀明　　国定　恭史
川村　聡志　　山本　康平

AI、特にディープラーニングと呼ばれる機械学習手法を用いて作られたAIモデルの社会実装が進んでいる。OKIは、ディープラーニングへの取組みとして、現場の近くで稼働する業務特化型AI（AIエッジ）に注力している。

ディープラーニングをAIエッジ領域に適用するには、解決すべきさまざまな課題がある。以下、その課題を三つ挙げる。

一つ目の課題は想定外データ入力への対応である。入力されたデータをいくつかの種別（クラス）に分類する、いわゆる分類タスクの場合、クラスの数やその場合分けは過去の経験などから想定して予め設計される。しかし実際に運用を開始すると、想定したどのクラスにも属さないデータが入力されるおそれがある。その場合、新たなクラスを追加したモデルの再学習が必要となる。ユーザーにモデルの再学習を促すには入力されたデータが想定の範囲外であることを検知する必要がある。

二つ目の課題はラベル*1)付与作業負荷の軽減である。高い精度のディープラーニングモデルを開発するには、大量のデータ、特に正解ラベルが付与された教師データを大量に用意しなければならない。ラベルを付与する作業には多大なコストがかかり、AI導入の障壁となっている。従って、より少ない教師データで高い精度のモデルを作る手法が必要となる。

三つ目の課題はモデルの軽量化である。近年のディープラーニングモデルは、その動作に膨大なメモリーや演算能力を必要とする。しかしAIエッジ領域でAI処理（推論）を行うマシン（AIエッジデバイス）の性能は、クラウドのそれに比べて劣る場合がほとんどである。そのような環境下でも安定して推論させるには、ディープラーニングモデルを軽量化する必要がある。

本稿では、これらの課題を解決するために、OKIが研究開発中のディープラーニング技術である分布外検知、少データ学習、モデル軽量化の三つの技術を紹介する。

分布外検知

(1) 分布外検知とは

分布外検知とは、学習データの分布から外れたデータ（分布外データ）を検知する技術である。特に、未知のデータやモデルが意図していない入力データを検知することを目的としている。

ディープラーニングは、分布外データを入力した時に誤った予測をすることがある（**図1**）。例えば、動物の種別（犬･猫･馬）を分類するモデルに、学習していない動物（鳥･魚など）の画像を入力した場合、モデルは正しい予測をすることができない。ディープラーニングは、分布外データに対する挙動は保証されず、場合によっては、誤判定の予測に対して高い分類確率を割り当ててしまう可能性がある。

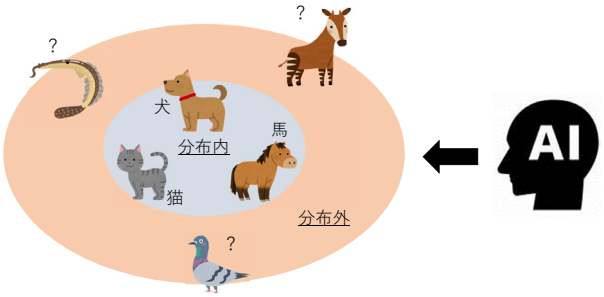

図1　分布内データと分布外データ

医療診断や不審者監視などのシステムでは、予測結果に高い信頼性が求められる。このようなシステムでは、分布外データが入力された際に誤った予測結果を返すのではなく、分布外データであることを検知し人に知らせる機能が求められている。誤判定によるトラブルを予防するためにも、分布外検知を行うことが重要である。

(2) 分布外検知の手法

分布外検知の手法としては、主に以下のものが提案されている。

- **分類ベース**

  分類モデルが出力するそれぞれのクラスの予測確率の最値が小さいものを分布外とする手法。学習済みの分類デルを転用できるので適用しやすい。一方で、分布外検知の精度が分類モデルの精度に依存するため、高精度な分類モデルが必要とされる。

- **密度ベース**

  学習データの分布の確率密度を、生成モデル*2)などを用いて近似し、密度が低いものを分布外とする手法。

*1) ディープラーニングの学習に必要となる、入力データに対する正解情報。　*2) 入力データの分布を推測し、その分布に従って画像などのデータをサンプリングすることができるAIモデル。

分類モデルには非依存だが、生成モデルの学習が難しく、チューニングに手間がかかる場合がある。

- **距離ベース**

学習済みモデルの中間層が出力する特徴マップを用いて学習データと入力データの距離を算出し、距離が大きいものを分布外とする手法。高精度な識別器を得るためは、少数の分布外データを用いた追加のフィッティングが必要になる場合がある。

(3)適用例:音響データに関する分布外検知

OKIでは、音響データに関する分布外検知技術に着目し研究開発を行っている。前述した既存手法は主に画像データの分類問題を扱っているが、音響データを扱った事例は少ない。音響認識は、工業製品の点検やソナーシステムなどで使用される重要な技術であるため、音響データに関する手法の開発が必要である。

そこで我々は、既存手法の中で最も高精度な手法[1]を音響データに適用した。これは、大規模画像データセット(ImageNet)によって事前学習したVision Transformerを利用することで、距離ベースの分布外検知精度を高める手法である。この手法を音響データに適用させるために、①事前学習には音響データの大規模データセットであるaudiosetを、②モデルにはVision Transformerをベースにした音響データ分類モデルであるAudio Spectrogram Transformer(AST)[2]を採用した(**図2**)。

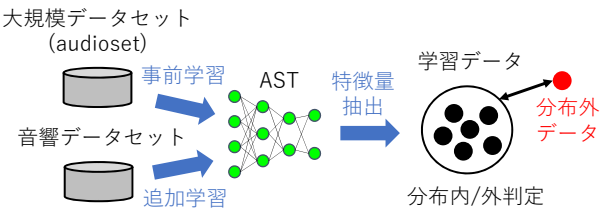

図2　提案手法

この提案手法により、OKIが保有する、ある機械に関する音を録音した音響データセットを用いて評価した。評価結果を**表1**に示す。結果として、提案法は既存法(ResNetを使用した距離ベース検知)に比べて7.0 pt(ポイント)程度の精度を改善されることが確認できた。

表1　分布外検知手法の評価結果

	既存法	提案法
精度	91.2 %	98.2 %

今回の評価で使用したデータセットは限定された環境条件で収集されたものであった。今後は、実運用の環境を考慮した、より実用的なデータセットを用いて評価し、さらに性能改善を進めていく予定である。

少データ学習

(1)少データ学習とは

少データ学習とは、ディープラーニングの学習に必要なデータの準備段階でのユーザー負担を削減しつつ、高い精度を維持するモデルを生成する技術の総称である。本稿では、少データ学習の一つである半教師有り学習を紹介する。

一般的なディープラーニングの学習では、入力データとラベルの組を大量に収集する必要がある。しかし、ラベルの付与作業は主に人手で行われ、大量の入力データにラベルを付与することはユーザーにとって大きな負担となる。この課題解決のため、半教師有り学習では、収集した入力データの一部にのみラベルを付与し、残りはラベルが存在しない状態で学習する。

画像認識タスクの半教師有り学習手法で、大きな成果を残したFixMatch[3]は、弱いデータ拡張*3)を施したラベル無しデータを推論し擬似的なラベルを作成した後、強いデータ拡張を施したラベル無しデータとの誤差を算出し、学習する手法である。

一般的な半教師有り学習では、ラベル付きデータとラベル無しデータはそれぞれ同一の分布から得ることを前提としている。しかし、ラベルを付与する入力データの選択にはユーザーの任意性があるため、この前提が担保されているとは限らない。また、Q.Wangら[4]は、ラベル付きデータが少数の場合、これらの経験分布*4)の不一致度が大きくなることを示している。すなわち、ラベル付きデータから得られる知識をラベル無しデータに十分に活用することができないと考えられる。これらの結果、認識精度にばらつきが生じるおそれがある。

そこで我々は、この認識精度のばらつきを軽減するため、ドメイン適応のコンセプトを半教師有り学習に応用する手法を提案する。ラベル付きデータとラベル無しデータを異なるデータ分布と考え、ドメイン適応の特徴である二つの分布を近づける仕組みを取り入れる。

(2)ドメイン適応を適用した半教師有り学習

提案手法について説明する。本稿では、ディープラーニングを用いた画像認識タスクを扱う。半教師有り学習にはFixMatch, ドメイン適応には DANN[5]を使用する。提案手法の概要図を**図3**に示す。

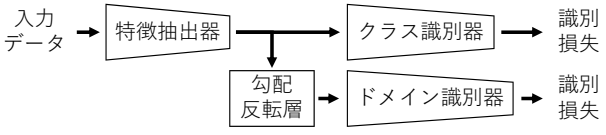

図3　提案手法の概要図

*3)画像の回転・拡大・色変換など、入力データを変換する作業。　*4)実際に得られたデータの分布。

まず、図3の特徴抽出器とクラス識別器は、画像認識タスクを行うためのモデルであり、FixMatchにより学習する。一方、特徴抽出器から出力される特徴量が、勾配反転層、ドメイン識別器の順に入力され、該当データがラベル付きデータかラベル無しデータかに判別される。ここで勾配反転層は判別結果を誤差逆伝播するときのみ作用し、誤差から得られる勾配を反転する。この作用により、ラベル付きかラベル無しかを判別できるほど、特徴抽出器はこれらを判別できないように敵対的に更新され、結果としてこれらの経験分布の不一致度を小さくできる。

手書き数字データセットMNISTを用いて識別精度のばらつき軽減効果を検証した。モデルとして特徴抽出器が畳み込み層2層、クラス識別器とドメイン識別器がそれぞれ全結合層2層で構成されたものを使用した。なお、全ての試行で、モデルの初期値は同じである。MNISTは手書き数字画像が学習用として合計6万枚用意されているが、そのうちの30枚にのみラベルを付与した。この30枚の組を5パターン用意し、それぞれの最高識別精度のばらつきを確認した。比較として、ラベルを付与した30枚のみを使用して学習する「教師有り学習」、「FixMatch」、「提案手法」でそれぞれ5パターンの学習を行った。

実験結果を表2に示す。識別精度は平均±標準偏差の形で表示する。提案手法の平均精度がFixMatchよりも0.7pt 増加し、標準偏差が1.9pt 減少したことから、ドメイン適応のコンセプトを応用することで,ラベルを付与したデータの組の違いによる精度のばらつきが軽減されていることが分かった。今後の課題として、更に精度向上やばらつき軽減できる技術の研究開発を進めていく。、ラベルを用意しなくても質の高い特徴抽出器を生成する教師無し学習の適用も検討していく。

表2　検証データセットでの実験結果

手法	識別精度
教師有り学習	71.1±6.2 %
FixMatch	95.3±7.5 %
提案手法	96.0±5.6 %

モデル軽量化

(1)モデル軽量化とPCAS*5)技術

モデル軽量化技術とは、モデルの精度を最大限維持しつつパラメーター数や演算回数を低減する手法の総称である。近年のディープラーニングモデルは、その動作に膨大なメモリーや演算能力を必要とすることからモデル軽量化技術の必要性が高まっている。

AIエッジデバイスなどの処理能力の限られる実行環境では、学習よりも演算リソースが少なくて済む推論機能だけを実装するのが一般的であるが、それでも高精度なモデルをAIエッジデバイス上で動作させることは難しい。その理由は、高精度なモデルほど膨大なパラメーター数や演算量を必要とする傾向があるためである。そこで、モデル軽量化技術を適用することにより、それらの制約を軽減し高精度なモデルの推論機能をAIエッジデバイス上で高速に動作させることができる。

OKIは独自のモデル軽量化技術であるPCAS技術を保有している[6),7)]。PCAS技術は、ディープラーニングモデル内に存在する不必要な演算を特定及び削減し、推論処理を高速化することができる。具体的には、PCAS技術は大規模なモデルのスクラッチ学習*6)後に重要度の低いニューロンを削減する「プルーニング（枝刈り）」と呼ばれるモデル軽量化手法の一種であり、プルーニングはニューロン削減後の精度劣化を回復させるためにファインチューニングと呼ばれる再学習プロセスを必要とする。

(2)プルーニングの実用上の課題

モデル開発の現場では、プルーニングのようなモデル軽量化技術を自前のモデルに適用するには、学習用ソースコードの大幅な修正が必要という課題がある。例えば、モデル内でプルーニングできる構造の特定や、ニューロンの重要度評価、重み係数の削減処理、ファインチューニング処理などを実装する必要があるほか、近年のディープラーニングモデルは多様な分岐結合構造をもつものが多く、各パターンに適した削減アプローチを用意する必要がある。従って、実装コストが高くモデル軽量化は手軽に適用しにくい。

また、モデル軽量化技術を効果的に適用するために、軽量化に関する専門知識が必要となる点も課題である。例えば、プルーニングではニューロンの削減率を層ごとに設定する必要があるが、通常その削減率は専門家によるモデルの分析作業によって適切な値を設定することが多い。さらに、より層数の多い大規模モデルに適用する場合には、削減率の必要設定数が多くなり、作業時間及び難度が飛躍的に高まってしまう。

(3)モデル軽量化ツールの提供

上記の二つの課題に対応するために、OKIは前述のPCAS技術を実装したモデル軽量化ツールの提供を予定している。本ツールは、ディープラーニングモデルの学習と軽量化をワンストップで実行する特長をもつ。具体的には、スクラッチ学習とファインチューニングの両方に対応したモデルの学習機能と軽量化機能の両方を搭載することでソースコード修正がほぼ不要となり、さらに、PCAS技術の特性を利用したニューロン削減率の自動化機能により軽量化の専門知識が不要で適用できる。また、軽量化後でも

*5)PCAS:Pruning Channels with Attention Statistics　*6)ニューラルネットワークの重みパラメーターなどにランダムな初期値を割り当て、一から学習を開始すること。

高い精度維持が期待できる「段階的な軽量化」や、モデルの分岐合流構造に対応する「分岐最適化」など有効なオプション機能を複数備えることで、軽量化のパフォーマンスを高めている。

昨今のモデル開発はオープンソース・ライブラリーを利用して実行されることが多いが、本ツールは、PyTorch[8]で実装された画像分類や物体検出、セマンティックセグメンテーション用のモデルを生成する強力なオープンソースとの連携でき、多様な種類のモデルの軽量化に対応している。また、モデルの学習機能はハイパーパラメーター*7)の最適化ライブラリーにも対応し、高い認識精度を得るために必要なパラメーターチューニング作業の自動化もできる。

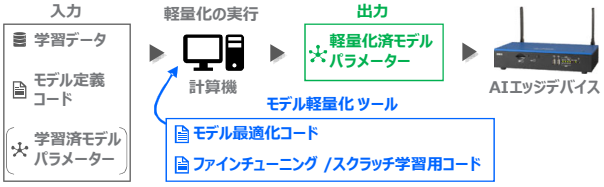

図 4　モデル軽量化ツールの使用例

ツール使用例としては、図4のように、まずユーザー側でモデル学習できる計算機環境及び学習用データ、モデル定義コードを用意し、諸設定を行った上でツールを実行すると、学習と軽量化が自動実行され、軽量化済みモデルパラメーターが生成される仕組みである。ただし、モデル定義コードは連携しているオープンソース・ライブラリーに同梱のものを使用する場合は不要である。

(4) 今後の展望

今後は、本ツールの提供を通じて、AIエッジ領域のモデル開発に貢献していく予定である。なお、ツール自体の改良も進め、対応するモデル種類を更に拡充し、量子化技術との併用による高効率化などを図っていきたい。

まとめ

OKIが研究開発中のディープラーニング技術である分布外検知、少データ学習、モデル軽量化の三つの技術を紹介した。今後も継続してAIの社会実装進展に資する技術を開発していく。 ◆◆

■参考文献

1)Fort, S., Ren, J., Lakshminarayanan, B.: Exploring the limits of out-of-distribution detection, Advances in Neural Information Processing Systems 34, 2021.

2)Gong,Y.,Chung,Y.-A.,Glass,J.:“AST:Audio spectrogram transformer”, Interspeech 2021,pp.571-575,2021.

3)Sohn,K.,et al.:“Fixmatch:Simplifying semi-supervised learning with consistency and confidence”,Advances in Neural Information Processing Systems,33,pp.596-608, 2020.

4)WANG,Q.,et al.:Semi-supervised learning by augmented distribution alignment, Proceedings of the IEEE/CVF international conference on computer vision, pp.1466-1475,2019.

5)Ganin, Y., et al.: Domain-adversarial training of neural networks,The journal of machine learning research,17, 1,pp.2096-2030, 2016.

6)Kohei Yamamoto,Kurato Maeno:“PCAS:Pruning Channels with Attention Statistics for Deep Network Compression”, British Machine Vision Conference (BMVC), September 2019.

7)山本康平、橘素子、前野蔵人:ディープラーニングのモデル軽量化技術、OKIテクニカルレビュー第233号、Vol.86 No.1、pp.24-27、2019年5月

8)Paszke, Adam et al.:“PyTorch:An Imperative Style, High-Performance Deep Learning Library”,Advances in Neural Information Processing Systems 32,pp.8024-8035, 2019.

●筆者紹介

玉井秀明:Hideaki Tamai. イノベーション推進センター AI技術研究開発部

国定恭史:Yasufumi Kunisada. コンポーネント＆プラットフォーム事業本部 開発本部 新規技術開発部

川村聡志:Satoshi Kawamura. イノベーション推進センター AI技術研究開発部

山本康平:Kohei Yamamoto. イノベーション推進センター AI技術研究開発部

【基本用語解説】

ドメイン適応
CG画像で学習し、実写画像で運用する場合など、あるドメインで学習した知識をもちいて、別ドメインでも高精度な推論ができる学習手法。

ニューロン
ニューラルネットワークを構成する基本的な要素。

モデル定義コード
ニューラルネットワークのモデル構成や動作を記述したソースコード。

*7)モデル学習において、学習率などユーザーが予め決定する必要のあるパラメーター。