

組み込みシステムにおける深層学習モデル応用の試み AN ATTEMPT TO APPLY DEEP LEARNING MODELS IN EMBEDDED SYSTEMS

劉 忠達*, 佐々木 慶文*, 川村 暁†

Zhongda Liu^{A)}, Yoshifumi Sasaki^{A)}, Satoshi Kawamura^{B)}

^{A)}Faculty of Science and Engineering, Ishinomaki Senshu University

^{B)}Faculty of Humanities, Morioka University

Abstract

In the field of image recognition, deep learning methods have made a marked advance. They have achieved more higher accuracy than previous methods. With the achievement of deep learning, it is expected to apply those methods to embedded system. However, many embedded systems generally have restrictions on hardware such as computing power and data storage capacity. In this work, regarding garbage separation issue, we created deep learning models to classify images using the transfer learning method and tried to apply deep learning models to embedded systems. We have applied models of different sized to not only the stand-alone embedded system but also the embedded system based on cloud computing. According to evaluation experiments, we found that it is hard to apply a large-scale deep learning model to the stand-alone system because of hardware constraints, and the embedded system based on cloud computing is not affected by those hardware constraints.

1. はじめに

2012年に提案された深層学習モデル AlexNet [1] は、画像認識において、それまでの手法を遥かに上回る認識精度を達成した。AlexNet の登場以降、深層学習は目覚ましい進歩を遂げている。例えば、Google 社の AlphaGo はハンディキャップ無しで有名なプロ囲碁棋士に勝利している。このような深層学習の進歩に伴い、自動運転や災害対応ロボットなどの組み込みシステムへの応用に期待が高まっている。しかし、一般に、組み込みシステムには計算能力や記憶容量などのハードウェア制約があり、既存の深層学習モデルをそのまま応用することは、性能的な側面から困難であると考えられる。

そこで、本研究では、既存の深層学習モデルから移転学習に基づき、組み込みシステム向けに再学習した深層学習モデルを作成し、応用することを検討した。一方で、最新の 5G などの高速ネットワークに着目し、クラウドサーバを利用することにより、組み込みシステムのハードウェア制約を受けにくい画像認識システムについても検討した。

本稿では、我々が日常生活で直面している問題である、ゴミの分別問題への応用について述べる。24 種類のゴミについて深層学習モデルを作成し、このモデルに基づく自動分別を行った。移転学習に基づく再学習には膨大な計算が必要であるため、組み込みシステム向けの深層学習モデルの作成は、高速な汎用計算機により行った。得られた深層学習モデルに基づくゴミの分別アプリを小型のシングルボードコンピュータによるスタンドアローンシステムに実装し、その認識時間を評価した。また、同様のアプリをクラウドサーバと Android 端末の連携によるシステムに実装し、その性能をスタンドアローンシステムと比較検討した。

2. 深層学習モデルの作成

画像の分類問題においては、畳み込みニューラルネットワーク (CNN:Convolutional Neural Network) に基づく深層学習の性能が高いことが知られている [1]。図 1 に CNN の構造を示す。CNN は、数層から数百層以上の畳み込み層とプーリング層を交互に積み重ねた後に、数層のパーセプトロンを接続した構造である。畳み込み層は画像の特徴を抽出し、プーリング層は抽出された特徴の移動不変性を維持する。抽出された特徴を用いて、最終段のパーセプトロンにより画像の分類が行われる。

高精度の CNN のモデルを新たに作成するには、膨大な数の高品質画像データを用意しなければならず、また、超大規模の高速計算機により長期間の学習を行う必要がある。これらの難点を克服するために本研究では、公開されている既存の学習済みモデルから、再学習によって新たなモデルを作成する移転学習法を採用した。

2.1 移転学習に基づく深層学習モデルの作成

表 1 に示すような規模の異なる 3 種類の学習済みモデル LeNet [2]、AlexNet [1] と VGG [3] を用いて移転学習を行う。これらのモデルは、公開されている大規模な画像データセット ImageNet [4] で学習されたものである。

本研究では、畳み込み層とプーリング層は、その構造を維持し、一部の層に学習済みのパラメータを適用する。最終段のパーセプトロン層は、ゴミの分別に適した構成に変更して再学習を行った。

Table 1: 移転元の学習済みモデル

CNN	層の数	パラメータ数
LeNet	6	6,191,424
AlexNet	13	58,379,672
VGG	23	122,887,256

* 石巻専修大学理工学部

† 盛岡大学文学部

組み込みシステムにおける深層学習モデル応用の試み

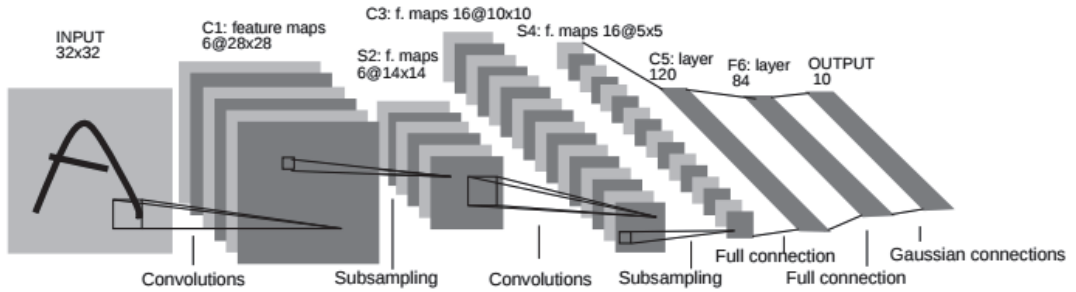


Figure 1: 畳み込みニューラルネットワーク LeNet の構造 出典 [2] P.7



Figure 2: 収集したゴミ画像のサンプル

学習と評価に用いるゴミ画像としては、インターネットで公開されている空き瓶、ペットボトルなど 24 種類のゴミ各 20 枚、計 480 枚の画像を用いている。(図 2)、収集したゴミ画像データを、8 対 2 の割合で訓練データ (384 枚) とテストデータ (96 枚) とした。訓練データにおいて、画像の量が少ないため、元画像を反転、回転や視点変換などの画像処理を行い、学習に用いるデータ数を増やしている。これらの画像を用いて高速の GPU ワークステーションを用いて再学習を行い、新たなモデル LeNet*、AlexNet* と VGG* を作成する (表 2)。

2.2 作成した深層学習モデルの性能評価

訓練データとテストデータを用いて作成したモデルによるゴミの分別を行い、その性能を評価した。表 2 に評価結果を示す。すべてのモデルにおいて訓練データの正解率は 0.98 以上となることが確認された。テストデー

タに対する正解率は、LeNet* と AlexNet* では 0.6 以下であったのに対し、VGG* では 0.9 程度と最も高くなった。テストデータに対しては、パラメータ数が最も多い VGG* の認識率が比較的高くなるという結果が得られた。

Table 2: 深層学習モデルの認識性能

モデル	ファイルサイズ	正解率 ^a	正解率 ^b
LeNet*	14MB	1.00	0.54
AlexNet*	218MB	0.98	0.56
VGG*	533MB	1.00	0.90

^a 訓練データに対する認識率

^b テストデータに対する認識率

3. 組み込みシステムへの応用

IoT (Internet of Things) といわれるように、組み込みシステムはインターネットに接続し、外部のシステムと連携して処理を行うことが一般的になっている。一方で、災害時等インターネットを利用できない、回線が逼迫している場合もある。そこで、スタンドアローンシステムとクラウドサーバとの連携システムの、2 つの異なる環境での深層学習モデルの応用を試みた。

3.1 スタンドアローンシステムでの応用

Table 3: 小型シングルボードコンピュータの仕様

	Raspberry Pi 3 Model B	Raspberry Pi 4 Model B
Soc	BCM2837 CPU: Cortex-A53 4 core 1.2GHz	BCM2711 CPU: Cortex-A72 4 core 1.5GHz
RAM	1GB (DDR2)	4GB (LPDDR4)
OS	Raspbian GNU/Linux 10 (Buster)	

スタンドアローンシステムには、表 3 で示した小型シングルボードコンピュータ Raspberry Pi 3 Model B と Raspberry Pi 4 Model B を用いた。

節 2.1 で作成した 3 種類のモデルについて、それぞれ分別実験を行い、処理時間を評価した。具体的には、深層学習モデルを読み込んでメインメモリ上に展開し、節 2.2 と同様に、テストデータを用いてゴミの分別を行うプログラムを作成し、その処理時間を測定した。

Table 4: スタンドアローンシステムの処理時間

モデル	Raspberry Pi 3		Raspberry Pi 4	
	リード時間	識別時間	リード時間	識別時間
LeNet*	0.14s	417ms	0.11s	216ms
AlexNet*	10.27s	179ms	1.12s	86ms
VGG*	-	-	2.70s	1.03s

評価結果を表 4 に示す。リード時間は深層学習モデルを読み込んで、メモリ上に展開する時間である。識別時間は CNN に 1 枚の画像を入力してから識別が完了するまでの平均時間を指す。モデルサイズ、すなわちパラメータ数が多くなるにつれてリード時間が増える。メモリ容量が 1GB である Raspberry Pi 3 では、モデルサイズが 500MB 以上の VGG*を読み込むことができなかった。これに対して、メモリ容量が 4GB である Raspberry Pi 4 は、すべてのモデルを読み込んで実行することができた。最も規模の大きい VGG*では、Raspberry Pi 4 でも 1 秒以上の識別時間を要した。以上のことから、ハードウェア制約があるスタンドアローンシステムに、規模の大きい深層学習モデルを応用することには限界があることが明らかになった。

次に、それぞれの深層学習モデルの実行時間について検討する。AlexNet* は、LeNet*に比べてニューラルネットワークの規模が大きく、また、畳み込み層にサイズが大きいカーネル（フィルタ）を使用しているにもかかわらず、識別時間が最も短い。このことから、カーネルサイズやニューラルネットワークの規模だけでなく、その構造も深層学習モデルの実行時間に影響を与える可能性があることがわかった。

3.2 クラウドサーバに基づくシステムでの応用

通信技術の進歩によって、組み込みシステムの構築は大きな変化を遂げようとしている。次世代通信規格である 5G は 10Gbps の通信速度と 1ms の超低遅延を実現できる [5] [6]。これによって、組み込みシステムはこれまでの計算能力やメモリ容量の制約を突破することが期待されている。

ここでは、ゴミの分別システムを次のように構築する。すなわち、深層学習モデルを応用した分別エンジンは、高い計算能力と大容量の記憶領域を有するクラウドサーバで実現し、利用者へのフロントエンドは Android 端末向けアプリで実現する。

表 5 で示したクラウドサーバ上には、節 2.1 で作成した深層学習モデルに基づきゴミを分別するためのサーバプログラムを実装する。組み込み端末からの画像を受信すると分別処理を行い、その結果を端末に返信する。

フロントエンドとして、図 4 に示すようなアプリを開

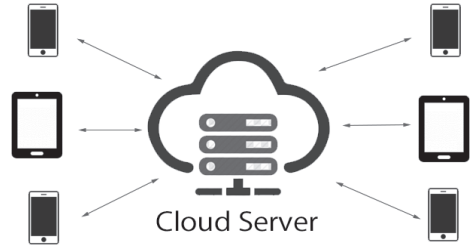


Figure 3: クラウドサーバによる分別エンジンの実現

発し、表 6 に示す Android 端末にインストールしておく。このアプリは、カメラで撮影したゴミの画像をクラウドサーバに送信し、その分別結果を受信し、ユーザに対して通知する。

Table 5: クラウドサーバの仕様

OS	Ubuntu 18
CPU	Intel i7-6900K 16 core 3.2GHz
GPU	NVIDIA TITAN V 12GB
メモリ	64GB

Table 6: Android 端末の仕様

OS	Android 6.0
CPU	Huawei Kirin 950 8 core (4x2.3GHz + 4x1.8GHz)
RAM	4GB

インターネットを模したネットワークを再現するため、有線のギガビットイーサネット (1000BASE-T) と無線 LAN (IEEE 802.11n) を用いて、表 5 のサーバと表 6 の Android 端末を接続し、性能評価実験を行った。

Table 7: クラウドサーバに基づくシステムの評価結果

モデル	識別時間	全体時間
LeNet*	36ms	979ms
AlexNet*	9ms	952ms
VGG*	101ms	1044ms

節 2.1 で作成した 3 種類の深層学習モデルの評価結果を表 7 で示した。識別時間は CNN に 1 枚の画像を入力してから識別が完了するまでの平均時間である。クラウドサーバに基づくシステムは、スタンドアローンシステ



Figure 4: Android アプリの画面

ム (表 4) と比較して、10～20 倍程度、高速であることが確認された。

また、全体の処理時間から識別時間を差し引いた時間は通信に要した時間であるが、送信画像のサイズが大きく、複数回に分割して送信しているため、約 943ms となっており、改善する余地がある。しかし、全体の処理時間はスタンドアローンシステムでの VGG* の識別時間と同程度であり、組み込みシステム特有の制約の影響は受けないシステムであると言える。また、さらに高精度かつ大規模の深層学習モデルを利用するシステムへの拡張も、比較的容易に可能であると考えられる。

4. まとめ

本研究では、日常生活で直面をするゴミの分別を組み込みシステムで実現するための、深層学習モデルの応用について述べた。公開されている 3 種類の深層学習モデルから、移転学習の手法に基づき、組み込みシステム向けの深層学習モデルを作成し、応用する方法を検討した。

スタンドアローンシステムとクラウドサーバに基づくシステムについて、規模の異なる 3 種類の深層学習モデルを作成し、実験的に識別時間や通信時間などの評価を行った。

その結果、ハードウェア制約があるスタンドアローンシステムでは、規模の大きい深層学習モデルを応用するのは難しいことがわかった。一方で、クラウドサーバに基づくシステムでは、組み込みシステム特有のハードウェア制約の影響を受けないため、高精度かつ大規模な深層学習モデルの応用が期待される。

謝辞

本研究の一部は IS 奨学研究助成を受けた。ここに謝意を表する。

参考文献

- [1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [2] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [3] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [4] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009, pp. 248–255.
- [5] P. Popovski, "Ultra-reliable communication in 5g wireless systems," in *1st International Conference on 5G for Ubiquitous Connectivity*. IEEE, 2014, pp. 146–151.
- [6] A. A. Esswie and K. I. Pedersen, "Opportunistic spatial preemptive scheduling for urllc and embb coexistence in multi-user 5g networks," *Ieee Access*, vol. 6, pp. 38 451–38 463, 2018.