

深層学習によるゴミ分別支援システムの構築 Construction of trash classification support system based on deep learning

劉 忠達[†]
Zhongda Liu

あらまし

日常生活で常にゴミの分別問題に直面している。近年、人工知能技術が飛躍的に進展した。特に深層学習は目覚ましい進歩を遂げている、様々な分野で大きな成果を上げた。画像認識において、人間以上の識別精度に達した。本研究では、物体をスマートフォンで撮影し、写真から自動的にゴミを分別する支援システムを構築する。インターネット上で公開されているゴミの画像を収集し、高速な汎用計算機により深層学習を行った。得られた深層学習モデルがクラウドサーバに應用され、モバイル端末からの画像を受信するとゴミ分別処理を行い、その結果を端末のアプリに返信してユーザに通知する。

1. はじめに

日常生活で常にゴミの分別問題に直面している。1人1日当たりのゴミ排出量は918グラムで年々減少傾向であるが、リサイクル率は僅か19.9%である[1]。リサイクル率を高めるためにゴミを正しく分別する必要がある。しかし、ゴミ分類の категорияが多く、市町村ごとに分別のルールが異なっている。手軽にゴミを分別できる支援システムが必要と考えられる。

近年、人工知能技術が飛躍的に進展した。特に深層学習は目覚ましい進歩を遂げている、様々な分野で大きな成果を上げた。画像認識において、人間以上の識別精度に達した。本研究では、物体をスマートフォンなどの端末で撮影し、写真から自動的にゴミを分別する支援システムを構築する。モバイル端末は、計算能力、メモリ容量などリソースによる制約があるため、そのリソースで直接に深層学習を應用することが困難である。

通信技術の進歩によって、モバイル端末への應用は大きな変化を遂げようとしている。本研究では、最新の5Gなどの高速なインターネットに着目し、リソース的な制約を受けないシステムについて検討した。

本稿では、インターネット上で公開されているゴミの画像を収集し、高速な汎用計算機で規模の異なる3種類の畳み込みニューラルネットワークを用いて深層学習を行った。学習結果を評価し、画像によってゴミの自動分別の可能性を確認した。

そして、クラウド上に高い計算能力と大容量の記憶領域を有するサーバー設置し、学習で得られた深層学習モデルを高性能のサーバーで実現する。モデル端末から撮影された画像を受信するとゴミ分別処理を行い、その結果を端末のアプリに返信してユーザに通知する。

2. 深層学習

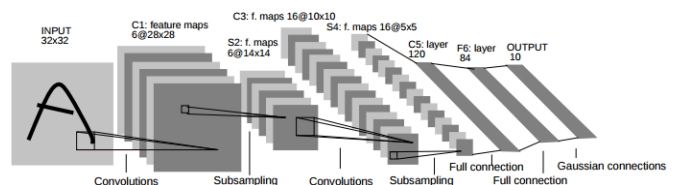


図1 畳み込みニューラルネットワーク LeNet の構造 出典[2]

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[3]。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。

2.1 学習データの作成

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 枚の画像を収集した (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とテストデータの2つに分ける。画像の量が少ないため、訓練データにおいて元画像を反転、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。訓練データを用いて GPU ワークステーションで深層学習を行う。テストデータを用いて学習したモデルの汎化能力を評価する。

[†] 石巻専修大学 理工学部

Faculty of Science and Engineering, Ishinomaki
Senshu University



図 2 収集したゴミ画像のサンプル

2.2 学習環境

表 1 GPU ワークステーションの仕様

CPU	Core i7 6900k 8core 16thread 3.2GHz
GPU	GeForceGTX 1080Ti 11GB × 2
Memory	48GB DDR4
OS	Ubuntu 14.04

表 1 の GPU ワークステーションを用いて、CNN の学習実験を行う。

2.3 学習実験

表 2 3 つ CNN の比較

	層数	パラメータ数
LeNet	6	6,191,424
AlexNet	13	58,379,672
VGG	24	20,024,384

DeepLearning4J[4]という Java の深層学習ライブラリを用いて、LeNet[2]、AlexNet[3]と VGG[3] 規模の異なる 3 つの CNN (表 2) を実装した。

そして、表 1 の GPU ワークステーションを用いて学習実験を行った。層数が少ない LeNet と AlexNet は順調に学習ができた。しかし、VGG のような層数が多く大規模の CNN は、順調に学習を行うために膨大な数の高品質画像データを用意しなければならない。2.1 の訓練データで学習を行うと図 3 の上部分の学習曲線で示したように長期間に学習を行っても損失関数の値が中々収束しない。この難点を克服するために本研究では、移転学習法を採用した。公開されている既存の学習済みモデルから、VGG のパラメータを取り入れ、再学習を行った (図 3 の下部分)。

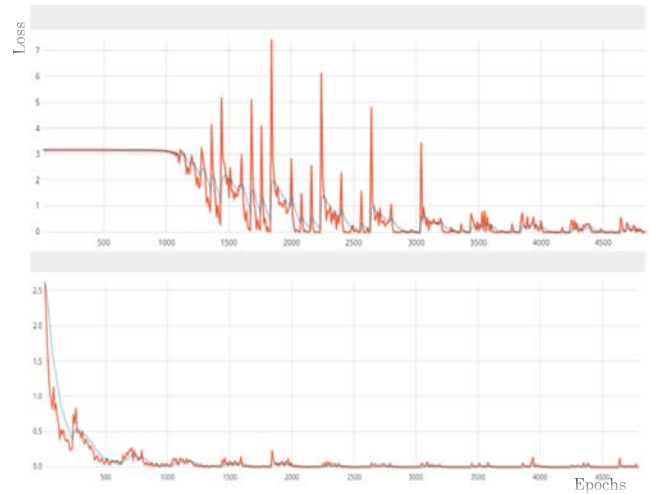


図 3 VGG の学習曲線の比較 (下は移転学習)

2.4 学習実験結果

表 3 学習実験結果

CNN	層数	訓練データの正解率	テストデータの正解率	モデルサイズ(MB)
LeNet	6	1.00	0.54	38
AlexNet	13	0.98	0.56	440
VGG	23	1.00	0.90	910

表 3 で学習実験結果を示した。すべての CNN において、訓練データの正解率は 0.98 以上になり、学習した画像を認識することができた。学習していないテストデータに対して、LeNet と AlexNet の正解率は 6 割以下、VGG は最も高く 9 割となっている。VGG は学習していない画像でも認識ができることが分かった。

これで画像によってゴミの自動分別の可能性が確認されたので、実験で学習した CNN のパラメータをモデル (以下、学習モデル) としてファイルに保存した。学習モデルのサイズは表 3 に示したようになっている。

3. ゴミ分別支援システム

表 4 Android 端末の仕様

OS	Android 6.0
CPU	Huawei Kirin 950
RAM	4GB

モバイル端末は、計算能力、メモリ容量などのリソースによる制約がある。CNN の学習モデルを読み込んでモバイル端末のメモリに展開して使用することには限界がある。

モデル端末（表4）では、サイズが小さい LeNet 学習モデルしかロードできないことを実験で明らかになった（表5）。

表5 学習モデルロード実験結果

読み込みの結果	
LeNet	○
AlexNet	×
VGG	×

IoT(Internet of Things)など通信技術の進歩によって、次世代通信規格である5Gは1Gbps以上の通信速度と1msの超低遅延を実現できる[6][7]。本研究は、モバイル端末のハードウェア制約を突破するために学習モデルをクラウド上のサーバーに設置し、モバイル端末がインターネットを通じてクラウド上のCNNを利用するゴミ分別支援システムを構築する。

3.1 ゴミ分別支援システムの構築

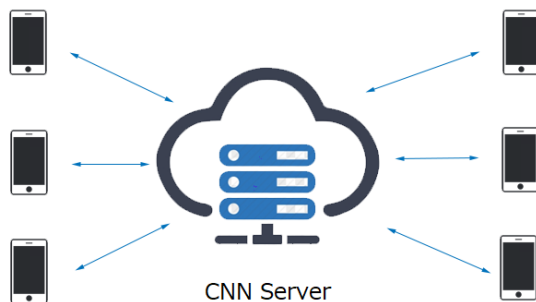


図4 3.1 ゴミ分別支援システムの全体図

図4に示したようにクラウド上に高い計算能力と大容量の記憶領域を有するCNNサーバーを設置し、節2.3で学習したCNNモデルをクラウドサーバに読み込んでモバイル端末からの画像を受信すると分別処理を行い、分別結果を端末へ返信する。

モバイル端末側では、カメラで物体を撮影するアプリ（図5）を実装する。撮影した画像の圧縮やサイズ調整などの前処理を行い、CNNサーバーに送信する。アプリはCNNサーバーからの分別結果を受信して画面に表示してユーザーに通知する。

3.2 ゴミ分別支援システムの評価

表6のサーバーに節2.3で学習したVGG学習モデルを読

み込んでゴミを分別するプログラムを実装した。そして、カメラで物体を撮影して画像をサーバーに送信するAndroidアプリを開発した。画面にタッチすると、撮影が行われ、撮影した画像が圧縮されて表6のサーバーに送信される。そして、サーバーから分別結果を受信して画面に表示する。

表6 CNNサーバーの仕様

CPU	Core i7 6900k 8core 16thread 3.2GHz
GPU	TITAN RTX 24GB
Memory	64GB DDR4
OS	Ubuntu 20.04

実験では、インターネット、5G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4のAndroid端末と表6のサーバーを接続し、性能評価を行った。



図5 Androidアプリ

図5で示したように、Androidアプリで物体を撮影すると、画面にゴミの識別結果（確率）が表示される。画像の受信開始からモバイル端末への分別結果返信完了までの全体処理時間は平均で約403msとなっており、その内に通信

時間は僅か約 28ms が掛かった。ユーザの体感では、撮影が終わるとリアルタイムで結果を表示することができたので、ゴミ分別支援システムの可用性が確認された。

4. まとめ

本研究では、日常生活で直面しているゴミの分別問題について、深層学習技術を用いて解決する方法を検討し、ゴミ分別支援システムについて述べた。画像データを収集し、移転学習の手法を用いて深層学習を行い、規模の異なる3種類の深層学習モデルを作成した。そして、モバイル端末への学習モデル応用について検討し、規模の大きい高精度の深層学習モデルをリソース制約の影響を受けないクラウドサーバーに应用された。モバイル端末からの受信画像から自動的にゴミを分別する支援システムを構築した。

評価実験によって、ゴミ分別支援システムは体感的にリアルタイムで分別処理ができことが明らかになった。

参考文献

- [1] 環境省：“一般廃棄物処理事業実態調査の結果（平成30年度）について”，<https://www.env.go.jp/press/files/jp/113665.pdf>
- [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] A. Krizhevsky, I. Sutskever, and G. E. Hinton: “Imagenet classification with deep convolutional neural networks”, *Advances in neural information processing systems*, pp. 1097–1105, (2012)
- [4] Deep Learning for Java: <https://deeplearning4j.org/>
- [5] K. Simonyan and A. Zisserman: “Very deep convolutional networks for large-scale image recognition”, *3rd International Conference on Learning Representations (ICLR 2015)*, (2015)
- [6] P. Popovski: “Ultra-reliable communication in 5G wireless systems,” *1st International Conference on 5G for Ubiquitous Connectivity. IEEE*, pp. 146–151, (2014)
- [7] A. A. Esswie and K. I. Pedersen: “Opportunistic spatial pre-emptive scheduling for URLLC and eMBB coexistence in multi-user 5G networks,” *IEEE Access*, vol. 6, pp. 38451–38463, (2018)