

# 深層学習を用いたセルフサービス型店舗における非包装商品の分類

情報科学科 林 慶二

指導教員：小林 邦和, 鈴木 拓央

## 1 はじめに

近年、小売業や外食産業を中心に企業の人手不足が深刻化している。そこで現場の人手不足を補う方法として、セルフレジの導入が進んでいる。しかし、セルフレジ化が進む中で、パン屋のパンや八百屋の野菜などの非包装商品にはバーコードや IC タグなどが付けられなく、セルフレジ化は困難である。さらにパン屋であれば、平均 50 種類以上のパンを取り扱っているので、従業員でさえ全種類のパンを覚え、スムーズに清算を行えるまで習得するには時間がかかる。そのため、バーコードなどが付けられない非包装商品の分類を自動で行うレジシステムの開発が必要になっている。先行研究として上から商品を撮影し、画像の特徴量を抽出し、パターン認識によって分類を行う方法が研究されている[1]。本研究では、多様な特徴の自動抽出が行えるディープニューラルネットワーク (DNN)[2] を用いて、さらに高度な分類を行うレジシステムの基幹技術の開発を目指す。

そして、非包装商品の一例として、パンを想定し、パン屋のような先に好きな商品をトレイに載せ、会計を行うセルフサービス型店舗に向けた非包装商品の自動分類が出来るレジシステムの基幹技術開発を目的とする。

## 2 提案手法

処理全体の流れは、まず図 1 のようなトレイに載ったパンの画像から、図 2 のように画像を二値化する。その後、個々のパンを検出し、図 3,4,5 のようなパンごとの 1 枚の画像に切り出す。その後、それぞれのパンの特徴を学習した DNN の入力サイズにパンの画像をリサイズし、画像を DNN に入力し、分類結果を得る。

### 2.1 検出アルゴリズム

パンの検出については、パンの輪郭線を取るために大津の二値化により、画像全体を二値化する。そして、Watershed アルゴリズムで輪郭線を描き出す。輪郭線に伴った外接矩形を検出し、外接矩形に沿って画像を切り出す。外接矩形と輪郭線の間は全て白色に塗りつぶす。その後、白色部分を透過し、元画像と照合することで、黒色部分にパンの画像が置換されパンの画像だけが切り出される。この手順により出力される画像をネットワークの入力サイズにリサイズしている。Watershed アルゴリズムとは、領域分割法のひとつで、接触する物体を分割する境界線を生成する。そのため、単体のパンだけでなく、複数の接触したパンの輪郭も抽出することが出来る。

### 2.2 深層学習

パンの分類については、切り出したパンの画像を入力とし、DNN を用いてパンの分類を行う。DNN の中でも、画像特徴ごとの小さな違いも学習することが出来る畳み込みニューラルネットワーク (CNN) を用いてパンの分類を行う。さらに、計算量が多く、学習に時間がかかるという欠点を CNN は有すが、転移学習 [3] を用いて、その欠点を補う。

転移学習とは、CNN の学習済みのパラメータを他のタスクに対し適用する手法である。ImageNet などの大規模データセットを用いて適切に学習させた CNN のパラメータは非常に汎用性が高く、強力な特徴抽出器として他のタスクへと転用することが可能である。一度学習済みなもので、初期の学習から適切なパラメータに近づいており、分類精度も高い。さらに、一部の層だけを再学習すればよいので、処理時間の短縮も図れる。

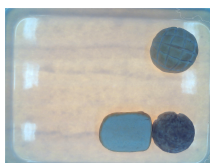


図 1 元画像

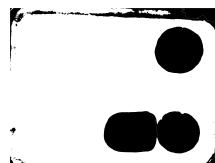


図 2 二値化画像



図 3 模擬パン 1



図 4 模擬パン 2



図 5 模擬パン 3

## 3 実験

### 3.1 実験環境

実験には本物のパンではなく、粘土で作成した模擬パンを使用する。模擬パンを 4 種類作成し各種、表 1 のように訓練データと訓練データに使用していない評価データに分割した。

表 1 データセットの作成

	訓練データ	評価データ
模擬パン	3 個	1 個
撮影画像	15 枚	5 枚
画像の水増し	210 枚	50 枚

撮影方法は以下の手順で行う。

1. 照明器具を上向きにし、その上にトレイを設置
2. トレイ上に模擬パンをランダムに配置
3. トレイの真上に Web カメラを設置し、RGB 画像を撮影

### 3.2 実験方法

ImageNet で学習させた学習済みネットワークである VGG16[4] を特徴抽出器として利用し、全結合層を本研究に合わせて学習させる。そして、提案手法と従来手法を以下のように定義し、正解率を算出し、精度の比較を行う。

- 提案手法: 転移学習で全結合層を再学習した VGG16
- 従来手法: 層の重みをランダムに初期化し学習した VGG16

### 3.3 実験結果

両ネットワークの正解率を表 2 に示す。とても少ない訓練データの数でも両ネットワークとも 90% 以上の正解率が出た。さらに転移学習を行うことで、正解率の向上が見られた。これにより実際の店舗で使用する場合、新商品が出てデータ数が少ない場合でも、高い正解率で分類が出来ることが予想される。

表 2 ネットワークの正解率

手法	正解率 [%]
従来手法	93.0
提案手法	95.5

## 4 おわりに

特徴の自動抽出が行える DNN を用い、さらに高度な分類をやるレジシステムの基幹技術開発を行った。そして、提案手法によって性能向上することが確認できた。今後は、接触したパンだけではなく、重なったパンのセグメンテーションも行い、パン以外の非包装商品にも対応出来るような技術の開発を目指したい。

## 参考文献

- [1] 森本 雅和, 「マイナー成分分析を用いたパンの画像識別」, 電子情報通信学会論文誌 A Vol.J94-A No.7 pp.548-551(2011)
- [2] 岡谷 貴之, 「深層学習」, 講談社, (2015)
- [3] 中山 英樹, 「深層畳み込みニューラルネットワークによる画像特徴抽出と転移学習」, 信学技報 115(146), pp.55-59(2015)
- [4] Karen Simonyan, Andrew Zisserman, "VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION", ICLR2015, P1130-1150(2015)