

画像認識による商品識別システムのための オクルージョンに頑健な認識手法

山田 泉[†] 波部 斉^{†, ††}

近畿大学大学院総合理工学研究科[†] 近畿大学理工学部情報学科[†]

近畿大学情報学研究所^{††}

1. はじめに

現在、セルフレジや RFID を使ったレジ等が普及してきている。その中で、画像認識による自動レジが提案され開発がされている。既にコンビニ等の一部店舗では実際に導入がされており、画像認識による自動レジは実用可能段階に入っているといえる。しかし、問題点がいくつかあり、商品同士が重なって検出が困難になるなどオクルージョンへの対処が挙げられている。実際に、開発がされている自動レジではカゴを使わず商品が重ならないように置くなど制約が多く非効率的なシステムになってしまう。そこで本研究では、画像認識による商品識別システムのためのオクルージョンを考慮した認識手法について提案する。今回は前段階として、物体検出の既存手法である YOLOv3[1] と CompotisionalNets (CompNets) [2] を用いてオリジナルデータセットを学習させ、商品の識別を行う。また、mAP, 検出速度, 学習速度の比較も同時に行う。

2. 提案手法

現時点で想定している、商品識別システムの概要を図1に示す。まず、商品を置く(a)。このとき置き方には制限を設けない。また、商品も1点ずつではなく、まとめて検出する。次に、商品を真上と真横の2方向からカメラで撮影を行う(b)。 (b)で撮影した画像を用いて物体検出を行う(c)。最後に(c)の検出結果から商品一覧を出力する。

本稿では、(c)の物体検出のモデルに 処理速度が速い YOLOv3 とオクルージョン下の物体を高い精度で検出が可能な CompNets を適応し実験することで、オクルージョンに頑健な識別システムに求められる識別モデルを考察する。

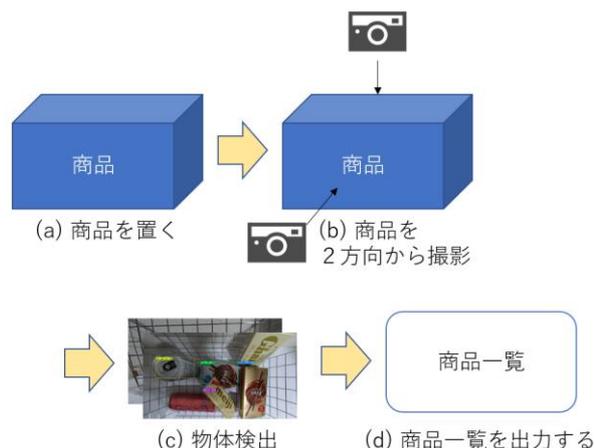


図 1 商品識別システムの概要

3. 実験

YOLOv3 と CompNets に図2のような画像を入力し、mAP, 検出速度, 学習速度の比較を行う。

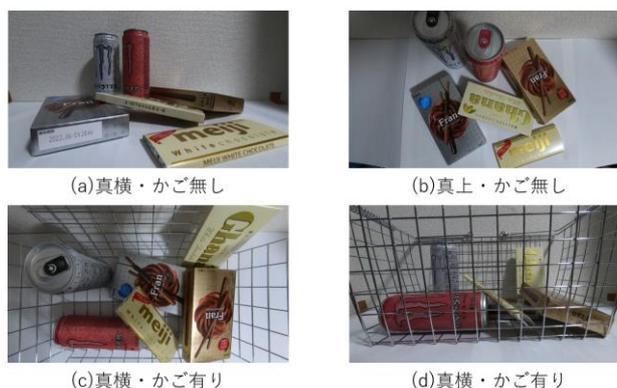


図 2 検出に使用する入力画像例

YOLOv3 と CompNets の学習は表1の条件で行った。このとき学習画像は、かご無し画像 80 枚、かご有り画像 80 枚を使用している。また画像内の商品数は 6 となっている。学習と検出に使用する画像は、表2の条件で撮影を行った。

表 1 検出モデルの学習条件

学習画像枚数	160
epoch	500

An Occlusion Robust Recognition Method for Identification of Merchandise System Based on Image Recognition

[†]ZUMI YAMADA, Graduate School of Science and Engineering, Kindai University.

^{††}HITOSHI HABE, Faculty of Science and Engineering, Department of Informatics, Kindai University.

表 2 撮影条件

使用機材	GoPro HERO7 CHDX-701-FW
解像度	2704 x 1520
フレーム数	60
撮影時間(秒)	40

4. 実験結果

まず、かご無し画像を入力した結果を図3に示す。かごによるオクルージョンが一切ないため、両モデルにおける検出結果に大きな差はなかった。次に、かご有り画像を入力した結果を図4に示す。かご有りでは、真横からの検出においてオクルージョンに強いモデルである CompNets の方が多く商品を検出できており、学習画像が少なくともオクルージョン下の物体検出が十分行えていることがわかる。

ここで、かご有り、かご無しの際の mAP、検出速度、学習速度を表3、表4に示す。まず mAP を比較すると、かご無しでは YOLOv3 と CompNets でそこまで差はなかったが、かご有りでは CompNets の方が高かった。これは CompNets がオクルージョンに強いモデルであるためこの結果になったと考えられる。次に検出速度を比較すると、YOLOv3 の方が 2 秒程早く検出できている。最後に学習速度を比較すると、約 20 分の差があった。店舗等での実用を踏まえると本実験で使用したデータ数よりも多いデータセットで学習するためこの結果よりもさらに大きな差になると考えられる。

これらの結果から、オクルージョンの有無にかかわらず、CompNets は高い精度で検出できているため CompNets をベースとして商品識別システムを構築することが望ましいと考えられる。しかし、実際に運用することを考えると学習速度が長いというのは、新商品が追加されたときに識別モデルの再学習等で大きな影響を及ぼすことが予想できる。したがって、精度を維持しつつ学習速度を短くする等の改良を今後していく必要がある。

表 3 かご無しの mAP・検出速度・学習速度

モデル	mAP	検出速度	学習速度
YOLOv3	0.623	3.2 s	約 30 分
CompNets	0.692	5.1 s	約 50 分

表 4 かご有りの mAP・検出速度・学習速度

モデル	mAP	検出速度	学習速度
YOLOv3	0.325	3.1 s	約 30 分
CompNets	0.538	5.2 s	約 50 分

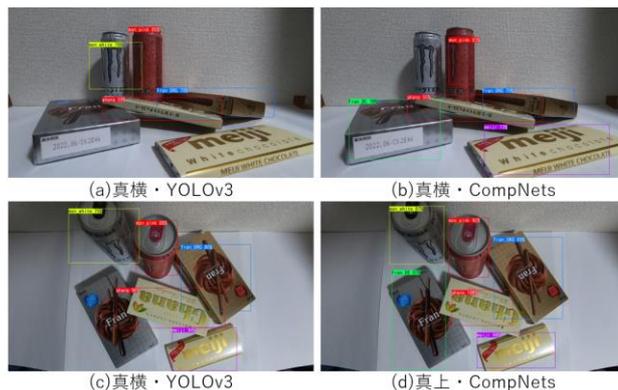


図 3 かご無し画像の検出結果

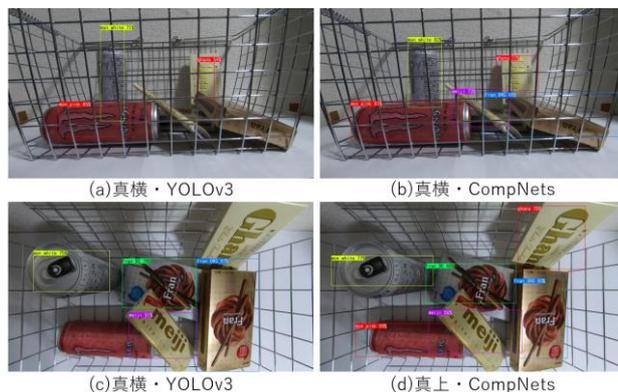


図 4 かご有り画像の検出結果

5. まとめ

本稿では、画像認識による商品識別システムのオクルージョンに頑健な認識手法を調査するにあたり、YOLOv3 と CompNets の比較を行った。オクルージョンの有無にかかわらず、YOLOv3 に比べ CompNets の方が少ない学習データでも、高い精度で商品の検出ができることがわかった。今後は、学習速度が長い問題点を解消するため、精度を維持しつつ学習速度を短くする等の改良を行う。

本研究の一部は、科研費 JP21H05302 の助成を受けた。

参考文献

[1] Joseph Redmon et al. YOLOv3: An Incremental Improvement. Tech report, 2018
 [2] Adam Kortylewski et al. Compositional Convolutional Neural Networks: A Deep Architecture with Innate Robustness to Partial Occlusion. CVPR, 2020.