

CNN を用いた画像中の物体の計数・計量

1190375 三浦 康寛 【 知能情報学研究室 】

1 はじめに

画像中に写り込む物の計量を行う研究は、屋内カメラでの車や人の認識や計測、工場でのベルト上の部品の認識・計測等多くの研究がある [1]。その中でも畳み込みニューラルネットワーク (Convolutional Neural Network: CNN) を用いた手法も多く提案されている。しかし、物体や計測状況の種類を問わない統一的な認識モデルは困難なため、応用分野ごとのモデルの構築をする必要がある。

本研究では、CNN を用いた計量として、分類と回帰の 2 つの問題を扱う。1 つ目は、分類問題の例として、玄関に置かれた靴の個数を計測し在室人数を推定する問題であり、プライバシーに配慮した在室管理としての応用が期待されている。2 つ目は、回帰問題の例として、ベルト上のカットされた食品の質量を推定する問題であり、低コストな簡易計量システムとしての応用が可能である。

2 提案する CNN モデル

本研究では、玄関画像とベルトコンベア画像に対し CNN を適用し、それぞれ画像中の靴の個数の推定とダンボール片の質量の推定を行う (図 1)。推定モデルは畳み込み層とプーリング層のペアを組み合わせた 6 層、8 層、10 層の 3 種類のモデルを設定する。靴の個数の推定ではフィルタ数を 30 と設定し、損失関数には交差エントロピーを用いる。ダンボール片の質量の推定ではフィルタ数を 20 と設定し、損失関数には平均二乗誤差 (MSE) を用いる。また、それぞれのタスクにおいて畳み込み層のフィルタサイズを 3、プーリング層のフィルタサイズを 2 と設定し、活性化関数には ReLU 関数、最適化には Adam を用い、学習エポックを 50 とする。

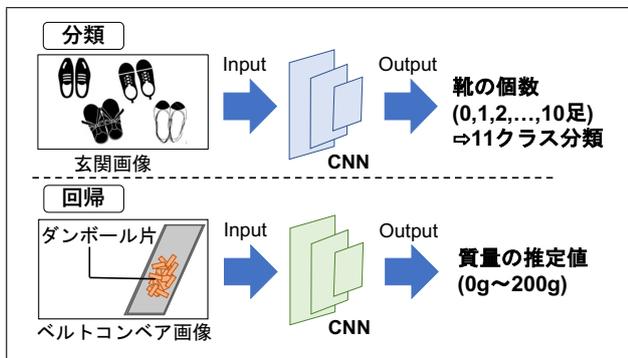


図 1 分類についての CNN 推定 (上) と回帰についての CNN 推定 (下)

表 1 分類精度と質量推定誤差

	背景差分法	CNN		
		6 層	8 層	10 層
靴の個数	66.6%	96.2%	98.6%	97.3%
ダンボール片の質量 (MAE)	4.65	3.94	3.05	3.49

3 検証実験

3.1 データセット

靴の個数の推定では、部屋の玄関の天井に取り付けたカメラで撮影した玄関画像を 5535 枚用いる。この玄関画像は、20 人の構成員と一部部外者が出入りする部屋のものとする。ダンボール片の質量の推定タスクでは、縦 100cm、横 20cm のベルトコンベア上のダンボール片を 50cm 上から撮影した画像を 6874 枚用いる。

3.2 検証方法

靴の個数の推定では、靴無しから靴 10 足までの 11 種類の分類を行い、正答率を算出する。ダンボール片の質量の推定では、0 から 20 グラムの間でランダムに撒かれたダンボール片の質量を推定し、実際の質量との平均絶対誤差 (MAE) を算出する。最終的な 3 つの推定モデルの正答率、MAE は 5 分割の交差検証法で算出する。

4 検証結果

表 1 に検証結果を示す。靴の個数の推定では背景差分が 66.6 % の正答率であるのに対し、CNN では 8 層のモデルで 98.1 % である。ダンボール片の質量推定では背景差分の MAE が 4.65 であるのに対し、CNN では 8 層のモデルで 3.05 である。

5 まとめ

本研究では、靴の個数の推定精度とダンボール片の質量の推定精度を背景差分法を用いた手法と CNN を用いた手法とで比較した。靴の個数の推定では、CNN を用いた手法が背景差分法を用いた手法に比べ、正答率が約 30 ポイント向上した。ダンボール片の質量の推定では、CNN を用いた手法が背景差分法を用いた手法に比べ、MAE が 1.5 減少した。これらのことから、画像中の物体の計数・計量において、CNN が有効であることが示された。

参考文献

[1] 高橋ら, “玄関の履物に対する画像認識を用いた在室管理手法の提案,” DICOMO 2016.