

移動ロボットのための領域ベース CNN と Triplet Loss を用いた人物再同定

村田 祐樹† 渥美 雅保‡

創価大学大学院工学研究科情報システム工学専攻† 創価大学理工学部情報システム工学科‡

1. はじめに

人物再同定とは一般に、視野を共有しない複数のカメラに映る人物像から同一人物を識別する問題である[1]。本研究では、特定の集団に定期的にサービスを提供する移動ロボットに搭載されたカメラによる人物再同定問題を扱う。本手法は、領域ベース CNN が検出する人物領域から Triplet Loss[2]に基づき人物特徴を抽出する特徴抽出器と人物特徴の転移学習を行う人物再同定器とから構成される。人物再同定器には服装等部分的に外見が異なる人物を定期的に再学習するための転移学習が組み込まれる。公開されている大規模データセットを用いた実験、及び移動ロボットにより同一集団に対して定期的に収集した自作データセットを用いた実験により、提案システムの性能を評価する。

2. 人物再同定手法

2.1. 概要

図 1 に示すように、3 種類の CNN で人物の検出と特徴抽出・再同定を行う。まず、人物検出のための領域ベース CNN で人物検出を行うと同時に、中間層から特徴量を取り出す。特徴量は人物領域に基づいてトリミングし、バイリニア補間法を用いてリサイズする。次に、この領域特徴を、人物特徴学習時は人物特徴抽出 CNN、人物再同定時は人物再同定 CNN へ入力する。2 つの CNN は前半部分が同一の構成となっており、特徴抽出 CNN 学習時の重みを再同定 CNN で利用する。

2.2. 人物検出

人物再同定を行うためには、画像中から適切な人物領域を取り出す必要がある。本手法では COCO データセットで事前学習された YOLOv2[3]を使用する。YOLOv2 は、セグメンテーションのタスクで用いられることの多い完全畳み込みネットワーク (FCN) 構造を採用し、精度の高い物体検出をリアルタイムに行うことができる。

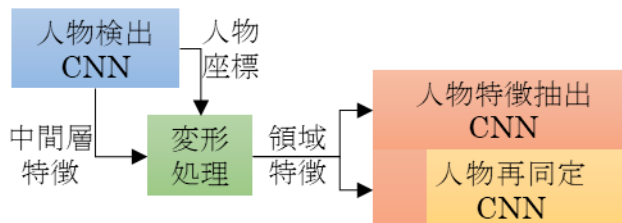


図 1 人物再同定手法

2.3. 人物特徴抽出

特徴抽出 CNN(図 2(a))は複数の畳み込み層とプーリング層で構成され、最後の全結合層で 512 次元の人物特徴を出力する。損失関数には、Triplet Loss[2]を用いる。Triplet は基準となるサンプル(anchor), それに類似するサンプル(positive), 類似しないサンプル(negative)の 3 つ組で構成され、それぞれが重みを共有した同じ特徴抽出 CNN に入力される。出力される特徴量同士を比較したときに、同一クラス間で距離が小さく、別クラス間で距離が大きくなるように学習が行われる。

2.4. 人物再同定

再同定 CNN(図 2(b))では特徴抽出 CNN で学習した重みを利用した転移学習を行う。出力層には人物クラスを分類するソフトマックス関数を、損失関数には交差エントロピーを用いる。学習時のパラメータ更新は、学習の収束を早めるために新たに追加した後半の層にのみ適用する。

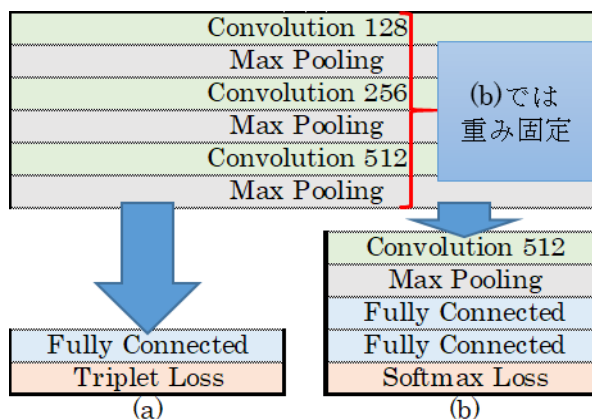


図 2 特徴抽出器(a)と再同定器(b)の CNN 構成
畳み込み層のフィルタサイズは 3×3, プーリング層のフィルタサイズは 2×2, ストライドは 2.
Convolution の横の数値はフィルタ数を示す。

Person Re-identification for Mobile Robot based on Region-based CNN and Triplet Loss

†Yuki Murata

Graduate School of Engineering Dept. of Information Systems Eng., Soka University

‡Masayasu Atsumi

Dept. of Information Systems Sci., Faculty of Sci., and Eng., Soka University

3. 実験

3.1. 実験概要

特徴抽出 CNN と再同定 CNN の性能評価実験を行う。特徴抽出 CNN の実験では、出力される特徴量間の類似度を利用して再同定精度を評価する。再同定 CNN の実験では、別のデータセットで学習した特徴抽出 CNN の重みを利用して、転移学習を実施した際の再同定精度を評価する。事前に行った複数の実験から、各 CNN へ入力する領域特徴は検出 CNN の 2 番目の畳み込み層から取得し、サイズは $20 \times 20 \times 64$ とした。

3.2. データセット

実験に使用したデータセットの構成を表 1 に示す。人物再同定のデータセットは、人物領域の画像だけで構成されている場合が多い。本手法では人物検出 CNN の中間層から特徴を取り出すため、人物領域が切り取られる前の画像全体のデータが必要である。

Person Re-identification in the Wild (PRW) [1] では画像全体と人物領域の座標情報を提供しているため、今回はこのデータセットを特徴抽出 CNN の再同定性能評価実験に使用する。Train データは学習時のみ使用し、評価には Gallery データと Query データを用いる。PRW データセットは検出器によって Gallery を作成するという想定の下で作成されているが、本実験では再同定性能のみ評価するため検出は行わず、与えられた人物領域に基づいて Gallery を作成している。

再同定 CNN の転移学習評価実験にはデータセット SOTARO-1 および SOTARO-2 を用いる。これらは、授業支援ロボット SOTARO [4] でそれぞれ別日に収集したデータを再構成したものである。この 2 つのデータセットには異なる服装の同一人物が含まれる。

表 1 データセット構成(サンプル数 / クラス数)

	Train	Gallery	Query
PRW	17690 / 483	16343 / 450	2057 / 450

	Train	Test
SOTARO-1	12789 / 41	3193 / 41
SOTARO-2	12825 / 48	3194 / 48

3.3. 特徴抽出器の評価

まず、Train データから Triplet を作成して、特徴抽出 CNN を学習する。次に、学習済み CNN を使って人物特徴を抽出し、コサイン類似度を求めて分類を行う。PRW データセットでは Gallery データと Query データ間で、SOTARO データセットでは Train と Test データ間で類似度を算出する。結果を表 2 に示す。ここで r1, r5, r10, r20 は類似度の高い上位 1, 5, 10, 20 サンプル中に 1 つでも正解サンプルが含まれていれば正しいと評価する指標で

表 2 類似度評価の結果

dataset	r1	r5	r10	r20	mAP
PRW	78.7	89.0	92.2	94.3	33.6
SOTARO-1	84.1	86.0	86.9	87.8	79.8
SOTARO-2	78.9	82.2	83.3	84.5	74.6

ある。また、mAP は mean Average Precision で、平均精度の平均を表している。

3.4. 人物再同定器の評価

SOTARO データセットを用いて、別日のデータで学習した再同定 CNN の転移学習性能を評価する。結果を表 3 に示す。

表 3 再同定評価の結果

feature / trans	r1	r5	r10	r20	mAP
SOTARO-2 / 1	81.3	93.6	97.2	98.4	86.8
SOTARO-1 / 2	76.4	90.0	95.3	98.2	82.4

ここで、1 列目の feature は特徴抽出 CNN の学習に使用したデータセットを、trans は再同定 CNN の転移学習に使用したデータセットを表す。どちらも、同じ日のデータを用いた評価とほぼ同等の精度の人物再同定器を作成することができた。また、転移学習の収束に要した時間は GPU 環境下で 3 分ほどであり、短時間で異なるデータセットに適した再同定器の学習ができることが分かった。

4. むすび

本論では、領域ベース CNN の中間層の特徴から Triplet Loss による特徴学習と転移学習により人物再同定を行う移動ロボット向けの手法を提案した。そして、大規模データセットと自作のデータセットに対して行った実験により、その有用性を確かめた。今後の課題として、移動ロボットがサービスを提供しながら人物再同定のためのデータの収集と再学習を繰り返し実行するシステムの構築があげられる。

参考文献

- [1] L. Zheng, H. Zhang, S. Sun, M. Chandraker and Q. Tian. Person Re-identification in the Wild, arXiv preprint arXiv:1604.02531, 2016
- [2] S. Florian, K. Dmitry, and P. James. FaceNet: A Unified Embedding for Face Recognition and Clustering. arXiv preprint arXiv:1503.03832, 2015
- [3] J. Redmon, and A. Farhadi. YOLO9000: Better, Faster, Stronger. arXiv preprint arXiv:1612.08242, 2016
- [4] 渥美雅保, 村田祐樹, 安川葵. SOTARO: オープンチャットとロボットの連携による演習授業支援システム—Human-in-the-loop 型機械学習によるアプローチ—. 情報処理学会第 79 回全国大会, 2017.