

# 姿勢不変なドメイン依存特徴の転移学習に基づく人物再同定

村田 祐樹<sup>†</sup> 渥美 雅保<sup>†</sup>

創価大学大学院工学研究科情報システム工学専攻<sup>†</sup>

## 1. はじめに

複数のカメラで撮影された人物画像から同一人物を見つける人物再同定において、人物姿勢の変化への対処と、新たなデータへの適合が課題とされている。本研究では、大規模データセットで学習された人物の姿勢不変特徴を維持しながら、新たなデータに適した特徴を獲得するモデルを提案する。提案モデルは、FD-GAN[1]を用いた姿勢不変特徴抽出器と、追加した複数の畳み込みレイヤーからなるドメイン依存特徴抽出器とで構成される。そして、少ない計算量で、追加したドメイン依存特徴抽出器のみ更新を行う転移学習を実施する。複数のデータセットを用いた実験から、提案モデルが優れた転移性を有することを示す。

## 2. 関連研究

深層学習によって、人物再同定の精度は飛躍的に向上した。しかし、現実的なアプリケーションへの適用を考えるためには、逐次収集される少量のデータを用いて短時間のうちにモデルを繰り返し調整できるようにすべきである。

人物再同定問題において少量データでの学習を困難にする要因の一つに、多様な姿勢がもたらす同一人物特徴の変化が挙げられる。Liqianら[2]は、GAN (Generative Adversarial Network) を用いて姿勢を変換した人物画像を生成し、学習用データを水増しする手法を提案した。Xuelinら[3]は、クラスタリングによって得られた数種類の標準姿勢に基づいて人物特徴を抽出するPN-GANを提案した。これらの手法では、特徴の抽出が画像の生成に依存するため、推論時にも姿勢情報に基づいた画像の生成が必要である。

本研究で用いるFD-GANは、姿勢変換画像の生成と人物固有の特徴を同時に学習するモデルである。人物特徴と姿勢特徴に基づいて姿勢変換画像を生成する中で、人物再同定に有用な姿勢不変特徴の抽出が可能となる。また、推論時には画像生成を行わないため、追加の姿勢情報も不要となる。

## 3. 提案手法

提案モデルを図1に示す。Pose Encoder( $E_p$ )は、変換画像の目標となる姿勢情報を受け取り、5つの畳

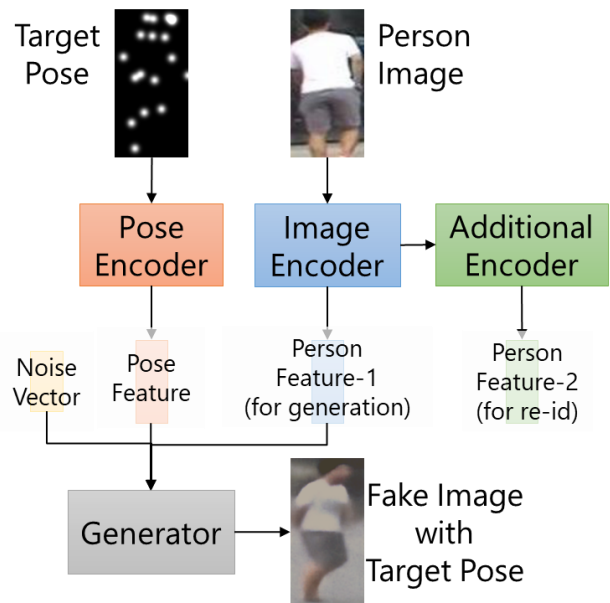


図1 提案モデル

み込み層を介して128次元の特徴ベクトルを出力する。入力される姿勢情報は、OpenPose[4]を用いて抽出したものを使用し、各チャンネルが1つの関節点の位置を表す特徴マップで表現される。Image Encoder( $E_I$ )は、入力された画像から人物固有の512次元の特徴ベクトルを出力する。これら出力と256次元のノイズベクトルを組み合わせられた特徴ベクトルがGenerator( $G$ )へ入力され、目標の姿勢をとった入力画像と同一人物の画像が生成される。Additional Encoder( $E_A$ )は、 $E_I$ の中間出力から得られる特徴マップを入力し、skip-connectionを適用した複数の畳み込み層を介して再同定用の特徴を出力する。提案モデルの $E_p$ および $G$ はFD-GANに基づいて構築している。 $E_I$ は、パラメータ数を削減する目的でResNet-50からResNet-18に変更した。

提案手法では、大規模データセットを用いてモデル全体の事前学習を行った後に、 $E_A$ に対して新たなデータに適合させるための転移学習を実施する。これにより、 $E_I$ は人物画像生成のための姿勢不変特徴を抽出し、 $E_A$ は人物再同定のためのドメイン依存特徴が抽出される。 $E_A$ の学習には、出力される2人の人物特徴が同一人物によるものか否かを判断する識別器 (Verification Classifier:  $C_v$ ) を使用する。 $C_v$ の構成を図2に、損失関数 (binary cross-entropy loss) を(1)式に示す。

$$L = -C \log d(x_1, x_2) - (1 - C)(1 - \log d(x_1, x_2)) \quad (1)$$

Person Re-identification based on Transfer Learning of Pose-invariant Domain-dependent Features.

<sup>†</sup>Yuki Murata, <sup>†</sup>Masayasu Atsumi

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

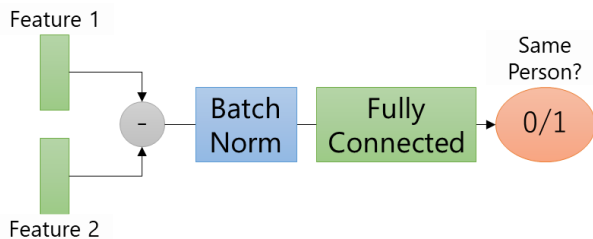


図2 識別器 (Verification Classifier) の構成

ここで $C$ は、画像 $x_1$ と $x_2$ が同一人物によるものであるとき $C = 1$ 、そうでないとき $C = 0$ となる正解ラベルを、 $d(x_1, x_2)$ はネットワークが出力する同一人物の信頼度スコアを表す。

#### 4. 実験

本実験では、提案モデルの人物再同定精度と転移性の評価を行う。実験には Market-1501[5]および DukeMTMC-reid[6]の2種類の公開データセットを使用した。評価指標には mean Average Precision(mAP)と Top-1 Accuracy を使用した。Additional Encoder ( $E_A$ )は、Image Encoder ( $E_I$ )の最終畳み込みブロックの後に畳み込み層を4層追加した Model A と、最終畳み込みブロックの前で分岐して畳み込み層を5層追加した Model B の2種類を構築し、それぞれ評価する。

##### 4.1. 再同定精度の評価

Market-1501 データセットを用いて FD-GAN と提案モデルを訓練し、人物再同定の精度を比較する。表1に示す通り、提案した Model B が FD-GAN を上回る精度で再同定できることがわかった。

##### 4.2. 転移性の評価

4.1で訓練済みのモデルと DukeMTMC-reid データセットを用いてドメイン依存特徴の転移学習を実施し、転移性の評価を行う。実験は全ての訓練データを使用した場合と、一定の訓練データをランダムに抽出した場合で行い、それぞれの再同定精度を評価した。また比較対象として、FD-GANの最終畳み込みブロックを Fine-Tuning したモデル (Baseline) を同様に評価した。

表2に示す結果から、Model B の転移性は Model A を大きく上回ることがわかった。また、訓練データを60%まで削減した場合でも、精度の低下は Top-1 Accuracy で3%以内に収まった。Baseline モデルは、再同定精度では提案モデルを上回る結果となったが、FD-GAN によって学習された姿勢不変特徴を維持できていない可能性がある。図3の各モデルが生成する姿勢変換画像からは、提案モデルの方が本人らしい画像の生成能力があり、姿勢不変特徴が維持できていることがわかる。

表1 再同定精度の評価

	Top-1	mAP
FD-GAN	86.7	70.2
Model A	86.6	68.7
Model B	<b>87.6</b>	<b>71.8</b>

表2 転移性の評価

訓練データ量	Model A		Model B		Baseline	
	Top-1	mAP	Top-1	mAP	Top-1	mAP
1.0	38.4	23.0	57.6	37.5	60.2	40.9
0.8	37.3	22.5	56.7	36.6	58.7	39.4
0.6	36.4	21.8	54.7	34.6	57.5	38.1
0.4	34.5	20.5	50.5	31.1	54.9	35.5
0.2	31.7	18.7	42.5	25.2	45.0	28.2

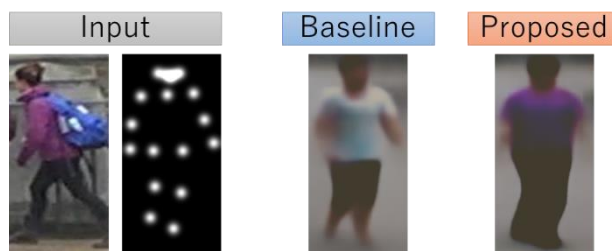


図3 各モデルが生成する画像

#### 5. むすび

本研究では、大規模データセットで学習された人物の姿勢不変特徴を維持しながら、新たなデータに適した特徴を獲得するモデルを提案し、その有効性を実験により評価した。これらから、より高い転移性を実現するために、ドメイン依存特徴を抽出する機構に対して更なる工夫が必要であることがわかった。

#### 参考文献

[1] Yixiao, Ge, et al. FD-GAN: Pose-guided feature distilling GAN for robust person re-identification. *NeurIPS*, 2018.  
 [2] Liqian, Ma, et al. Pose guided person image generation. *NeurIPS*, 2017.  
 [3] Xuelin, Qian, et al. Pose-normalized image generation for person re-identification. *ECCV*, 2018.  
 [4] Zhe, Cao, et al. Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *CVPR*, 2017.  
 [5] Liang, Zheng, et al. Scalable Person Re-identification: A Benchmark. *ICCV*, 2015.  
 [6] Zheng, Zhedong, et al. Unlabeled Samples Generated by GAN Improve the Person Re-identification Baseline in vitro. *ICCV*, 2017.