歩容特徴の抽出精度向上のための 異なる人物間の特徴交換を用いた歩容認証

○吉野 弘毅 中嶋 一斗 安 正鎬(九州大学) 岩下 友美(NASA/Caltech JPL) 倉爪 亮(九州大学)

1. はじめに

人の歩く様子を撮影した歩容映像は、対象者に特別な動作を強いることなく非接触で取得できる利点から、犯罪捜査やスムーズな入退室管理に有効な生体情報として期待されている。歩容を用いた個人識別タスクである歩容認証は、パターン認識の観点では、学習データを構成する既知クラスに含まれない、未知クラスのデータを評価対象とするオープンセット認識問題の一つであり、認証対象者に特化した特徴や識別器を直接学習することはできない。そのため既知人物の歩容の像を使って、評価対象となる未知の人物にも適応する汎化性能の高い特徴抽出器を作ることが歩容認証の目的となる。しかし、歩容映像には服装や背景等の歩容に固有でない情報(共変量)が多く含まれ、認証を困難にしている。

従来多くの歩容認証手法は、背景差分手法や骨格抽出手法等の前処理によって共変量を除去してきたが、近年では、データの潜在的に独立な属性に従ってニューラルネットワークの内部表現を分離する Disentanglement Representation Learning(分離表現学習)が活発に研究されている。この手法は、属性ラベルを教示する必要なく解釈可能な特徴表現を学習できる。この利点から、分離表現学習を用いて共変量や歩容に関係する特徴を入力から直接抽出する手法 [2, 3] が提案されている。しかし、分離表現学習は一般に教師なし学習であるため、分離性能は学習データの量への依存性が高い、歩容認証は他のタスクに比べてデータセットの量が少ないため、分離性能は限定的であると考えられる。

この問題に対して、我々の先行研究 [1] では、分離表現学習の分離性能向上に伴う識別精度の向上を目的として、分離表現学習に基づく見えに関する共変量の特徴(以下、スタイル特徴)の除去、および敵対的生成に基づくデータ拡張による学習データの量とバリエーションを増加させる歩容認証手法を提案している。実験では、データ拡張によって識別精度が向上することが示されたが、異なる映像間での特徴交換によるデータ拡張は同一人物間に限定されている。

そこで本研究では、先行研究 [1] によるデータ拡張を 異なる人物間での特徴交換へと拡張する手法を提案す る.提案手法では、人物ごとの歩容映像の数に応じて、 学習データのバリエーションおよび量を増やすことが できるため、分離性能および特徴抽出器の汎化性能の 向上が期待できる.具体的には、まず歩容映像は、歩 容に依存する特徴(姿勢特徴)および外観に関する共 変量の特徴(スタイル特徴)に分離される.次に、異 なる人物の二つの歩容画像間で、姿勢特徴を交換して

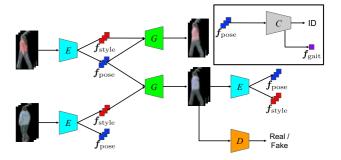


図1 提案手法の概要図. (1) アンカーとなる歩容映像, (2) アンカーと同一 ID で異なる歩容映像, (3) アンカーとは異なる ID の歩容映像の三つの歩容映像から,二つの映像をペアとして入力する.

生成された歩容画像を、学習データに追加する. 最後に、時系列の姿勢特徴群から抽出される歩容特徴を用いて識別を行う. 識別学習と生成学習は同時に最適化される. 実験では、識別精度の算出によって提案手法におけるデータ拡張の有用性を検証する.

2. 提案手法

提案するネットワークを図1に示す.提案手法において,姿勢特徴は,時系列の歩容画像において,時間で変化し,かつ人物に依存する特徴(全身の形状や関節の角度など)であり,スタイル特徴は,時間で不変の,人物に依存しない特徴(服装・歩行方向など)と定義する.提案手法は先行研究[1]の同一人物間での特徴交換生成によるデータ拡張を,異なる人物間での特徴交換生成に拡張し,学習の安定化のために損失を一つ追加している.以下では提案手法のパイプラインについて述べるが,損失の定義などの詳細な情報は先行研究[1]を参照されたい.

2.1 歩容画像から姿勢特徴とスタイル特徴へのエンコード

入力された歩容映像は、エンコーダによって姿勢特徴およびスタイル特徴に分離・抽出される.歩容画像の特徴が適切に分離・抽出されるように、同一人物で歩行条件の異なる二つの歩容画像からそれぞれ得られた姿勢特徴から、姿勢特徴の類似性損失を計算する.

2.2 姿勢特徴とスタイル特徴を用いた歩容画像の生成

次に、エンコーダEによって抽出された姿勢特徴およびスタイル特徴から、生成器Gによって歩容画像を生成する。生成された画像および元画像は識別器Dに入力され、生成画像であるか実画像であるかを判別さ

れる. 画像の生成には元画像の再構成と仮想歩容画像 の新規生成の2種類が存在する.

再構成の際には、同一の歩容映像の異なるフレームから抽出した姿勢特徴とスタイル特徴から、姿勢特徴の元となった画像が生成されるように再構成損失を定義する. 再構成損失によって、歩容画像中の時間で変化しない情報がスタイル特徴として、時間で変化する情報は姿勢特徴として抽出されるようになる.

新規生成の場合には,識別器 D が入力された画像が 実画像か生成画像であるかを識別し,敵対的損失を計 算する.

2.3 生成された歩容画像の再エンコード

生成された画像も識別学習に活用するために,生成画像は再度エンコーダEに入力され,2.1節と同様に姿勢特徴とスタイル特徴に分離される.

分離された特徴の元となった画像の特徴を,生成画像に正確に保持させるために,再度エンコードする前後での両特徴の差分から,エンコーダEの一貫性を保証する一貫性損失を計算する.

2.4 元画像の循環再構成

先行研究では再エンコードされた後,姿勢特徴のみ識別学習に用いられてスタイル特徴は破棄されるが,提案手法では再エンコードした特徴から元画像を再構成する.異なる二つの歩容画像 I^A , I^B に対して,エンコード,特徴交換生成,再エンコードの順で処理することで, I^A の姿勢特徴 $E_{\rm pose}\left(G(f^B_{\rm style},f^A_{\rm pose})\right)$, I^A のスタイル特徴 $E_{\rm style}\left(G(f^A_{\rm style},f^B_{\rm pose})\right)$ が得られる.これら二つの特徴をもとに,次の循環再構成損失を定義する.

$$\mathcal{L}_{\text{cycle}} = \sum_{\substack{a,b \in \{A,B\}\\a \neq b}} \left\| G\left(E_{\text{style}} \left(G\left(f_{\text{style}}^{a}, f_{\text{pose}}^{b} \right) \right), \right.$$

$$\left. E_{\text{pose}} \left(G\left(f_{\text{style}}^{b}, f_{\text{pose}}^{a} \right) \right) \right) - I^{a} \right\|_{1}$$

$$(1)$$

この損失は、Zhuら [5] が報告するように教師データのない画像生成の高品質化への貢献が期待され、提案手法における特徴の分離性能の向上にも寄与すると考えられる.

2.5 姿勢特徴による人物識別

実画像・生成画像の両方から抽出された姿勢特徴を、 分類器 C に入力し、被験者 ID の確率分布を出力する。 分類器 C によって出力された人物 ID の確率分布をも

表 1 比較手法

手法	データ拡張	$\mathcal{L}_{ ext{cycle}}$		
ベースライン	なし			
先行研究 [1]	同一人物間			
提案手法	異なる人物間	\checkmark		
w/o $\mathcal{L}_{\mathrm{cycle}}$	異なる人物間			

表 2 CASIA-B [4] における識別精度 [%]

手法	NM	BG	CL
ベースライン	90.2	86.1	24.3
先行研究 [1]	$\underline{94.4}$	<u>88.8</u>	29.2
提案手法	93.3	86.5	24.0
w/o $\mathcal{L}_{\mathrm{cycle}}$	94.0	85.0	30.0

とに、入力フレーム数に応じて重み付けをした交差エントロピーを計算する.

学習では、以上全七つの損失に対して重みをかけて 足し合わせて同時に最適化する.

3. 実験

実験では、CASIA-B [4] データセットを用いて、識別精度の評価を行った。CASIA-B は、NM (標準)・BG (バッグを所持)・CL (他とは異なる服装) の三つの設定から構成される。表 1 に実験に用いた手法を示す。

比較手法として,先行研究 [1] およびそのベースライン [2],提案手法と提案手法から循環再構成損失 1 を除いた設定の計 4 手法の精度を算出した.ただし,本実験で用いたベースラインは,Zhang [2] らの論文を元に筆者らが実装したものである.

3.1 識別精度の定量評価

NM, BG, CL の各設定における Rank-1 認証率の値を表 2 に示す.表 2 より,比較手法の中で先行研究が最も精度が高く,再構成循環損失による精度の向上は限定的であることがわかる.

異なる人物間での特徴交換によるデータ拡張が識別精度の向上に寄与しなかった理由として、生成する画像の多様性が向上したことで識別学習が困難になったことが挙げられる.よって改善策としては、拡張したデータを現状の枠組みで識別するのではなく、効率的に識別学習するように再設計することが考えられる.例えば、特定の歩き方をする人物の歩容画像を多様なスタイルで生成できる利点を活用して、現状の識別学習の代わりにTriplet Loss [6] による距離学習を採用することが考えられる.

また循環再構成損失は、特徴交換生成による教師データのない画像生成の高品質化と特徴の分離性能の向上に寄与すると考えられる.しかし識別精度の向上に寄与しなかったのは、損失が合計7種類となっていて学習が複雑になっていることが大きな原因の一つであると考えられる.この問題に対しては、同じく特徴の分離性能の向上を目的とする一貫性損失との比重の調整を行っていく予定である.

4. まとめと今後の予定

本研究では、分離表現学習の分離性能を向上させる ために、分離表現学習とデータ拡張を組み合わせた新 しい歩容認証手法を提案した.提案手法では、歩容画 像から姿勢とスタイル特徴を分離・抽出する.分離し た姿勢特徴を、異なる人物の映像間で交換することに よって仮想の歩行画像を生成し、学習データに加える. 実験の結果、提案手法の識別精度の向上には課題があ ることがわかったため、その原因の考察を行った.

今後は,損失間の比重の再設定および拡張したデータを活用した距離学習の実装を行っていく予定である.

謝 辞 本研究は JST 科学技術イノベーション創出 に向けた大学フェローシップ創設事業 JPMJFS2132 の支援を受けたものである.

参考文献

- [1] Koki Yoshino, Kazuto Nakashima, Jeongho Ahn, Yumi Iwashita and Ryo Kurazume, "Gait Recognition using Identity-Aware Adversarial Data Augmentation." *IEEE/SICE International Symposium on System Integration (SII)*, pp.596–601, 2022.
- [2] Ziyuan Zhang, Luan Tran, Xi Yin, Yousef Atoum, Xiaoming Liu, Jian Wan, and Nanxin Wang. "Gait Recognition via Disentangled Representation Learning." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4710–4719, 2019.
- [3] Xiang Li, Yasushi Makihara, Chi Xu, Yasushi Yagi, and Mingwu Ren. "Gait Recognition via Semi-supervised Disentangled Representation Learning to Identity and Covariate Features." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 13309–13319, 2020.
- [4] Shiqi Yu, Daoliang Tan, and Tieniu Tan. "A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition." In *Proceedings of the IEEE International Conference on Pattern Recognition (ICPR)*, vol. 4, pp. 441–444, 2006.
- [5] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. "Unpaired image-to-image translation using cycle-consistent adversarial networks." In *Proceedings* of the IEEE international conference on computer vision (ICCV), pp. 2223–2232. 2017.
- [6] Alexander Hermans, Lucas Beyer, and Bastian Leibe. "In defense of the triplet loss for person reidentification." arXiv preprint arXiv:1703.07737, 2017.