

# 歩容認証のための容姿・姿勢特徴の合成によるデータ拡張

## Data Augmentation by Synthesizing Appearance and Pose Features for Gait Recognition

村上斗吾<sup>1\*</sup>  
Toa Murakami<sup>1</sup>

原田翔太<sup>1</sup>  
Shota Harada<sup>1</sup>

満上育久<sup>1</sup>  
Ikuhisa Mitsugami<sup>1</sup>

<sup>1</sup> 広島市立大学

<sup>1</sup> Hiroshima City University

**Abstract:** 本研究は、異なる人物の容姿特徴と姿勢特徴の合成による、歩容認証のためのデータ拡張を提案する。機械学習技術を活用した高精度な歩容認証の実現には、多様かつ大量の歩容画像が必要となる。しかし、多様かつ大量の歩容画像の収集には多くのコストを伴う。本研究では、少量の歩容画像から容姿特徴と姿勢特徴を抽出し、異なる人物のそれらの特徴を組み合わせた画像を生成する手法を提案する。実験では、歩容画像からの年齢分類実験により提案手法の有効性を確認した。

### 1 はじめに

人の歩容には、年齢・性別・健康状態といった人物の属性や状態を特定できる様々な情報が含まれている。そのため近年では、機械学習技術を用いて歩容画像から年齢層の推定・性別の識別、うつ病やパーキンソン病などの疾患を検出する研究が多く実施されている。しかし、機械学習技術のための歩容画像の学習データを収集するためには、実験参加者の同意や実験環境の整備が不可欠であり、多様な歩容パターンの取得は容易ではない。この問題を解決する手段の一つとして、既にある歩容画像をもとに深層学習技術を用いて新たな歩容画像を生成し、学習データを増やす方法（データ拡張）がある。

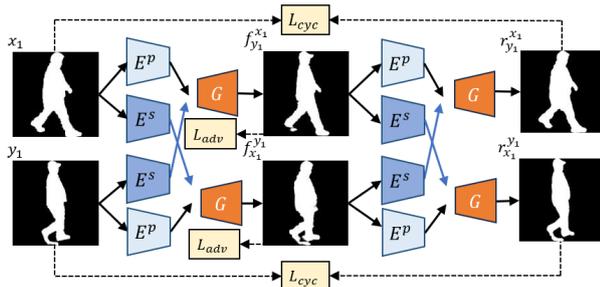
そこで本研究では、異なる人物の見た目の特徴（容姿特徴）と歩き方の特徴（姿勢特徴）の合成による、歩容認証のためのデータ拡張の手法を提案する。提案手法は、まず、複数の要素が混在しているデータから各要素に対応する独立した特徴を抽出する Disentangled Representation Learning (DRL) [1] を歩容画像に応用して、容姿特徴と姿勢特徴にそれぞれ分離する。そして、異なる人物のそれらの特徴を組み合わせて、新しい人物の歩容画像を生成する。本稿では、歩容画像からの年齢分類実験を通して、歩容画像のデータ拡張としての提案手法の有効性を示す。

### 2 提案手法

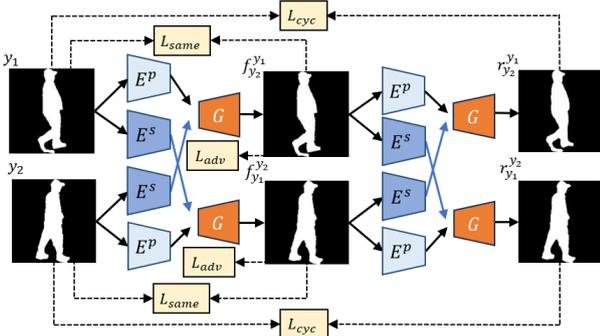
提案手法の概要を図1に示す。提案手法は容姿特徴抽出器  $E^s$ 、姿勢特徴抽出器  $E^p$ 、生成器  $G$ 、識別器  $D$  の4種類のモジュールから構成される。まず、ある人物の歩行シルエット画像  $x_1$  と他の人物のそれぞれ異なる姿勢の歩行シルエット画像  $y_1, y_2$  から、容姿特徴を  $E^s$  で、姿勢特徴を  $E^p$  で抽出する。次に、容姿特徴を入力画像のペア  $\{x_1, y_1\}, \{y_1, y_2\}$  の中で交換し、 $G$  を用いて合成画像  $f_{y_1}^{x_1}, f_{x_1}^{y_1}, f_{y_2}^{y_1}, f_{y_1}^{y_2}$  を生成する。最後に、合成画像から容姿・姿勢特徴を抽出し、入力画像ペアの容姿特徴を再び交換することで、再構成画像  $r_{y_1}^{x_1}, r_{x_1}^{y_1}, r_{y_2}^{y_1}, r_{y_1}^{y_2}$  を生成する。

容姿特徴と姿勢特徴の抽出と合成を実現するために、提案手法は3種類の損失関数を導入する。第一の損失関数は、入力画像と再構成画像の再構成損失  $L_{cyc}$  である。この損失関数により、入力画像と合成画像から一貫した姿勢特徴を抽出し、その特徴に基づいた再構成画像の生成が可能となる。第二の損失関数は、同一人物間の合成画像と入力画像との再構成損失  $L_{same}$  である。この損失関数により、同一人物の画像から一貫した容姿特徴を抽出することが可能となる。第三の損失関数は、深層生成モデルの1つである DCGAN [2] でも使用される敵対的損失  $L_{adv}$  である。この損失関数により、 $D$  が入力画像と合成画像を識別できないような合成画像を  $G$  が生成するようになり、これによって合成画像の品質を高めることができる。提案手法は、これらの損失関数を用いることで、容姿・姿勢特徴の抽出と合成が可能となる。

\*連絡先： 広島市立大学  
〒731-3194 広島県広島市安佐南区大塚東 3-4-1  
E-mail: f20189@e.hiroshima-cu.ac.jp



(a) 異なる人物ペアの画像生成



(b) 同一人物ペアの画像生成

図 1: 提案手法の概要

表 1: 年齢分類におけるマクロ平均の  $F$  値の比較

拡張手法	学習データ数	マクロ平均の $F$ 値
-	100%	0.700
-	50%	0.680
Blur	50%	0.679
DCGAN	50%	0.675
Ours	50%	<b>0.697</b>

### 3 実験

提案手法によるデータ拡張の効果を検証するために、歩容画像からの年齢分類実験を実施した。実験には、OUMVLP データセット [3] を使用した。年齢分類器は、1 歩行周期のシルエット画像系列を平均して得られる Gait Energy Image (GEI) [4] という歩容画像を入力として、若者 (~14 歳)・成人 (15~64 歳)・高齢者 (65 歳 ~) の 3 クラスを分類する。本実験では、データ拡張によって学習データ数を増やした時の分類精度の変化から有効性を確認した。

提案手法によるデータ拡張は以下の手順で実施した。最初に、OUMVLP データセットの 50% の学習データを用いて、クラス毎に提案手法を学習する。その後、図 2 に示すように、学習データの各人物に対して同一クラスからランダムにペアを作成し、提案手法によってシルエット画像系列を合成する。最後に、合成画像系列

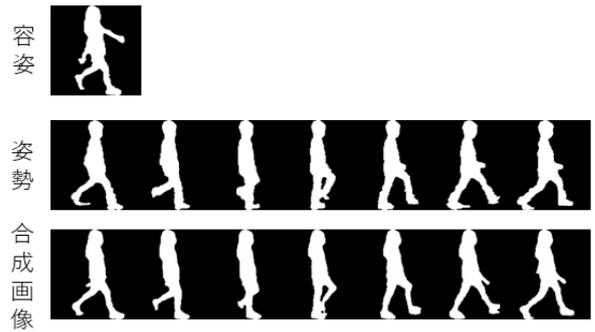


図 2: 容姿・姿勢特徴を合成した画像系列

から作成した GEI を学習データに加え、それを用いて年齢分類器を学習する。本実験では、ガウシアンフィルタによるぼかし (Blur) と、DCGAN による生成画像を用いたデータ拡張を適用した場合と、データ拡張を適用していない学習データによる学習結果と比較した。

表 1 に年齢分類の結果を示す。表 1 より、提案手法によって学習データ数を 50% から 100% に増やすことで分類精度が向上し、学習データ数が 100% の場合と同程度の精度まで向上していることが分かる。一方、比較手法はデータ拡張前の学習データ数が 50% の場合と比較して精度が向上しなかった。このことから、提案手法によるデータ拡張の有効性が確認された。

### 参考文献

- [1] Zhang, Z., Tran, L., *et al.*: Gait recognition via disentangled representation learning, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4710–4719 (2019).
- [2] Radford, A., Metz, L., *et al.*: Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks, *International Conference on Learning Representations*, (2016).
- [3] Takemura, N., Makihara, *et al.*: Multi-view Large Population Gait Dataset and its Performance Evaluation for Cross-view Gait Recognition, *IPSJ Transaction on Computer Vision and Applications*, Vol. 10, No. 4, pp. 1–14 (2018).
- [4] Han, J. and Bhanu, B.: Individual Recognition Using Gait Energy Image, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 2, pp. 316–322 (2006).