

歩容データによる個人の識別と認証の検討

前田 真由[†] 岡本 泰秀[†] 水野 秀之[‡]公立諏訪東京理科大学大学院 工学・マネジメント研究科[†]公立諏訪東京理科大学 工学情報応用工学科[‡]

1. はじめに

近年、生体特徴に基づく個人識別技術の一つである歩容を用いた識別や認証が監視システムへの応用という面から期待されている。解像度が低い映像においても適用可能なシステムが望まれていることから、これまで歩容シルエット画像を用いた方法などが提案されている^[1]。

本稿では、歩行動作を撮影した時系列の動画データを用いて個人の識別と認証を検討する。

2. 使用したデータと特徴量

一定の方向に歩行を行っている実験対象者を側面から一人あたり 50 回撮影した歩行動画から、それぞれ約 2 歩分(約 1 秒間)を切り出し 50 本の動画データを作成した。実験対象者は 20 名、フレームレートは 29.97 fps である。

専用の機器を用いることなく通常の画像から骨格情報を取得可能な OpenPose の COCO モデル^[2]を用いて右側面から観測可能な 10 点の Keypoint を取得した。No.1(Neck)の座標を基準とした各 Keypoint の相対角度、No.0(Nose)と No. 1 の距離値を 1 とした相対距離を求め特徴量として用いた。さらに各特徴量の 3 フレームずつの平均値を求め、1 フレームずつシフトし 14 フレーム単位で取り出すことで、1 動画データから 15 サンプル取得した。データ量は一人あたり 750 サンプル、20 名分で 15000 サンプルとなる。

3. 識別方式

個人識別のため、1次元の CNN と RNN を組み合わせたニューラルネットワークを用いた。図 1 にネットワーク構造を示す。識別対象者の 1 サンプル分のデータを入力すると、出力として 20 名に対する確率が得られる。

学習と評価については 10 クロスバリデーション

により行ったが、学習データ量が少ないため学習データにのみ 5 種類のランダムノイズを加えデータ量を 5 倍にすることで識別対象者一人あたり 3375 サンプルのデータを学習に用いた。テストデータは 75 サンプルである。従って実験に用いた総学習データ数は 67500 サンプル、総テストデータ数は 1500 サンプルとなる。

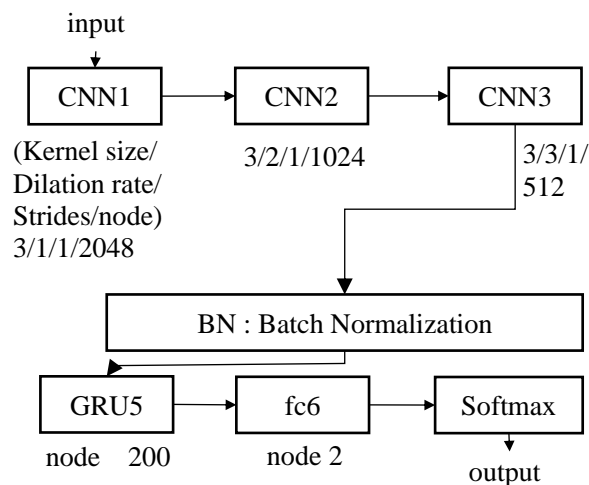


図 1 識別方式のネットワーク

識別対象者 20 名に対する再現率、適合率の最大値、最小値、平均値を表 1 に示す。正解率は約 99% となり、高い精度を得ることが出来た。再現率、適合率ともに高い数値を得られたことから、提案手法が、歩容データからの個人識別に有効であることが確認できた。

表 1 再現率、適合率 (各最大最小平均値)

	再現率	適合率
最大値	1.000	1.000
最小値	0.964	0.966
平均値	0.987	0.988

4. 認証方式

個人認証のため、Auto Encoder と Siamese Network^[3]を組み合わせたニューラルネットワークを用いた。図 2 にネットワーク構造を示す。

Personal identification and recognition based on gait information

[†]Mayu Maeda, Taishu Okamoto, Suwa University of Science

[‡]Hideyuki Mizuno, Suwa University of Science

認証対象者を A としたときの学習方法は、以下の通りである。

- 1) A の学習データのみを用いて Auto Encoder を学習する。
- 2) 全学習データを Auto Encoder に入力して再構築データを取得する。このとき、Auto Encoder の再構築誤差が最も少ない A 以外の実験対象者（以下 Z とする）を求める。
- 3) Z を除く実験対象者の Auto Encoder の入力データを Siamese Network の input1, 再構築データを Siamese Network の input2 に入力し、入力データが A か A でないかの 2 値分類で学習を行う。

上記を全 20 人分行うことで個人別に 20 個のネットワークを得た。

用いたデータは個人識別方式と同じであるがクロスバリデーションは行わず、あらかじめランダムに 10% の動画データ（5 本）を評価用に除外し、残り 90%（45 本）を学習に用いた。また認証対象者の学習データのみ個人識別と同様な方法で Data Augmentation を行った。

実験結果として、認証対象者の FRR と、Siamese Network の学習に用いた認証対象以外の 18 名の FAR(in), 学習に用いていない 1 名の FAR(out) の最大値, 最小値, 平均値を表 2 に示す。

実験結果から、平均値は比較的良好な値が得られ、対象者によっては FRR, FAR(in) とともに 0 という非常に良い結果が得られている一方、FAR, FRR(out) では 10% を超える例も存在した。特に FAR(out) の精度が低い理由として FAR(in) との精度差が極めて大きいことから、限られた人数のため学習データ内に学習から除外した実験対象者に類似したパターンが含まれず、十分に学習されていないことが原因として考えられる。

FRR の精度が FAR(in) と比べて極めて悪い例がある点については、Data Augmentation を行っているものの認証対象者とそれ以外でデータ量に偏りがある点が原因だと考えられる。

表 2 FRR, FAR(in), FAR(out)
(各最大最小平均値)

	FRR	FAR(out)	FAR(in)
最大値	0.213	0.52	0.031
最小値	0.000	0.000	0.000
平均値	0.033	0.146	0.007

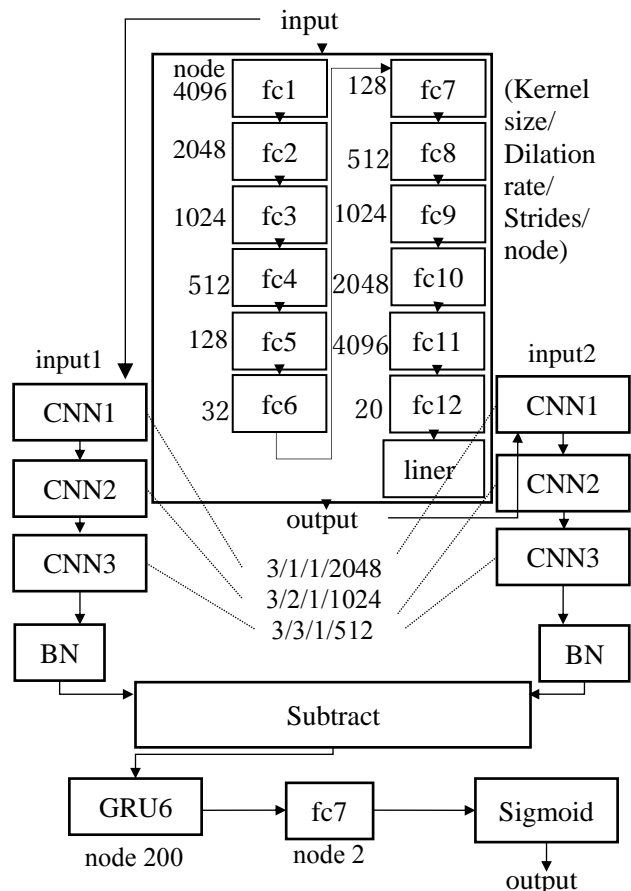


図 2 認証方式のネットワーク

5. おわりに

歩容データから時系列として得られた特徴量を用いた個人識別方式と個人認証方式を提案した。個人識別、個人認証それぞれの精度について検証し、高精度な個人識別が可能であることが確認できた。今後は、さらなる認証精度の向上とともに様々な角度から撮影した動画からの個人識別と個人認証の検討を行う予定である。

参考文献

- [1] N. TAKEMURA, et al., “Multi-view large population gait dataset and its performance evaluation for cross-view gait recognition.”, IPSJ T Comput Vis Appl 10, 4 (2018).
- [2] Z.Cao, et al., “Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields”, In Proc. of CVPR, pp. 7291-7299, (2017).
- [3] S. Chopra, et al., “Learning a similarity metric discriminatively, with application to face verification.”, In Proc of CVPR, pp.539-546, (2005).