

多段階畳み込みニューラルネットワークを用いた 歩容に基づく年齢推定

阪田 篤哉^{1,a)} 武村 紀子^{1,†1} 八木 康史¹

概要: 歩容に基づく年齢推定はその実用性の高さから盛んに研究されており、幅広い応用が考えられる。本稿では、畳み込みニューラルネットワーク (CNN) を用いた歩容特徴を入力とする年齢推定手法を提案する。性別や年齢などの人物属性により歩容特徴に現れる特性が大きく異なるため、性別や大まかな年代層を推定した上で各属性に特化した年齢推定を行う。大規模歩容データベースを用いた評価実験の結果、既存手法を上回る性能が示された。

キーワード: 畳み込みニューラルネットワーク, 歩容, 年齢推定

SAKATA ATSUYA^{1,a)} TAKEMURA NORIKO^{1,†1} YAGI YASUSHI¹

1. はじめに

人物の画像から年齢を推定する技術は様々な応用が考えられ、近年、盛んに研究されている。年齢推定に関する従来手法の多くは、シワやたるみ等、経年変化が出やすい顔画像を用いたものである [1], [2], [3], [4], [5], [6]。しかし、顔画像を用いた年齢推定では、正面を向いた高解像度な顔画像が必要であるため、その使用用途が限られる。例えば、酒類やたばこを購入する際の年齢確認やデジタルサイネージをはじめとした広告の注視者の人物属性取得など、近距離から正面顔を撮影できる状況でしか年齢推定ができない。

一方、人間の歩き方の特徴である歩容特徴は、遠方から撮影した低解像度の画像からユーザの協力なしに抽出することができる。また、歩容特徴を用いた個人認証の研究において、歩容特徴は個人を識別できるだけの特徴量を有していることが実証されており [7]、年齢に関する情報も含まれることが期待される。そこで、本研究では歩容特徴に基づいた年齢推定を行う。歩容の利点を活かして遠方から人物年齢の推定を行えば応用の範囲が更に広がる。例として、広域監視カメラを用いた犯罪捜査や迷子・徘徊老人の

発見だけでなく、ショッピングモールまたは駅・空港など大規模施設における人々の往来を解析し、年齢層別計数などを行うといったマーケティング分野などの応用が期待されている。

歩容特徴を用いた年齢推定に関する研究もいくつかなされている。横原ら [8] はガウス過程回帰に基づいた年齢回帰アルゴリズムを提案している。また、Lu ら [9] は同じ年齢でも男女間で歩き方の外観が異なることから、年齢と性別との相関を判別空間を用いて学習し推定を行う手法を提案している。さらに、Lu ら [10] の別の研究では年齢差から形成された多様体を学習し、多重線形回帰を行う手法を提案している。これらの手法では、歩容画像の各画素値を互いに独立した特徴量として扱うため、本来画像がもつ空間的な局所性・近接性といった情報が失われ、結果、過学習を起こしやすいという問題がある。これらの問題に対処するため、本研究では、近年画像認識の分野において大きな成功を収めている畳み込みニューラルネットワーク (CNN) を用いた年齢推定手法を提案する。

歩容特徴は性別や年代に応じて様々な特性が現れると言われており [11]、全ての性別・年代を網羅する年齢推定器を学習するのは容易ではない。そこで、本研究では、CNN を用いて各入力の性別および年代を個別に推定し、性別・年代ごとに学習された年齢推定器を用いて年齢推定を行う。

本稿では、2 歳から 90 歳までの男女のデータからなる世界最大の歩容データベース、OU-ISIR Gait Database、

¹ 大阪大学産業科学研究所
The Institute of Scientific and Industrial Research, Osaka University

^{†1} 現在、大阪大学データビリティフロンティア機構
Presently with The Institute for Datability Science, Osaka University

^{a)} sakata@am.sanken.osaka-u.ac.jp

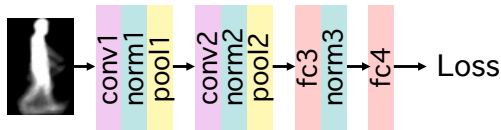


図 1 ネットワーク構造

表 1 各層の設定

層	カーネル数	サイズ/ストライド	ユニット数
conv1	81	5×5/1	-
pool1	-	3×3/2	-
conv1	45	7×7/1	-
pool2	-	2×2/2	-
fc3	-	-	1024
fc4	-	-	1

Large Population Dataset with Age (OULP-Age) [12] を用いて性能評価実験を行い、本提案手法の有用性を示す。

2. CNN を用いた年齢推定

本稿では、歩容特徴 Gait Energy Image(以下, GEI) [13] を入力とした CNN による年齢推定手法を提案する。GEI は手脚の振り方といった動的特徴と姿勢や体型といった静的特徴を併せ持つ特徴であり、歩容を用いた個人認証で一般的に用いられる特徴である。

2.1 単一の CNN を用いた年齢推定手法

年齢推定器のネットワーク構造を図 1 に、各層のパラメータを表 1 に示す。本 CNN は、GEI を入力画像とし畳み込み層 (conv), 最大プーリング層 (pool), バッチ正規化層 (norm) の 3 つ組 2 層と、全結合層 (fc), バッチ正規化層の 2 つ組 1 層, および全結合層 1 層からなる。畳み込み層および全結合層の活性化関数として、ReLU を採用する。

学習の際は、CNN の重みパラメータの初期値を He の手法を用いて設定し、バイアスパラメータの初期値は 0 とする。パラメータの更新は Adam を用いて行い、学習率は初期値を 0.001 とする。さらに、fc3 層, fc4 層において Dropout (選出確率はそれぞれ 0.8, 0.5) を適用する。最終層の出力を推定年齢と見なし、推定年齢と実年齢の平均絶対誤差が小さくなるように年齢推定器を学習する。

文献 [7] でも言及されているように、GEI を入力とする識別問題では、入力の多様性が少ないため、物体認識に用いられる一般的な CNN と比較して、浅いネットワーク構造で特徴を表現することが可能である。

2.2 多段階の CNN を用いた年齢推定手法

歩容データベース OULP-Age に含まれるデータについて、各性別、各年代の GEI の平均画像を図 2 に示す。図 2 より、頭と体の比率、髪型、体型、姿勢など、性別および年代によって歩容特徴が異なることが分かる。

以上より、性別、年代ごとに年齢推定器を学習すること

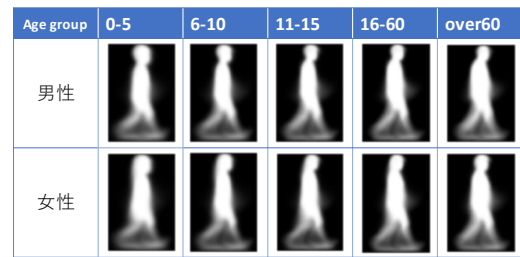


図 2 男女別、年代層別の平均 GEI

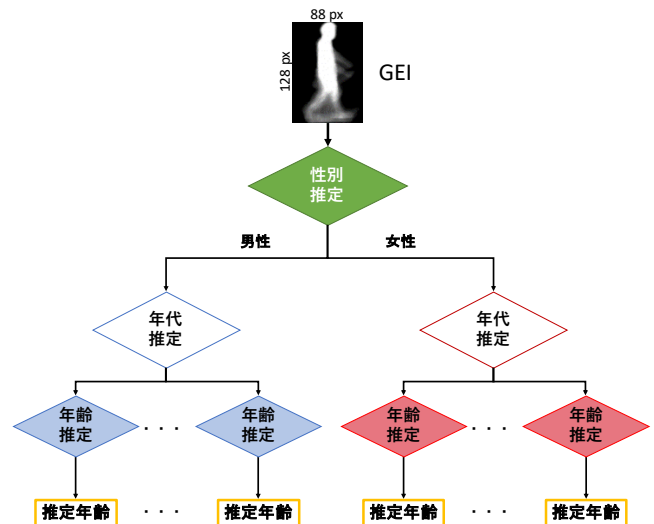


図 3 多段階年齢推定器のフローチャート

で、年齢推定精度の向上が期待できる。本稿では、図 3 のように、CNN を用いた、性別推定器、年代推定器、年齢推定器を順に適用する多段階 CNN により、年齢推定の精度向上を図る。ただし、推定器の順序は、識別が容易な順に性別推定器、年代推定器と配置した。

各推定器のネットワーク構造は図 1 と同様のものを用いる。ただし、性別推定器では、最終層出力にシグモイド関数を適用したときの交差エントロピーを Loss 関数とした。また、年代推定器では、fc4 層のノード数を 5 (年代クラス数) とし、最終層出力にソフトマックス関数を適用したときの交差エントロピーを Loss 関数とした。

2.2.1 年齢推定器の学習手順

多段階 CNN における年齢推定器の学習手順を以下に示す。

- (1) 全ての性別・年代が含まれるトレーニングセットを用いて性別推定器を学習する。
- (2) 学習済みの性別推定器に (1) の学習に用いたトレーニングセットを入力し、性別推定を行う。
- (3) 各性別に推定されたトレーニングセットを用いて推定性別ごとに年代推定器を学習する。
- (4) 学習済みの年代推定器に、(3) の学習に用いたトレーニングセットを入力し、性別ごとに年代推定を行う。
- (5) 各性別、各年代に推定されたトレーニングセットを

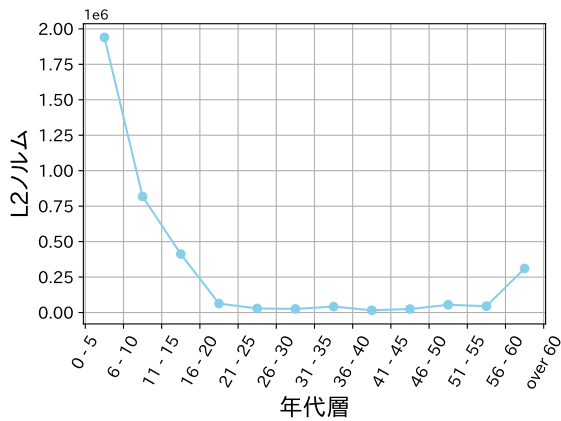


図 4 各年代の平均 GEI の差による L2 ノルムの推移 (OULP-Age)

用いて、推定性別・推定年代ごとに年齢推定器を学習する。

年代推定器は推定性別ごとに、年齢推定器は、推定性別および推定年代ごとに学習するため、学習データ数が絶対的に少なくなり、過学習を起こす可能性が高くなる。これを防ぐために、年代推定器を学習する際は、全性別・全年代のデータを用いて学習した年代推定器からファインチューニングを行い、年齢推定器を学習する際は、性別ごとのデータを用いた年齢推定器からファインチューニングを行った。

2.2.2 年代クラス的设计

多段階 CNN において、年代を推定する際の年代クラスの設定方法について述べる。本稿では、GEI の類似性に着目して年代クラスを設定する。まず、OULP-Age を 5 歳ごとのグループに分割し、各グループの GEI の平均画像を作成する。ただし、60 歳以上のデータが他の年齢層と比較してかなり少ないため、60 歳以上は同一グループとする。次に、隣り合うグループの平均 GEI の L2 距離を求める (図 4)。L2 距離が一定値以内のグループを同じクラスと定義し、0-5, 6-10, 11-15, 16-60, over 60 の 5 クラスを設定した。

3. 評価実験

3.1 データベース

本稿では、OU-ISIR Gait Database, Large Population Dataset with Age (OULP-Age) [12] を用いて、年齢推定手法の性能評価を行う。OULP-Age は 2 歳から 90 歳の 63,846 人 (男性 31,093 人, 女性 32,753 人) の歩容画像データからなる大規模なデータベースである。各被験者について、側面から撮影された歩行映像から生成された GEI が 1 枚ずつ存在し、各 GEI に対して性別と年齢のラベルが付与されている。

Xu らの行った OULP-Age に対する既存の年齢推定手法の実行結果を提案手法との比較として用いるため、ベンチマークのプロトコルを揃えて実験した。このプロトコルで

表 2 各手法の平均絶対誤差

手法	MAE [歳]
Single-CNN	6.22
Multi-CNN	5.84
GPR [8]	7.30
SVR [2]	7.66
OPLDA [10]	8.45
MLG [9]	10.98

は、無作為抽出によりトレーニングセットとテストセットを男女比および年齢比が 1 : 1 になるように調整してある。また、トレーニングセットをさらに学習用とバリデーション用に 8 : 2 の割合で分割し学習を行った。

3.2 評価指標

実験結果の評価は年齢推定器により得られた年齢に対して行う。この評価ではテストセットに対する実年齢 t_n と推定年齢 y_n との平均絶対誤差 (MAE: Mean Absolute Score) を用い、他手法との比較を行う。

$$MAE = \frac{1}{N} \sum_{n=1}^N |t_n - y_n| \quad (1)$$

さらなる評価指標として累積スコア (CS: Cumulative Score) を用いる。累積スコアは N 人のテストデータのうち推定年齢と実年齢との絶対誤差が l 歳以下となる被験者数 N_l の割合として定義される。

$$CS(l) = \frac{N_l}{N} \times 100\% \quad (2)$$

3.2.1 比較手法

本実験では、文献 [12] で用いられているプロトコルを用いて実験を行い、以下の 6 手法について比較を行う。

Multi-CNN : 単一の CNN を用いて年齢回帰を行う手法 (提案手法)

Single-CNN : 多段階の CNN を用いて性別推定、年代推定、年齢回帰を順に行う手法 (提案手法)

GPR [8] : ガウス回帰に基づく手法

SVR [2] : サポートベクター回帰に基づく手法

OPLDA [10] : 年齢差から形成された多様体を学習し、多重線形回帰を行う手法

MLG [9] : 年齢と性別との相関を判別空間を用いて学習し、年齢推定を行う手法

3.3 推定結果

提案手法および従来手法での推定結果の MAE を表 2 に示す。

表 2 によると、多段階年齢推定器の MAE が最小となった。これにより年齢のみについて回帰推定を行うよりも、性別や年代層を考慮した上で回帰推定を行うことが精度向上につながる事が示された。Single-CNN に関しても、

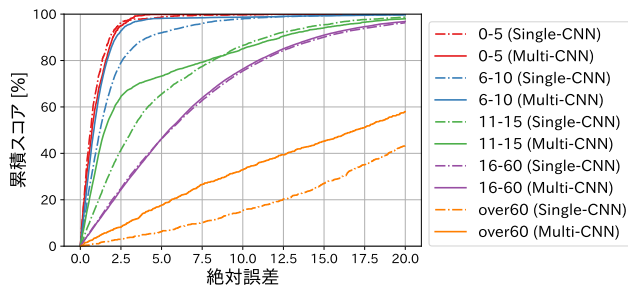


図 5 提案手法と Single CNN の年代層別累積スコア

CNN を用いない既存手法に比べ MAE が小さくなっており、CNN の歩容に基づく年齢推定においてその有効性が示された。

図 5 は 2 つの提案手法の累積スコアを年代層クラス別に表したものである。0-5、16-60 の 2 クラスではほとんど累積スコアは変わらなかったが、6-10、over 60 の 2 クラスに対して多段階モデルで精度が大きく改善された。反対に 11-15 クラスでは絶対誤差が 7.7 歳を超えた時点で 1 段階モデルが累積スコアを上回っている。このクラスは成長期にあたるため体型変化の個人差が著しく、16-60 に誤分類されたデータが累積スコアに大きく影響を及ぼしていることが原因と考えられる。

3.4 考察

3.4.1 散布図の考察

図 6 はテストセットの各データの実年齢を横軸に、推定器による推定年齢を縦軸にプロットした散布図である。各データは推定年代層ごとに色分けしてある。

実年齢 0-5、6-10 の 2 クラスに属するデータは、年齢推定の誤差が小さく、年代分類の精度も高い。これは GEI の作成時に高さに基づくスケリングを行っているために頭の大きさが特徴的に現れ、年代推定および年齢推定が容易となることが推測される。

また 11-15 クラスに分類されたデータのうち分類を誤っているものは、年代推定においても 11~15 歳程度と推定されている。さらに over 60 クラスに分類されたデータのうち分類を誤っているものも同様に、年齢推定において 60 歳以上程度と推定されている。このことから、年齢推定は年代層分類の精度に大きく依存しているものと考えられる。

加えて実年齢が大きくなるにつれ推定年齢のばらつきが大きくなっていることが見て取れる。これは高齢者ほどデータ数が少ないことに加え、年代層推定器の推定結果が被験者の体型や撮影時の歩行姿勢（例：60 歳以上の被験者が姿勢を正して歩いている）により左右されやすいことが原因と考えられる。

3.4.2 失敗例の考察

図 7 に推定年齢を誤った例として 4 つのケースを示す。各ケースに該当する GEI を上段：男性、下段：女性として

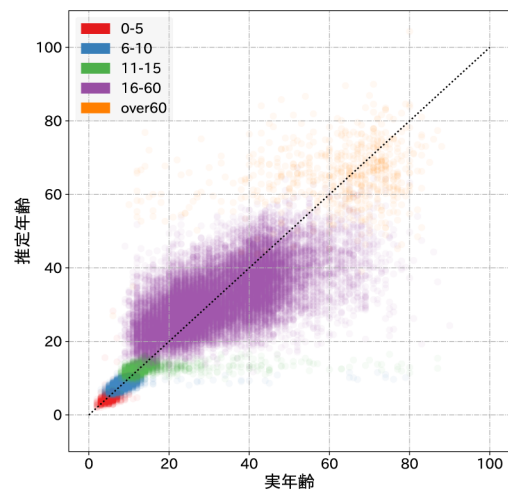


図 6 テストセットの実年齢と推定器による推定年齢との散布図。黒点線は実年齢と推定年齢が等しい場合を表す。

示す。ただし、推定年齢は小数点以下を四捨五入して表記してある。

まずケース 1 は、高齢であるにもかかわらず 13 歳や 28 歳など若年者として推定されている例である。これらの GEI を観察するといずれも背筋を正して歩いた様子がわかり、これが推定に影響を及ぼしていることが考えられる。

ケース 2 は成人年齢にもかかわらず 10 歳未満として推定されている例である。図 2 にも示されているように、10 歳未満の GEI は平均的に全身に対して頭の大きさの占める割合が大きい。そのため成人であるにもかかわらず、このような特徴を持つデータは 10 歳未満に誤推定されてしまう。

ケース 3 は実年齢に対して高く推定された例である。男性はコートを着用している影響で身体と頭の大きさの比率が成人に近くなったことにより成人として推定されている。女性も同様に服装が原因で体の部分が over 60 クラスの特徴に近くなっているため、高齢者として推定されたと考えられる。

ケース 4 は年齢推定に加え性別推定で誤りを生じた例である。男性の例を見ると髪型が長いように見えるが実際はコートのフード部分が重なっており、髪の部分とつながって見えるようになるため、女性に誤分類されたものと考えられる。対して女性の例は図 2 で示したように高齢であるほど髪型が短くなる傾向があるため、男性に誤分類されやすいものと考えられる。

4. おわりに

本稿では、CNN を用いた歩容特徴 GEI を入力とする年齢推定手法を提案した。この手法は年齢だけでなく、まず性別や年代層といった別の人物属性を推定する CNN をそれぞれ学習させて段階的に構成することにより各特性を踏まえた推定を実現した。本手法を用いて大規模歩容デー


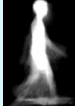






	ケース1	ケース2	ケース3	ケース4
男性				
実年齢	77	31	9	30
推定年齢	13	9	23	44
推定性別	男性	男性	男性	女性
女性				
実年齢	61	25	16	72
推定年齢	28	8	60	30
推定性別	女性	女性	女性	男性

図 7 推定失敗例。

データベースにより学習・テストを行った結果、平均絶対誤差 (MAE) が 5.84 歳と既存の年齢推定手法を上回る性能を示した。今後の展望として推定器の精度向上が挙げられる。CNN に年齢の単一のラベルではなく性別・年代層ラベルを加えて同時に学習を行うマルチタスク学習や、目的関数および年代層境界の再定義などが考えられる。

参考文献

[1] Geng, X., Yin, C. and Zhou, Z. H.: Facial age estimation by learning from label distributions, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 35, No. 10, pp. 2401–2412 (online), DOI: 10.1109/TPAMI.2013.51 (2013).

[2] Guo, G., Fu, Y., Dyer, C. R. and Huang, T. S.: Image-based human age estimation by manifold learning and locally adjusted robust regression, *IEEE Transactions on Image Processing*, Vol. 17, No. 7, pp. 1178–1188 (online), DOI: 10.1109/TIP.2008.924280 (2008).

[3] Fu, Y. and Huang, T. S.: Human age estimation with regression on discriminative aging manifold, *IEEE Transactions on Multimedia*, Vol. 10, No. 4, pp. 578–584 (online), DOI: 10.1109/TMM.2008.921847 (2008).

[4] Zhang, Y. Z. Y. and Yeung, D.-Y. Y. D.-Y.: Multi-task warped Gaussian process for personalized age estimation, *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 2622–2629 (online), DOI: 10.1109/CVPR.2010.5539975 (2010).

[5] Niu, Z., Zhou, M., Wang, L., Gao, X. and Hua, G.: Ordinal Regression with Multiple Output CNN for Age Estimation, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4920–4928 (online), DOI: 10.1109/CVPR.2016.532 (2016).

[6] Escalera, S., Fabian, J., Pardo, P., Baro, X., Gonzalez, J., Escalante, H. J., Misevic, D., Steiner, U. and Guyon, I.: ChaLearn Looking at People 2015: Apparent Age and Cultural Event Recognition Datasets and Results, *Proceedings of the IEEE International Conference on Computer Vision*, Vol. 2015-Febru, pp. 243–251 (online), DOI: 10.1109/ICCVW.2015.40 (2015).

[7] Takemura, N., Makihara, Y., Muramatsu, D., Echigo, T. and Yagi, Y.: On Input/Output Architectures for Convolutional Neural Network-Based Cross-View Gait Recognition, *IEEE Transactions on Circuits and Sys-*

tems for Video Technology, Vol. PP, No. 99, pp. 1–1 (online), DOI: 10.1109/TCSVT.2017.2760835 (2017).

[8] Makihara, Y., Okumura, M., Iwama, H. and Yagi, Y.: Gait-based age estimation using a whole-generation gait database, *2011 International Joint Conference on Biometrics, IJCB 2011*, (online), DOI: 10.1109/IJCB.2011.6117531 (2011).

[9] Lu, J. and Tan, Y. P.: Gait-based human age estimation, *IEEE Transactions on Information Forensics and Security*, Vol. 5, No. 4, pp. 761–770 (online), DOI: 10.1109/TIFS.2010.2069560 (2010).

[10] Lu, J. and Tan, Y. P.: Ordinary preserving manifold analysis for human age and head pose estimation, *IEEE Transactions on Human-Machine Systems*, Vol. 43, No. 2, pp. 249–258 (online), DOI: 10.1109/TSMCC.2012.2192727 (2013).

[11] 万波秀年, 榎原 靖, 八木康史: 歩容における性別・年齢の分類と特徴解析, *電子情報通信学会論文誌 D*, Vol. J92-D, No. 8, pp. 1373–1382 (2009).

[12] Xu, C., Makihara, Y., Ogi, G., Li, X., Yagi, Y. and Lu, J.: The OU-ISIR Gait Database Comprising the Large Population Dataset with Age and Performance Evaluation of Age Estimation, *IPSSJ Trans. on Computer Vision and Applications*, Vol. 9, pp. 1–14 (online), DOI: 10.1109/TIFS.2012.2204253 (2017).

[13] 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 (online), DOI: 10.1109/TPAMI.2006.38 (2006).