

## 服装と顔画像にもとづく年齢推定手法の提案

兼子 大起<sup>†</sup> 郷古 学<sup>†</sup><sup>†</sup>東北学院大学 工学部 情報基盤工学科

## 1 はじめに

近年、機械学習を用いて画像から人物属性（年齢や性別）を推定する研究が行われている [1, 2, 3]. 特に、年齢推定はセルフレジでの年齢認証 [4] や、パチンコ店の入店制限など、幅広い応用が期待できる。

これまでに、顔画像にもとづく属性推定手法が提案されている [1]. 従来手法には、推定にはカメラに対する顔の向きや、一定の解像度を要する等の制約がある。

そこで本研究では、人物の全身画像を用いることで、服装の情報にもとづく属性推定手法を提案する。提案手法では、顔と服装の画像から個別に推定を行い、それらの結果を統合して、最終的な推定結果とする。同手法を実現するため、本研究では新たに服装からの属性推定モデルを構築した。本稿では、服装からの属性推定モデルの性能評価と、提案手法の有効性について報告する。

## 2 従来研究

従来研究では、手や顔などの特定部位の画像を用いた属性推定手法が提案されている [1, 2]. これらの手法では、各部位がカメラに正対している必要があり、また解像度が低い場合や、マスクやサングラスなどの着用も推定精度の低下を招く。

全身画像を用いた属性推定手法も提案されている [3]. この手法は、全身画像と姿勢推定アルゴリズム (OpenPose) で取得した骨格情報から属性推定を行っている。同研究では、学習データとして用いた服装の種類が乏しい点が問題として挙げられている。

## 3 提案手法

図1に提案手法の概要を示す。提案手法は、入力される人物の全身画像から、その人物の属性（成人女性、成人男性、未成年女性、未成年男性の4種類）を推定する。入力画像は人物の全身画像とし、入力画像のサイズは224×224ピクセルとした。

提案手法では、入力される「全身画像（服装）」とそこから抽出された「顔画像」から、個別に属性を推定し、それらの結果を重み付けの上、統合し最終的な推定結果とする。提案手法は、顔と服装から個別に推定

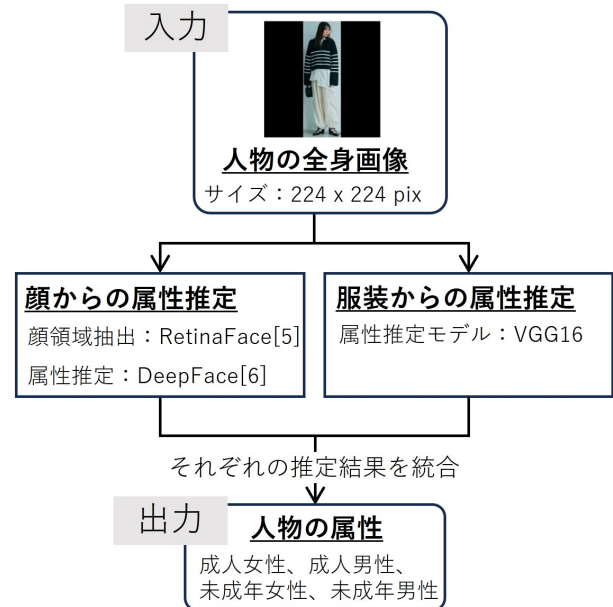


図1: 提案手法

を行うため、例えばマスクの着用により顔からの推定結果の信頼性が低い場合でも、服装からの推定結果を用いることで、全体として頑健な推定が期待できる。

## 3.1 顔画像にもとづく属性推定

顔画像にもとづく属性推定は、既存技術を用いて実現した。入力画像からの顔領域の抽出には RetinaFace [5] を、抽出された顔画像からの属性推定は DeepFace [6] を用いた。

## 3.2 服装にもとづく属性推定

機械学習により、服装からの属性推定モデルを構築した。はじめに、推定モデルの学習データとして、ファッション誌とコーディネートサイトから人物のスナップ写真を収集した。各属性800枚ずつ、合計3,200枚の学習データを収集した。

また、学習データとは別に評価用データとして、各属性200枚ずつ、合計800枚の画像を同様に収集した。なお、学習データおよび評価データの教師ラベルは、データ収集に使用したファッション誌の購買層およびファッションモデルの年齢から判断して作成した。

収集した学習データを用いて、属性推定モデルの学習を行った。モデルには畳み込みニューラルネットワークの VGG16 を使用した。

Age estimation based clothing and face image

<sup>†</sup> Daiki KANEKO (s2045129@g.tohoku-gakuin.ac.jp)<sup>†</sup> Manabu GOUKO (gouko@mail.tohoku-gakuin.ac.jp)

Department of Information Technology, Faculty of Engineering, Tohoku Gakuin University (†)

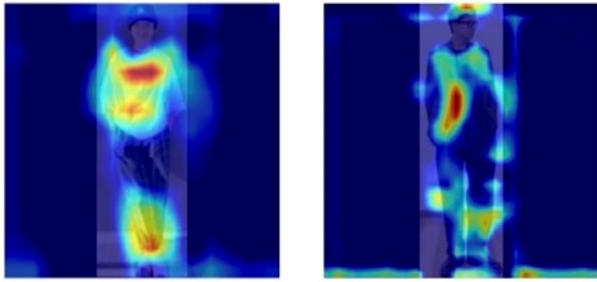


図 2: ヒートマップ

#### 4 推定精度の検証および考察

服装からの属性推定モデルと提案手法による推定精度を検証した。

##### 4.1 服装からの属性推定

服装からの属性推定モデルの推定精度は 69.5%であった。同モデルにおいて、入力画像のどの領域が推定に大きく影響するのかを確認するため、モデルに可視化技術の1つである Grad-CAM を適用した。

Grad-CAM で得られた入力画像に対するヒートマップの一例を図 2 に示す。これらの入力画像はいずれも正しく推定できた画像である。ヒートマップは推定に大きく影響を与えるピクセルほど赤色で表示される。

図 2 から、推定に影響を与える領域は主に上半身(肩から胸元にかけて)と足元であることを確認した。一方で、顔領域が推定に与える影響は比較的小さいことも確認した。

##### 4.2 提案手法(顔 + 服装)による属性推定

提案手法を用いた属性推定精度の検証を行った。検証には、新たに収集した 200 枚の人物画像を用いた。画像の収集法は 3.1 と同様である。

表 1 に、提案手法の推定精度と、顔画像および服装の情報のみを用いた場合の属性推定の精度を示す。同表で、「年齢」は未成年か成人かを、「性別」は女性か男性かを推定した結果であり、「全体」はそれらを組み合わせた 4 種類の属性の推定結果である。なお本検証では、提案手法における顔と服装からの推定結果の統合については、実験者がマニュアルで調整した。

提案手法による「年齢」、「性別」、「全体」の推定精度は、それぞれ約 84%、91%、80%であった。また、顔画像のみを用いた場合の推定精度が全体的に低かった。これは、本検証で用いた画像の多くで、顔領域が小さかったことが原因だと考えられる。提案手法では、このように顔領域が小さい(解像度が低い)場合でも、服装の情報から正しく推定することが期待できる。

表 2 に提案手法による属性ごとの推定精度を示す。これを見ると、成人男性に対する推定精度が他の属性

表 1: 提案手法による推定精度 (%)

	提案手法	顔のみ	服装のみ
年齢	<b>84.0</b>	51.0	85.0
性別	<b>90.5</b>	61.0	89.5
全体	<b>79.5</b>	34.0	78.5

表 2: 属性ごとの推定精度 (%)

成人女性	成人男性	未成年女性	未成年男性
86.0	74.0	86.0	90.0

に比べて低いことがわかる。これは、他の属性と比較して、成人男性の学習データは、簡素なデザインの服や無地の服が多く、服装のバリエーションが乏しいことが影響していると考えられる。

#### 5 まとめと今後の課題

本研究では、顔と服装の画像を用いた属性推定手法を提案した。服装からの属性推定モデルを構築し、提案手法の推定精度を検証した結果、約 80%の精度で属性推定が可能であることを確認した。

今後の課題として、顔と服装からの推定結果の統合の自動調整の実現が挙げられる。また推定精度向上に向けて、学習データにおける、成人男性の服装のバリエーションを増やすとともに、バッグなどのアイテムについても考慮する必要がある。

#### 参考文献

- [1] Levi G. and Hassner T., "Age and gender classification using convolutional neural networks," Proc. of the IEEE conf. on computer vision and pattern recognition workshops, pp. 34-42. 2015.
- [2] M. A. Abderrahmane, *et al.*, "Human age prediction based on hand image using multiclass classification," 2020 Int'l conf. on data analytics for business and industry: way towards a sustainable economy (ICDABI), pp. 1-5. IEEE, 2020.
- [3] 井上, 梅原, 今井, 神谷, 田中, 米田, "ビデオカメラを用いた人物属性の推定に関する研究," 情報処理学会第 84 回全国大会, pp.629-630, 2022.
- [4] 日本フランチャイズチェーン協会, デジタル技術を活用した酒類・たばこ年齢確認ガイドライン, 2022-11, <https://www.jfa-fc.or.jp/particle/4061.html>, (参照 2023-07-31).
- [5] J. Deng, *et al.*, "Retinaface: single-stage dense face localisation in the wild," Proc. of the IEEE/CVF conf. on computer vision and pattern recognition, pp. 5203-5212, 2020.
- [6] Y. Taigman, *et al.*, "Deepface: Closing the gap to human-level performance in face verification," Proc. of the IEEE conf. on computer vision and pattern recognition, pp. 1701-1708. 2014.