

時空間情報を活用した人物属性推定手法の精度改善

手嶋はずき[†] 神山剛[†]

[†]長崎大学 情報データ科学部

1. はじめに

近年、スマートシティの実現に向けて、都市における様々な情報を取得し、サービス連携させる基盤システムの構築が進められている。本研究では、ドライブレコーダで収集したデータから街中の歩行者と、性別や年代などの属性を推定し、道路上の人流情報を得るための AI の実現を目指している。本稿では、属性推定の精度改善のためのアプローチの 1 つとして、学習・推定時のデータセットに時空間情報を活用する手法を提案する。

2. 現状の人物属性推定手法とその課題

2.1. 従来の人物属性推定手法

我々が開発中の人物属性推定手法は図 1 に示すとおりである。ドライブレコーダの動画から、YOLO[1]による物体検出により人物画像を抽出し、この画像を属性推定 AI に入力することで画像の人物の属性推定結果を得る。属性推定 AI は、YOLO で切り取った人物画像と、人物属性を付与したアノテーションデータを含む教師データを学習させることで作成している。なお、この AI の実装には ALM[2]を用いている。

2.2. 従来の人物推定手法の問題

従来の人物属性推定モデルには、精度の点で課題がある。一般的にこのような推定を行うモデルは、正解率が 80%以上、適合率と再現率がともに 70%以上で精度がよいとされるが、我々のモデルの正解率は平均で 60%弱、精度が最も低い項目である年代の推定では 20%弱であり、さらなる精度改善が必要である。また、我々以外の先行研究[3]においても、属性の種別によって正解率が 50~80%台と大きな開きがある。十分な推定精度が得られない原因として教師データの品質や量の問題などいくつかの要因が考えられるが、このアプローチだけで精度向上させるのは容易ではない。

3. 時空間情報を活用した手法の提案

本稿では、人物属性推定の精度改善のための手法として、人物画像が撮影された時間と場所に関する情報（以下、時空間情報）を活用することを提案する。提案する人物属性推定手法は図 2 に示すとおり

である。モデルの作成時に用いる教師データに人物画像が撮影された時空間情報を追加、また、推定時のデータにも時空間情報を追加するという手法である。例えば空間情報には撮影された場所の POI 情報、時間情報には、曜日や時間帯、季節を含む情報などを用いることを提案する。

4. 時空間情報の有効性の検証

人物属性推定に対する時空間情報の有効性を検証するため、時間や場所の違いでそこに現れる人物の属性に異なる傾向が確認できるか検証する。

検証には、図 3 に示す長崎市内の国道 206 号線上を、2023 年 11 月の平日の各曜日 1 日ずつ合計 5 日間の、8-15 時台に 1 時間おきにドライブレコーダで撮影したデータを用いる。分析にあたっては、周辺に商店街がある住吉、教育機関である長崎大学、オフィスや商業施設が集まる大波止、観光地であるグラバー園付近の 4 区間にわけ、各区間における歩行者の属性の傾向を確認する。

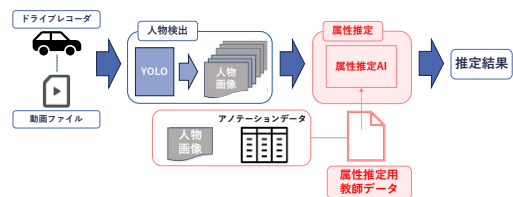


図 1 従来の人物属性推定手法

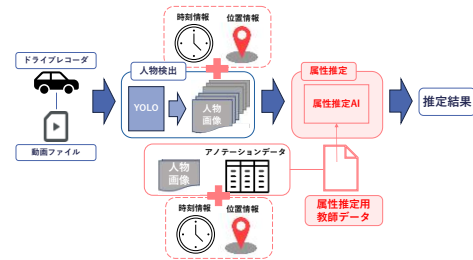


図 2 提案手法



図 3 検証に用いる長崎市内の道路

Improving the accuracy of Pedestrian Attribute Recognition using spatiotemporal information

Hazuki TESHIMA[†] Takeshi KAMIYAMA[†]

[†]School of Information and Data Science, Nagasaki University

また、あらかじめ人物属性推定 AI の教師用データを作成するのと同様に、検証に用いるドライブレコーダのデータに対しても人物画像の抽出と人手によるアノテーション作業により、3938 件の人物画像に対する属性情報を付与した検証用データセットを作成した。

4.1. 空間情報の有効性

前述した 4 区間における人物の属性の比較を行った。場所ごとの男女比、年代比をそれぞれ図 4、図 5 に示す。なお、男女比、年代比ともに、データ収集期間の合計人数(アノテーション結果が不明であるものを含む)に対する各項目の割合を算出した。男女比に関しては、商店街が近くにある住吉付近ではほかの場所に比べて女性の割合が多かった。年代比に関しては、グラバー園周辺では 12-19 歳、長崎大学周辺では 20-34 歳、オフィスが多い大波止では 35-49 歳、住吉では 50 歳以降の割合が多いという傾向がみられた。

4.2. 時間情報の有効性

前項と同じ 4 区間それぞれについて、8:00~16:00 の各時間帯における人物属性の傾向を分析した。

属性の違いが特に顕著であった大波止周辺における男女比と年代比をそれぞれ図 6、図 7 に示す。

男女比については、朝の通勤時間に当たる 8 時台や、昼の 12-13 時台の時間帯は特に男性の割合が高く、9 時台や 14 時以降の時間帯では女性の割合が高いという傾向がみられた。

年代比については、8-9 時台や 12-13 時台では 20-40 歳代の割合が高く、10-11 時台では、50 歳以上の年代の割合が高いという偏りがみられた。

その他の区間では、住吉では 8 時台と 12 時台で、男性と 34 歳以下の若い世代の割合が高く、グラバー園では 14 時台のみ 20-34 歳の割合が高かった。長崎大学では時間による大きな属性の違いがみられなかった。

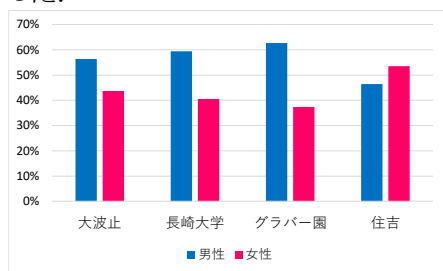


図 4 場所の違いによる男女比の比較

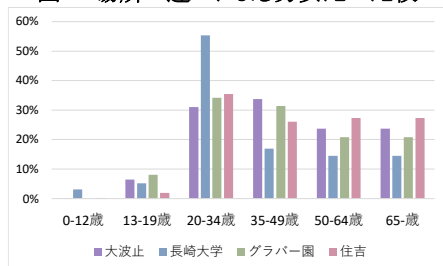


図 5 場所の違いによる年代比の比較

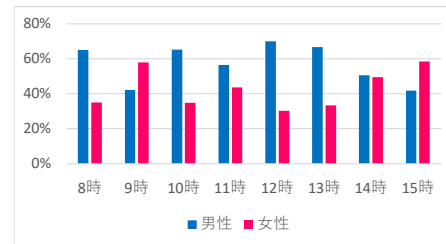


図 6 大波止:各時間帯の男女比の推移

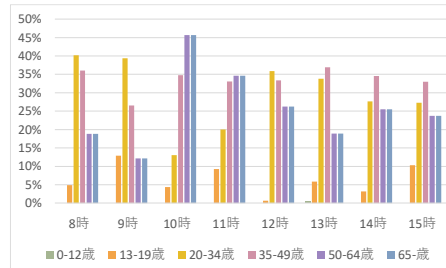


図 7 大波止:各時間帯の年代比の推移

5. おわりに

本稿ではドライブレコーダの動画から歩行者の人物属性推定を行う際の精度を改善するための新たな手法として時空間情報を活用することを提案した。

検証の結果、時空間の違いによって、その場所の人物の属性に偏りが確認できたことにより、提案手法による精度改善の可能性が示された。今後は、教師データに追加すべき時空間情報を具体的に定め、学習と精度評価を実施する予定である。

謝辞

本研究の一部は、科研費 22H03580 の支援を受けたものである。

本稿の執筆にご助言くださいました株式会社リクルートの塩澤繁様、鈴木教嗣様、山本慎也様に心から感謝いたします。

参考文献

- [1] Joseph Redmon, Santosh Divvala, Ross Girshick and Ali Farhadi. You Only Look Once: Unified, Real-Time Object Detection. In arXiv:1506.02640v5 [cs.CV] 9 May 2016
- [2] Chufeng Tang, Lu Sheng, Zhaoxiang Zhang and Xiaolin Hu. Improving Pedestrian Attribute Recognition With Weakly-Supervised Multi-Scale Attribute-Specific Localization. In arXiv:1910.04562v1 [cs.CV] 10 Oct 2019
- [3] Deng, Y. Luo, P. Loy, C. C. & Tang, X. Pedestrian attribute recognition at far distance. In Proceedings of ACM Multimedia (ACM MM), 2014