

深層学習を応用した都市空間での人流解析

よしだりゅうと たかもり まき こ おおくぼじゅんいち ふじいじゅんいちろう べつぷともや
吉田龍人¹・高森真紀子¹・大久保 順一¹・藤井純一郎¹・別府知哉¹

¹八千代エンジニアリング(株) (〒111-8648東京都台東区浅草橋5-20-8)

「人中心のストリートへの転換」が新たな都市政策の重要課題として掲げられ、ストリーットの質を評価するうえで人の行動把握の重要性が示されている。行動把握の方法は種々あるが、ビデオカメラによる映像取得と分析は、簡便性等の観点からも適用性が高いと考えられる。ただし、人手による映像分析とデータ整理は膨大な作業時間を要することが課題であった。この課題を解決する足掛かりとして、本研究では動画から人流に関する基礎的な情報（方向別通行人数、通行時間、軌跡）をAIによって解析する手法を開発し、精度を評価した。手法には改善点も確認されたが、実適用に向けた一定の成果が得られた。

Key Words : 人流解析, 深層学習, AI, Re-Identification, 公共空間利活用

1. はじめに

まちなかの街路などの公共空間を人中心の空間へと転換するまちづくりが注目されている。2020年には国土交通省から「ストリートデザインガイドラインver1.0」¹⁾が公開されるなど、様々な取組が展開されている。人中心のまちづくりに向けた公共空間利活用に関する社会実験や実践を進める中で、空間の変化が人々の行動に与える影響の把握が望まれるが、行動把握のための調査は人手に依存しがちな状況がある²⁾。人手による調査では、現地の直接的な観測や、道路上に設置したビデオカメラ映像の目視判別など様々な方法が取られるが、いずれも人がチェックする作業を要するため、多大なコストや人的資源を要する。同時に、コストの制約によって調査地点数や調査日数などが限定されるといった課題がある。

これに対して、近年の技術発展やサービスの多様化に伴って調査を自動化する様々な手法検討の動きがある³⁾。代表的なものに携帯電話の位置情報に基づく推計手法が挙げられる。この手法は自動的に大規模なデータが取得されるといったメリットがある反面、利用できるデータの空間メッシュが粗く、歩道空間の調査には向かない。また、センサーによる計測手法も一部地域で実用化が進んでいる⁴⁾が、観測に専用機器が必要であり、性別などの属性情報の取得が困難といったデメリットがある。これらの手法に対して、ビデオカメラ映像解析による手法は、一般に普及しているカメラの映像でも適用可能であり、様々な属性情報の取得可能性にも期待がある。

以上に基づき、本研究の目的を都市空間のビデオデータを対象とする行動把握のための解析技術を構築することとした。特にまちづくりに関する施策の

実施主体となる自治体などが利用することを考慮して、①一般的な画質（1920×1080px）のビデオデータを用いることと、②一般に普及している非GPU搭載PCによって解析を実施できることといった条件を満たす手法の開発を目指した。

本論文では、目的の達成に向けた足掛かりとして新たに開発した方向別通行人数、通行時間、軌跡が取得可能な人流解析手法を提案する。さらにその手法で集計した人流観測データと人手で観測したデータに基づく精度検証を行い、得られた結果から手法の有用性を検討した結果を示す。

2. 提案手法

本研究で提案する人流解析手法は、人を検出する「物体検出」および物体検出結果の軌跡を取得する「物体追跡」で構成される。入力データは動画のみであり、解析によって動画のフレームに紐づいた人物ごとの軌跡データが出力される。なお出力に関して個人を特定する情報は含まれない。

「物体検出」では動画のフレームごとに通行者の検出を行う。物体検出には深層学習モデルのSSD⁵⁾を用いた。SSDは1度の推論で画像全体の物体検出ができるという点から、処理速度に強みがある。本研究ではさらなる処理速度向上を目指してSSDのベースネットにMobileNet⁶⁾を用いた。SSDで画像を推論すると、検出対象物に対してBoundingBox（以後、BBoxと呼ぶ）と呼ばれる矩形が付与される。

「物体追跡」では各フレームで出力されたBBoxの座標を基に、異なるフレームに写った同一人物の軌跡を追跡する。本研究では、物体追跡に深層距離学習モデルを用いたRe-Identification⁷⁾を採用した。Re-Identificationとは、直近のフレームで検出さ

れたBBox内部の特徴量の類似度を算出することで同一物体の軌跡追跡を可能とする手法である。Re-Identificationは、各人物の見た目によって同一性を判定することから、移動方向が複雑な場合の物体追跡においても頑健性の高い追跡が可能となる。

本手法では「物体検出」および「物体追跡」の各工程で2つの異なる深層学習モデルを使用するにあたり、開発コスト削減の観点から、既存の事前学習済公開モデルを適用することとした。さらに、一般的なPCで解析することを念頭に、CPUに最適化したモデルを用いてシステムの構築を行った。

3. 実証実験

(1) 実験内容

実験は三鷹駅北口において撮影した60FPSの動画によって行う。図-1に実際に撮影した動画の1コマを示す。本動画は社会実験の様子を撮影したものである。その社会実験は歩道上に設置したカフェセットの効果を検証するものであり、その一環としてカフェセット設置前後における通行状況調査を行った。調査では、調査区間内を25メッシュ（図-1）に区分し、表-1の内容について観測している。本研究ではカフェセット設置時を社会実験期間内と称する。

本章では人手による動画の解析結果と2章で示した提案手法（以後、AI解析手法と呼ぶ）による解析結果を比較する。比較する項目は通行人数、通行時間、通過メッシュの適合性である。

AI解析は動画の全フレームに実施されるため、得られる軌跡は人手による作業では現実的に取得が困難なほど細かな粒度のデータとなり、挙動を敏感に把握する点が特徴の一つとなる。瞬間的なメッシュ間移動もすべて把握されるが、人手の解析における25メッシュの調査結果との比較を行うため、通過位置の判別を人手解析と同程度に丸める処理を行った。具体的には通過したメッシュを判定する頻度を減じることとした。判定頻度は、10・15・20・30フレームごとの4条件で実験をし、人手による調査結果に最も近い出力を得た20フレームを採択した。



図-1 調査箇所の状況及びメッシュ設定

表-1 調査概要

撮影日時	2020.10.14 17:00~17:10 (社会実験期間外) 2020.10.28 17:00~17:10 (社会実験期間内)
観測内容	・ 歩行者、自転車の方向別通行者数 ・ 通行者別通行位置 (25メッシュ) ・ 調査区間の入出時間 (in, out)
解析時間	10分の動画で約120分 (約1分/人)

(2) 実験結果

本節では10分の動画を解析し、データを取りまとめた結果を示す。

AI解析によって得られた軌跡を画像に描画した結果を図-2に示す。10分の動画に対するAI解析時間は約30分であり、人手での作業に比べて1/4となった。特に、AI解析の実施にかかる人の作業時間はプログラムの実施にかかる数分のみまで短縮された。

通行人数の集計結果を表-2に示す。なお、今回用いたモデルでは、自転車運転者も人として検出されるため、歩行者と自転車を合わせた通行者として集計している。以後に真値として記載するものはいずれも人手で計測した結果である。

北向き・南向き・方向不明者の総和を取った総通行人数（総ID数）は、実験期間外において103%となり、高い精度となった。他方、期間内の精度は低下（115%）しており、主因は、歩道に設置したコーンの誤検出にある。期間内外の両ケースにおいて、計測値が真値より多くなる現象が観測されており、「同一人物に異なるIDを付与」「誤検出により人以外の物体にIDを付与」が主な原因となる。なお、この結果は「検出漏れによる下振れ」と「誤検出による上振れ」が相殺されている点に注意が必要である。



(a) 社会実験期間外



(b) 社会実験期間内

図-2 歩行軌跡の可視化結果

表-2 通行人数の比較

項目	真値(人)		AI解析(人)		精度(%)	
	期間外	期間内	期間外	期間内	期間外	期間内
北向き (奥方向)	65	67	61	58	94	87
南向き (手前方向)	60	66	48	54	80	82
方向不明	—		20	41	—	
総和	125	133	129	153	103	115
方向別通行人数の平均精度					86	

方向別通行人数のAI解析による集計結果は真値に比べて平均86%となり、少なく計上された。この原因は以下の2点と推察する。1つはプログラム上で1~5もしくは21~25番のメッシュで人が検出できなかった場合、方向不明とするアルゴリズムになっていたためである。特に、1~5番や25番のメッシュのように人の特徴量が取得しにくい箇所での検出漏れが影響した可能性が高い。もう1つの原因はRe-Identificationのみを用いた物体追跡によって、服装など見た目の類似した人物に同一IDが割り振られてしまったためである。フレーム内に同時に存在する2人に同一IDが振られることは仕組み上ないが、ある人が対象区間から出たタイミングで似た見た目の人（同一人物と判定）が対象区間に入った場合に同一IDが割り振られる場合があり、方向別人数が1人少なく計上されていた。

通行時間の集計結果を表-3に示す。AI解析の通行時間計測精度は真値に比べて期間外で77%、期間内で79%と低い。この結果は方向不明者など、軌跡を十分に取得できていない人も含めて平均を算出した影響と考えられる。今後、AI解析結果から通行時間の分布を把握する際には、調査結果の全数を評価するのではなく、十分に軌跡を取得できたデータだけを評価対象とし、母集団の分布を推計する方が実態に即した分析になる可能性が示唆されている。

表-3 通行時間の比較

項目	真値(秒)		AI解析(秒)		精度(%)	
	期間外	期間内	期間外	期間内	期間外	期間内
平均通行時間	13	12	10	9	77	79
歩行者平均	13	12	—			
自転車平均	8	11	—			

表-2に示した全ての通行人数のうち、各メッシュを通過した人数の割合を評価した結果を表-4に示す。この結果より南向きの通行者は車道側、北向きの通行者は店舗側を通る実際の傾向を概ね捉えており、AI解析の一定の妥当性を確認できている。

なお、AI解析結果と真値の差分に着目すると、人通りの多い縦ライン（北向きでは店舗側、南向きでは車道側）および画面奥で符号が負になる傾向にあった。人通りの多いラインでは人同士の重なりによるAIの検出漏れ、画面奥では対象物の小ささに起因する検出漏れが確認され、これらの現象が負の値となる要因であったと分析する。最大誤差である社会実験期間内の14番のメッシュにおける41%の誤差は、コーンの誤検出によるものであった。図-2に示されるコーン周辺の軌跡の乱れもその影響である。

によるAIの検出漏れ、画面奥では対象物の小ささに起因する検出漏れが確認され、これらの現象が負の値となる要因であったと分析する。最大誤差である社会実験期間内の14番のメッシュにおける41%の誤差は、コーンの誤検出によるものであった。図-2に示されるコーン周辺の軌跡の乱れもその影響である。

(3) 精度向上に向けた今後の課題

前節では人手での解析との比較を行い、AI解析手法の精度が低下する要因を整理した。この結果を踏まえて、本手法の実適用に向けた精度向上の取組として、以下の方策が有用と考える。

a) 動画の撮影角度の改善

画像認識タスクにおいて、手前の物体が奥の物体を隠すオクルージョンと呼ばれる現象が、奥の物体の認識精度を下げることは通説となっている。本研究の動画においては、人混みおよびカフェセットによるオクルージョンが確認された。オクルージョンは、物体検出の精度を低下させるだけでなく、Re-Identificationによる物体追跡の精度も低下させる。例えばオクルージョン発生時にかろうじて物体検出ができたとしても、類似度を算出する段階で手前に存在する物体の特徴量が悪影響を及ぼす。カメラの設置位置や撮影角度は現地の状況にも依存するが、できるかぎりオクルージョンの影響を低減できるような配慮・工夫を行うことが望ましい。

b) 誤検出低減型物体検出モデルの構築

コーンの誤検出が人流の評価精度を低下させる様子が確認された。一般にAIのモデルは学習時の目的関数によって、誤検出を許容して検出漏れを低減するか、検出漏れを許容して誤検出を低減するかといった方針を定めて作成する。今回使用したモデルは前者の方針であることが、コーンの誤検出と精度低下に繋がった可能性がある。今後、後者の方針で学習した誤検出の少ないモデルを開発し、精度について比較することが有用と考える。なお、学習データは公開されているため、開発コストの縮減は十分可能と考える。

c) 物体追跡アルゴリズムの改善

Re-Identificationによる見た目の一致度に依存した物体追跡の結果として、画面奥へと通り抜けた人物が突如画面手前に出現するといった現実的

通行位置番号

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

表-4 通過メッシュ評価結果

	(a) 社会実験期間外					(b) 社会実験期間内														
	AI解析		AI解析-真値			AI解析		AI解析-真値												
北向き	北向き-AI		σ	平均		北向き誤差		σ	平均											
			0.15	23%				0.10	0%											
	31%	18%	11%	8%	3%	-20%	-20%	-1%	-1%	-1%	40%	17%	9%	5%	2%	-33%	-2%	0%	-2%	-1%
	46%	25%	11%	11%	2%	-5%	-11%	-5%	4%	0%	40%	17%	9%	10%	0%	-33%	1%	-2%	7%	0%
	51%	56%	31%	28%	10%	-5%	17%	16%	16%	8%	72%	57%	21%	31%	12%	-8%	32%	13%	24%	11%
	30%	43%	20%	16%	8%	-9%	0%	7%	7%	7%	62%	36%	17%	24%	5%	-8%	21%	13%	15%	4%
30%	46%	16%	21%	5%	-18%	12%	-2%	12%	2%	59%	34%	10%	22%	2%	-3%	12%	4%	16%	2%	
南向き	南向き-AI		σ	平均		南向き誤差		σ	平均											
			0.19	25%				0.09	2%											
	17%	10%	13%	50%	33%	13%	-5%	-9%	-12%	-2%	33%	19%	19%	9%	7%	27%	2%	-18%	-35%	-18%
	15%	6%	8%	52%	19%	5%	0%	-13%	4%	-3%	20%	13%	24%	26%	4%	14%	1%	-23%	-9%	4%
	13%	13%	17%	79%	33%	4%	4%	-5%	24%	10%	20%	35%	50%	81%	9%	10%	19%	3%	41%	9%
	6%	10%	17%	46%	33%	5%	-3%	-7%	3%	8%	15%	13%	17%	70%	2%	9%	-1%	-14%	23%	2%
10%	19%	6%	60%	27%	0%	5%	-7%	17%	9%	13%	9%	11%	69%	0%	2%	6%	-16%	26%	0%	

り得ない軌跡が数件確認された。よってBBox間のユークリッド距離を考慮したcentroid truckingと呼ばれる手法も組み合わせることで、上記の現象の発生を防止し、精度向上に繋がるものと考え。

d) 方向判定方法の見直し

対象区間の端部（本研究における1～5および21～25番のメッシュ）を通過していない軌跡を方向不明とする今回のアルゴリズムでは、何らかの要因による検出漏れによって方向不明者が増加することが確認された。これに対して出現時と消失時の座標を結んだベクトルのなす角から移動方向を定義すると、方向別通行人数の精度向上に繋がるものと考え。

e) 速度評価の実現

通行時間を同条件で計測するためには、対象区間を出入りする瞬間を捉える必要があり、検出漏れの影響を受けやすい。通行時間を代替・補完する指標として速度を用いることで、短い軌跡でも計測可能となり、通行状況の実態に即した評価手法として、実適用の可能性向上が期待できる。速度の計測において必要となる画像上のBBoxの移動距離を実空間上の移動距離と対応づける処理は、画像の射影変換によって実施可能である。

4. おわりに

(1) まとめ

本研究では、人の行動把握ツールを構築することを目指した第一歩として、AIを用いた人流解析手法を開発した。本手法は一般的なPCでも解析を実行できるものとし、解析に必要となる時間の大幅な削減を達成した。また、人手によるビデオ解析結果に基づく精度検証によって、概ね人流に関する傾向を表したデータが取得できていることを確認した。

精度向上が今後の課題として残るが、理論的に説明性のある改善策も検討・整理しており、今後の実適用に向けた一定の成果が得られた。

(2) AI解析手法の応用可能性

AIによる人流解析手法は現地の状況や調査の目的によって柔軟に解析方法を変更することができ、様々な都市空間分析への応用可能性が期待される。

今後は、3章で示した精度向上の取組に加え、①都市政策の課題を検討する際に活用しやすい取得データの集計・表現の工夫や、②人流に関する属性情報の付与に取り組むたい。

①の試行例を図-3に示す。上段の歩道を模した挿絵に描いた軌跡は、画像の射影変換技術を活用して図-2に示す軌跡を俯瞰図として表したものである。

射影変換では実空間上のサイズに基づいた変換を実施することで、画像の各ピクセルの持つ空間解像度を統一することができる。これにより、軌跡データを地図情報に投影することが可能となる。さらに画像の空間解像度が統一されると、各人物の軌跡から移動距離が、移動距離と移動時間から速度が算出可能となる。デジタルデータは、集計や表現など、加工のしやすさが特徴の一つと認識しており、その

工夫・検討を通じて、都市空間の課題・特徴の抽出や施策検討に寄与したいと考える。

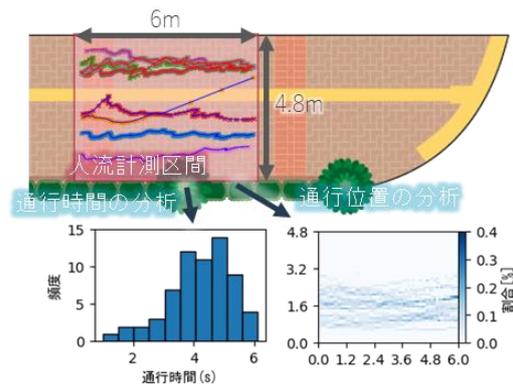


図-3 人流の俯瞰図および分析例

②に関しては、今後のストリートの質評価に繋げる観点から、質的データへの展開として、軌跡に加えて属性情報を付与した分析を可能とする技術開発を進めたい。直近の取組としては自転車検出モデルを作成し、自転車と歩行者を区別した分析を目指す。その他、年齢、性別、グループ人数など順次分析可能な属性項目を追加検討する予定である。また、立っている・座っているなど、人の状態を評価する技術の検討も進め、都市空間評価のために人の行動を詳細に把握できる手法の構築を実現したい。

謝辞：研究に掲載した人手解析調査は、著者らが武蔵野市都市整備部まちづくり推進課より業務委託を受けて2020年に実施したものです。AI解析の実施にあたり、ビデオ調査データ及び人手解析結果の利用について快諾いただき、本研究を進めることができました。ここに謝意を表します。

参考文献

- 1) 国土交通省：ストリートデザインガイドラインver2.0, 2021.
- 2) 柴田久：大分 昭和通り・交差点四隅広場のリノベーションプロジェクトレポート, 2018, http://www.tc.fukuoka-u.ac.jp/tc/labo/keikan/works/project/pj%20report/reborn_pj%20report.pdf (閲覧日:2022/07/28).
- 3) 国土交通省：地域課題解決のための人流データ利活用の手引きVer1.0, 2022.
- 4) 国土交通省：レーザーセンサーによる高精度でリアルタイムな人流計測, <https://www.mlit.go.jp/plateau/use-case/activity-monitoring/1-004/> (閲覧日:2022/07/28).
- 5) Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., Berg, A. C.: SSD:Single Shot MultiBox Detector, arXiv:1515.02325, 2016.
- 6) A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam.: MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications, arXiv:1704.04861, 2017.
- 7) Wang, H., Du, H., Zhao, Y., Yan, J. : A Comprehensive Overview of Person Re-Identification Approaches., IEEE Access, Vol.8, pp. 45556 - 45583, 2020.