

転移学習と Faster R-CNN を用いた熱赤外画像を対象とする 車両および人物動作検出手法の改善

劉 亜儒 松井 解 景山 陽一 白井 光 石沢 千佳子

秋田大学大学院理工学研究科

1. 背景・目的

現在、日本は高齢化率が最も高い国である。そのため、高齢者の人口が増加しており、全年齢層における交通事故の死者数は減少傾向にある一方で、65歳以上の高齢者が占める割合は依然として高い。また、交通事故における昼夜別の死傷者数に着目すると、夜間の死者数の割合は圧倒的に高く、昼間の死傷者数の約2.4倍である[1]。したがって、運転手、歩行者などの道路利用者の安全、安心を支援する技術を開発することは、持続可能な地域社会を構築する上で不可欠である。

本研究グループはこれまでに、夜間屋外環境下における歩行者の行動パターンを解析するため、熱赤外カメラを用いたロバスト性の高い夜間人物検出法を検討した[2]。具体的には、外気温や照度の異なる熱赤外画像データを対象とし、畳み込みニューラルネットワークを用いて、夜間における人物の動作解析を行った[3]。さらに、車両走行時の道路利用者をリアルタイムに支援するため、熱赤外画像データを用いて夜間における車両走行のパターンを取得し、Faster R-CNN[4]を用いたリアルタイムの人物動作検出および車両検出手法に関して検討を行った[5]。しかしながら、これまでの検討では、熱赤外画像データ内における歩行者の人数が一人の場合のみを対象としており、複数人を対象とした検討を行うまでには至っていない。

そこで本研究では、熱赤外カメラを用いて複数人が動作を行う様子を新たに取得し、夜間における車両検出および複数人を対象とする人物の動作認識を行った。

2. 使用データ

2.1. 熱赤外線カメラ

夜間屋外環境下は、日中とは異なり、車両のライトや街灯の有無などに伴って、照明条件が大きく変化する。そこで本研究では、照明条件に影響を受けない熱赤外画像データを活用する。夜間における歩行者および車両を対象とし、熱赤外カメラ(株式会社 D-eyes 製、「超高感度+遠赤外線」2in1 カメラ WCAM001-AU、640×480 画素)を用いて動画像データを取得した。

2.2. 学習済みモデル用データ

歩行者が行う動作を想定し、被験者がカメラから約 10m 離れた場所で“しゃがみ、屈み、直立、な

らびに歩行”の4種類の動作を行う様子を一人ずつ撮影した。なお、歩行者に加えて、走行する車両を同時に撮影した。被験者は4人である(20代男性2名および女性2名、イーストアジア人)。

2.3. 追加データ

複数人を対象とした検討を行うため、被験者2名と被験者3名がカメラから約10m~20m離れた距離で直立、歩行および車両の動作パターンを追加データとして取得した。取得した追加データの例を図1に示す。なお、本研究で用いたデータは「秋田大学手形地区における人を対象とした研究に関する倫理規程第6条第2項」に基づき、被験者の同意を得て取得した。

3. 提案手法

3.1. 提案手法の流れ

提案手法の流れを図2に示す。本稿で使用した熱赤外カメラは、撮影画像に温度バーが表示される。はじめに、前処理として撮影した画像データに含まれる温度バー領域を除去した。次に、LabelImg ツールを用いて人物動作パターンおよび車両のラベル付けを行い、アノテーションデータセットを設定した。さらに、構造のバックボーンを ResNet50[6]とする Faster R-CNN モデル(学習済みモデル)を作成し、学習済みモデル用データを学習した。加えて、学習済みモデルを転移学習した転移学習モデルを作成する。転移学習モデルは、学習済みモデルの重みを固定し、特徴抽出器として利用する。最後に、Anchor Box サイズを調整し、追加データセットを用いて転移学習モデルを学習した。本研究では、得られたモデルに対してテストデータを用いた検証を実施し、車両検出および人物の動作(直立、歩行)認識の精度に関する考察を行った。

3.2. データセットの処理

学習済みモデル用画像データは、LabelImg を用いて、6079 枚の画像に対して手動で検出目標をラ



図1 データ取得例(被験者3人)

Improvement of Vehicle and Human Action Detection in Infrared Images Based on Transition Learning and Faster R-CNN

Yaru Liu, Kai Matsui, Yoichi Kageyama, Hikaru Shirai, Chikako Ishizawa

Graduate School of Engineering Science, Akita University

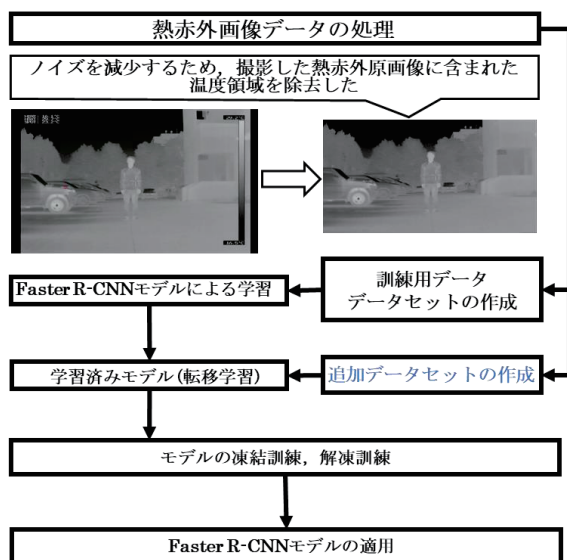


図 2 提案手法の流れ

ベリリングした。その中から、アノテーションデータセットを作成し、訓練データ 4257 枚、検証データ 1822 枚に分割した。また、追加データは 1651 枚の画像であり、訓練データ 1320 枚、検証データ 331 枚に分割した。

3.3. 転移学習

学習済みモデルの性能を利用するため、転移学習を施した。追加データは、再学習せずに学習済みモデルを利用した特徴抽出器を用いて、特徴ベクトルに変換した。

3.4. Faster R-CNN モデルの改善

入力画像から特徴マップを出力するため、特徴抽出器はネットワーク ResNet50 をモデルのバックボーンとして流用した。また、新たな検出対象に対応するため、Anchor Box のサイズはそれぞれ 64×64 , 256×256 , ならびに 512×512 を検討した。なお、学習済みモデルの重みを保護するため、学習済みモデルを凍結し、エポック数を 50 回に設定して凍結訓練および解凍訓練[7]を行った。

4. 評価実験

訓練した Faster R-CNN モデルを評価するため、追加データから 413 枚の画像をテストデータに設定した。車両および 2 種類の人物動作(直立, 歩行)の 3 つのパターンを検証し、各パターンの精度 (Precision), 再現率(Recall), 平均精度(AP), ならびに評価指標(mAP)を算出した。ここで、比較手法は改善前の手法とした[5]。

5. 結果および考察

評価結果を表 1, 表 2 にそれぞれ示す。提案手法は、各パターンにおいて、改善前よりも高い精度を示した。平均精度の向上は、直立で約 46.5%, 歩行で 63.9%, 車両で 51.9%である。また、最終的な評価指標である mAP は、提案手法で 96.5%を示し、改善前の手法における検出結果より高い精度を得た。

表 1 テストデータの検証結果(改善前)

	Precision	Recall	AP
直立	36.1%	38.2%	18.2%
歩行	27.7%	15.4%	7.1%
車両	43.7%	69.1%	51.4%
mAP			25.6%

表 2 テストデータの検証結果(改善後)

	Precision	Recall	AP
直立	82.6%	91.0%	90.4%
歩行	91.6%	98.6%	99.1%
車両	95.6%	100%	99.9%
mAP			96.5%

以上の結果は、提案手法は転移学習と Faster R-CNN モデルを用いた、夜間における車両および複数人の人物動作検出を行う上で、検出精度の向上に寄与することを示唆している。

6. まとめ

本稿では、提案手法を用いた夜間の人物動作検出および車両検出に関する検討を行った。その結果、提案手法は複数人の動作を検出可能であり、従来手法と比較して、検出精度向上に寄与することを明らかにした。今後は、カメラと撮影対象の距離を考慮した手法に関して検討を加える予定である。

7. 謝辞

本研究は、スズキ財団科学技術研究助成を受けて行われた。

参考文献

- [1] 警察庁：令和 3 年における交通事故の発生状況等について、<https://www.npa.go.jp/bureau/traffic/bunseki/nenkan/040303R03nenkan.pdf> (Accessed:2022/12/23)
- [2] Y. Liu, K. Matsui, Y. Kageyama, H. Shirai, C. Ishizawa: Analysis of Human Action in Infrared Images at Nighttime Using CNN, The Ninth International Conference on Materials Engineering for Resources, ICMR2021 AKITA, B-8 (2021)
- [3] Y. Liu, K. Matsui, Y. Kageyama, H. Shirai, C. Ishizawa: man Action Analysis at Nighttime under Different Temperature Conditions in Infrared Images Using CNN, the 10th IIAE International Conference on Industrial Application Engineering, OS2-4 (2022)
- [4] S. Ren, K.He, R. Girshick, J. Sun: Faster r-cnn: Towards real-time object detection with region proposal networks, Advances in neural information processing systems 28 (2015)
- [5] 劉亞儒, 松井解, 景山陽一, 白井光, 石沢千佳子: Faster R-CNN モデルを用いた熱赤外面像を対象とする夜間の人物動作認識および車両検出, 映像情報メディア学会 2022 年冬季大会, 13B03 (2022)
- [6] K He, X Zhang, S Ren, J Sun: Deep residual learning for image recognition, Proceedings of the IEEE conference on computer vision and pattern recognition, pp.770-778 (2016)
- [7] R. Girshick: Fast r-cnn, Proceedings of the IEEE international conference on computer vision, pp.1400-1448 (2015)