

# 二輪車および歩行者に対応した交通計測 AI の開発および地点別自己学習手法の提案

尾原 健太郎<sup>1</sup>・柳沼 秀樹<sup>2</sup>・寺部 慎太郎<sup>3</sup>・海野 遥香<sup>4</sup>・鈴木 雄<sup>5</sup>

<sup>1</sup>学生非会員 東京理科大学院 創域理工学研究科社会基盤工学専攻 (〒 278-8510 千葉県野田市山崎 2641)

E-mail: 7623507@ed.tus.ac.jp

<sup>2</sup>正会員 東京理科大学准教授 創域理工学部社会基盤工学科 (〒 278-8510 千葉県野田市山崎 2641)

E-mail: yaginuma@rs.tus.ac.jp (Corresponding Author)

<sup>3</sup>正会員 東京理科大学教授 創域理工学部社会基盤工学科 (〒 278-8510 千葉県野田市山崎 2641)

E-mail: terabe@rs.tus.ac.jp

<sup>4</sup>正会員 東京理科大学助教 創域理工学部社会基盤工学科 (〒 278-8510 千葉県野田市山崎 2641)

E-mail: unoharuka@rs.tus.ac.jp

<sup>5</sup>正会員 東京理科大学助教 創域理工学部社会基盤工学科 (〒 278-8510 千葉県野田市山崎 2641)

E-mail: yusuzuki@rs.tus.ac.jp

近年、自動車交通量計測では AI による画像解析手法などが取り入れられ自動観測の体制が整いつつある。一方で、二輪車および歩行者に関しては計測精度が低いことが未だ課題であるが、実務においては交通量のみならず移動速度や移動軌跡などが取得可能な交通計測 AI の開発が期待されている。本研究では、二輪車および歩行者に対応した交通量と移動速度が推定可能な交通計測 AI の開発を目的とする。具体的には、新たな学習ラベルを用いて畳み込みニューラルネットワーク (CNN) を学習し、物体追跡により取得した情報を活用することで交通量計測と速度推定を可能とした。また、地点別の自己学習手法の提案についても取り組み、地点毎に生じる追加学習データのコスト低減と精度向上を試みた。その結果観測地点の背景画像を使用した合成画像による自己学習の有効性が示された。

**Key Words:** *traffic volume measurement, vehicle speed, Convolutional Neural Network, domain adaptation, self learning*

## 1. はじめに

国土交通省が実施している全国道路・街路交通情勢調査(道路交通センサス)は、交通状況の現状把握や道路の新たな整備計画を策定する上で必要不可欠な基礎資料である。しかしながら、平成 27 年度調査においても、交通量調査員主体による人手観測が主な計測方法であり、時間的・人的コストが高い現状にある。また、二輪車や歩行者を対象とする全国規模での調査が行われておらず、昨今の交通結節点整備や道路空間再配分などの新たな施策を検討する基礎的なデータが不足していることが課題である。

そこで、国土交通省は、従来の 5 年に 1 度の調査体形から常時観測体制への移行のために、常設トラフィックカウンターに加えて道路管理用の CCTV カメラの映像を AI 解析することで、常時観測区間を拡大することを目指している。バイクや自転車などの二輪車や歩行者は磁気に反応しないことから既設のトラフィックカウンターを使用した交通量の観測は不可能である。そのため、CCTV や防犯カメラ映像の AI 解析を実施し、二輪車・歩行者の交通量や速度の算出が可能になれば、全

国規模で詳細なデータを得ることができると期待されている。また、CCTV 画像に AI を用いた解析を行うことで、自転車・二輪車・歩行者の交通量や速度だけでなく、移動軌跡の取得も可能になる。横断歩道上の自転車や歩行者の移動軌跡を取得し、自転車の有無で歩行者の移動軌跡が変化するか否かなどを比較することができれば、交通安全対策にも応用可能である。

しかし、現状の AI を用いた交通量の解析精度は国土交通省の調査<sup>1)</sup>によると、自動車と比較し、二輪車・歩行者においてはほとんど検知できていない点が現状の AI 解析の課題といえる。そこで、本研究の目的は二輪車・歩行者の交通量を導出するアルゴリズムを開発することを基本とし、移動軌跡を活用した新たな速度推定手法についても提案する。

## 2. 既往研究と本研究の位置付け

道路管理用カメラである CCTV の活用方法についてはこれまでも多くの検討がされてきており、近年では機械学習や深層学習などを用いた手法の導入の検討が進んでいる。

CCTV により取得した映像データの日常的利用方法について、片岡ら<sup>2)</sup>は既存施設である CCTV 映像を用いた交通流計測を行なっている。背景差分を基本にした画像処理アルゴリズムを用いて 6 通りの車類分類を行い、通過した車種分布のマッチングを行うことで旅行時間が計測可能であることを示すと共に、更なる精度向上のためには車種区分を増加させるという改善策が挙げられている。Caballo ら<sup>3)</sup>は、フィリピンではメジャーな交通機関であるトライシクルの映像処理モデルの検討を行った。衛生画像や航空画像における、自動車検知などに一定の実績があった、畳み込みニューラルネットワーク (以下 CNN) を用いた、深層学習ベースの物体検出モデルである YOLO を使用し、精度検証を行った。214 枚の学習画像を作成し学習させることで、物体検出タスクの評価基準である mAP が、0.3791 を達成したことが示された。より正確なモデルの実現のためには、トレーニングデータセットを増やすことに加えて、データ収集地点のバリエーションを増やすことが必要であると指摘されている。また、物体検出モデル同士の処理速度の比較を、Jia ら<sup>4)</sup>は行い、モデルの精度の高さは mAP、処理速度は 1 秒間処理できる画像の枚数を表す、フレームスピード (FPS) 指標を用いられている。主要なモデルの比較を行った結果、YOLOv5 というモデルについて、精度と処理速度共に高いことが示された。

映像情報を使用した車速推定アルゴリズムにおいてもいくつかの手法が存在する。Kurniawan ら<sup>5)</sup>は、CCTV 映像に背景差分処理を施して物体検出を行った後に、射影変換を行い距離情報を取得し車速の推定を行った。射影変換とは画像の視点を変化させるアルゴリズムである。また、Wang ら<sup>6)</sup>物体検出に深層学習モデルである YOLOv3 を使用し、物体追跡においては状態空間モデルの一つであるカルマンフィルタを適用し、車速推定手法の検討を行った。

以上より、本研究では精度と処理速度が共に高い YOLOv5 というモデルを基本とした交通計測が可能なアルゴリズムを構築し、実用面を考慮した新たな学習手法を提案する。加えて、カルマンフィルタと深層学習を組み合わせた物体追跡アルゴリズムを使用し、速度推定手法の開発を行う。

道路の常時観測体制の実現のために、映像から物体を検出して追跡、そして追跡結果を分析するといった一連の流れを提案することが本研究の意義である。

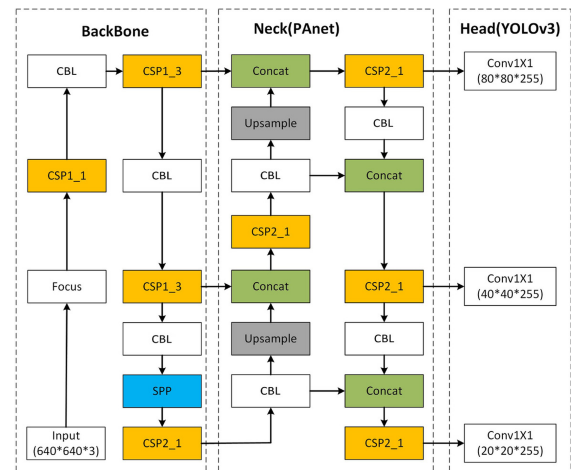


図-1 YOLOv5 のアーキテクチャ<sup>4)</sup>

### 3. 本研究で使用するモデルの概要

#### (1) CNN

##### a) CNN の基本的構造

畳み込みニューラルネットワーク (convolutional neural network: CNN) とは、画像処理の分野で使用されることが多い手法であり、ニューラルネットワークを応用したモデルである。画像データは、縦横の 2 次元の数値に加えて、色情報が格納された、3 次元データとなる。そのため、通常のニューラルネットワークを使用することでも画像の処理は可能だが、縦横に並んでいる画像データを分解して、縦一列の情報に変換する必要がある。画像データは、周囲のデータの位置関係が重要な意味を持つため、画像データに適したニューラルネットワーク構造である CNN が考案された。CNN が通常のニューラルネットワークと異なる点は、畳み込み層、プーリング層、全結合層の実装のされ方である。これらの実装によって、画像内の特徴を抽出し、画像内物体の位置変化の影響を受けづらい物体検出が可能になる。

##### b) YOLOv5 モデルの概要

YOLOv5<sup>7)</sup>は、2020年に発表された5番目のYOLOモデルのバージョンである。YOLOv5のフレームワークは、図-1でも示しているように、特徴を抽出するBackBone、情報を伝えるNeck、情報を出力するHeadの主に3つに分かれた構造となっている。YOLOv5の特徴は、モデルサイズを柔軟に変更可能であり、物体追跡手法などの他モデルとの拡張性が高い点である。このような観点より、本研究ではYOLOv5の軽量モデルであるYOLOv5sを用いる。

##### c) mAP

mAP(mean Average Precisions)とは物体検出モデルの正確さを評価するために使用される指標である。物体検出では、物体の位置を矩形領域 (Bounding Box) で

表現するため、Bounding Box が正しく予測できているか、判断する指標が必要である。Bounding Box の予測の判定には、IOU(Intersection Over Union) という指標が使用され、IOU は予測した Bounding Box と正解の Bounding Box の重なっている領域の比率を示す。本研究では、IOU が 0.5 より大きい時に正解とみなすこととする。

Precision (適合率) を縦軸, Recall (再現率) を横軸にとり、プロットしたものが PR 曲線 (Precision Recall Curve) である。PR 曲線が右上に近いほど、検出すべきボックスが検出され、その上ボックスが正しいことを示しているため優秀なモデルとすることができる。PR 曲線の下側の面積を求めたものが、AP(Average Precision) であり、mAP(mean Average Precision) は全てのクラスの AP の値の平均を取ったものである。

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

表-1 TP,FP,FN の定義

	正解	不正解
検出	TP(True Positive)	FP(False Positive)
検出されない	定義しない	FN(False Negative)

## (2) StrongSORT

StrongSORT<sup>8)</sup> は複数の特定の物体を追跡する手法である。カルマンフィルタと呼ばれる時系列状態空間モデルを適用することで、位置情報を基に物体同士を紐づける。加えて、CNN を組み込み視覚的な特徴量の抽出も行うことで、位置情報と視覚情報という複数の情報を用いて物体の追跡を行う点が StrongSORT の特徴である。StrongSORT には外特徴抽出器に ResNest50 をバックボーンとするモデルである BoT が導入されているため、個体同士の特徴抽出能力が向上している。加えて、カルマンフィルタについても物体の検出信頼度に応じて観測ノイズを変化させる役割を果たす、NSA カルマンが導入されていることも特徴である。

## 4. 学習ラベルの拡張と転移学習結果

### (1) 学習ラベルの拡張

物体検出タスクにおいて一般的に用いられているクラス分類ラベルは、COCO データセットを基準にしている。COCO データセットは 80 個のクラスから構成され、人間、自転車、バイクといったクラス分けが成さ

表-2 NSA カルマンフィルタの入出力とアルゴリズム

入力	
観測値	$z_k$
観測信頼値	$c_k$
予測状態推定値	$\hat{x}_{k k-1}$
観測モデル (観測係数ベクトル)	$H_k$
測定値のノイズの共分散	$R_k$
出力	
更新状態推定値	$\hat{x}_{k k}$
事後誤差共分散行列	$P_{k k}$
アルゴリズムの流れ	
1: 事前状態誤差の導出	$\hat{y}_k = z_k - H_k \hat{x}_{k k-1}$
2: NSA 共分散の導出	$\hat{R}_k = (1 - c_k) R_k$
3: 事前誤差共分散の導出	$S_k = H_k P_{k k-1} H_k^T + \hat{R}_k$
4: 最適カルマンゲインの使用	$K_k = P_{k k-1} H_k^T S_k^{-1}$
5: 状態推定値の更新	$\hat{x}_{k k} = \hat{x}_{k k-1} + K_k \hat{y}_k$
6: 事後誤差共分散行列の更新	$P_{k k} = (I - K_k H_k) P_{k k-1}$

れている。COCO データセットのラベルを使用して二輪車の検出を行う場合、人間の部分と二輪車の部分に分かれて検出されるため、二輪車や歩行者の交通量を正確に計測することに適していない。また、自転車やバイクなどの二輪車は、通常人が乗っていることが当然であるため、二輪車は人間と共にまとめてラベル付けを行うことが適切であると考えた。

そこで、本研究においては表-3 に示す通り、二輪車と歩行者に注目して 7 種類のラベルを用いて学習用データを作成する。

表-3 使用するラベルの種類

ラベル名	状態
person	歩行者
person-on-motorcycle	バイクに乗車している人
person-on-bicycle	自転車に乗車している人
person-by-bicycle	二輪車を押している人
bicycle	放置されている自転車
motorcycle	放置されているバイク
car	自動車

学習画像については、二輪車及び歩行者が多い地点を中心に 10 地点を選定し、1760 枚の画像のアノテーションを行った。各ラベルのインスタンス数については表-5 に示す。

### (2) 転移学習の実施

転移学習 (Transfer Learning) とは、学習済みのモデルを利用し、出力層のみ再学習させる方法である。転移学習はデータセットの数が十分に揃っていない場合において有効だという特徴が存在する。本研究では、

表-4 転移学習後の交通量計測結果

地点	時刻	歩行者		自転車		自動二輪	
		going	coming	going	coming	going	coming
地点 1	8:00-8:30	100%	87.1%	90.9%	78.3%	77.8%	69.1%
	12:00-12:30(学習画像含む)	95.0%	98.1%	89.7%	88.9%	-	-
	18:00-18:30(学習画像含む)	93.8%	94.6%	85.3%	92.9%	-	-
地点 2	8:00-8:30	91.3%	94.3%	22.2%	55.6%	85.8%	
	8:30-9:00	81.3%	98.1%	30.0%	50.0%	97.7%	
地点 3	12:00-12:30	46.2%	68.4%	95.4%	85.4%	-	-

表-5 アノテーションデータ内容

ラベル名	インスタンスの個数(個)
person	19847
person-on-motorcycle	2259
person-on-bicycle	1073
person-by-bicycle	204
bicycle	220
motorcycle	383
car	6947

YOLOv5s という学習済みモデルの backbone 層を固定して、boundingbox の出力とクラス分類に大きく関わる Neck 層と出力層の転移学習を行う。

学習データと検証データの割合を 8:2 で分割を行い、エポック数が 200、バッチ数 16 で backbone の 10 層を固定して学習を行った。学習結果は mAP=0.506 と示され、ラベル数が少ない bicycle, motorcycle, person-by-bicycle を除くと、mAP=0.662 という値になる。

転移学習後のモデルを用いて交通量計測を実施した。歩行者、自転車、自動二輪(バイク)の3つのクラスで実施し、実測値と観測値の一致率を精度として表す。3地点において観測を実施し、地点1については転移学習実施時の学習画像が含まれている地点であり、12:00-13:00, 18:00-19:00 の2時間の画像が学習されている。地点2,3は学習画像が含まれていない。

交通量計測結果を表-4に示す。なお、画面奥側に向かう方向を going 方向、画面手前側に向かう方向を coming 方向と表している。学習画像が含まれている地点については、学習画像が含まれていない時間帯においても一致率が高い結果となった。一方で、学習画像が含まれていない地点は、歩行者や自転車一方のみ高いなどばらつきが生じるという結果が得られた。

## 5. 速度推定手法の概要

速度推定を行うためには、画像内の距離を取得する必要がある。しかし、画像内の距離は、カメラから遠ざかるにつれて1ピクセルあたりの距離が大きくなるという特性がある。1ピクセルあたりの距離を合わせるために、画像の画角を真上から見た状態に変換することが可能になれば、画像内の距離の比率は一定になる。

このような変換は、射影変換(ホモグラフィ変換)を行うことで可能になり、変換前の座標を  $(x, y)$ 、変換後の座標を  $(x', y')$  とすると、ホモグラフィ行列は式(3)と表される。式(4)より、8個の未知のパラメータを求めることになるため、変換前の座標4点、変換後の座標4点を定める必要がある。本研究においては、撮影に用いた画像と GoogleMap から得られる航空写真の位置情報を合わせて、ホモグラフィ行列を取得した。

$$s \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ e & h & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3)$$

$$\begin{cases} x' = \frac{ax+by+c}{gx+hy+1} \\ y' = \frac{dx+ey+f}{gx+hy+1} \end{cases} \quad (4)$$

速度推定手法の手順について説明する。まず、YOLOv5 と StrongSORT を組み合わせることによって物体の検出と追跡を行い、座標を取得する。そして、取得した座標を算出したホモグラフィ行列を用いて変換する。また、各 ID ごとに追跡開始時と終了時のフレーム差分を取得して、追跡フレーム数を算出する。移動距離は各 ID ごとに前後フレームのユークリッド距離を計算し、足し合わせることで総移動ピクセル距離が算出できる。最後に、移動ピクセル距離を地図の縮尺情報などから実距離に変換し、動画のフレームレート情報を取得した上で、総移動距離をフレーム数で割ることで移動速度の算出が可能になる。

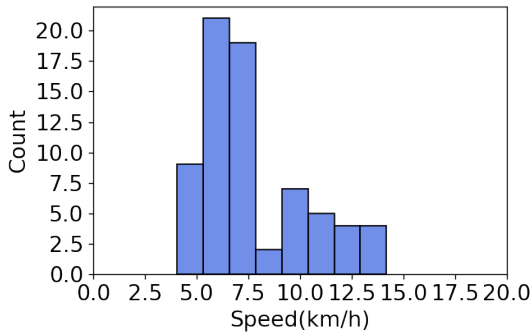


図-2 歩行者の移動速度

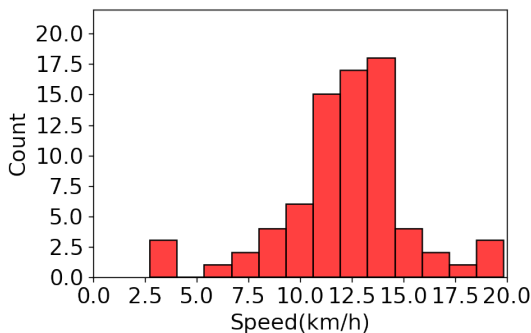


図-3 自転車の移動速度

提案手法を用いて歩行者、自転車の平均速度を導出した結果のヒストグラムが図-2、図-3であり、平均速度は歩行者が7.7km/h、自転車が12.3km/hと算出された。歩行者の平均速度に注目すると、Waterら<sup>9)</sup>の調査によれば、通常時の平均歩行速度は4.8km/h、速い歩行速度は6.3km/hであると示されており、今回導出された移動速度は少々早い。考えられる原因は、物体検出時において歩行者と自転車の誤判定が発生している点と、今回用いたGoogleMAPの地図の縮尺の精度が正確でないことが挙げられる。一方で、歩行者と自転車の速度差が確認できていることから、今回の手法は時間帯別やモード別の速度差の導出といった分析には有用であると考えられる。

## 6. 地点別自己学習手法の提案

### (1) Domain Adaptation の応用

#### a) 現状モデルにおいての問題点

アノテーションデータセットを作成し転移学習を行うことで、学習を行った地点の検出精度については上昇することが確認された。一方で、学習を行っていない地点、横断歩道上や白いタイルの上などにおいては、

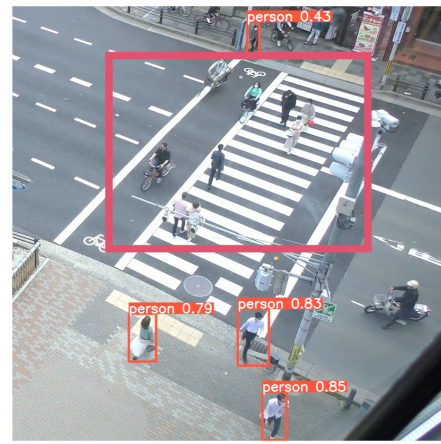


図-4 横断歩道上での検出率の低下例

検出精度が低下するという問題が存在する。解決策としては、検出精度が低い条件や地点においてのアノテーションデータセットを作成し、学習することが挙げられるが、全ての観測地点において学習データセットを作成することは不可能である。このような問題を解決するために Domain Adaptation という手法を応用する。

#### b) Domain Adaptation

機械学習モデルは、学習データと観測データの分布に差が生じている場合、適切な性能を発揮できない場合が生じる。このような現象を解消するための手法が Domain Adaptation(DA)である。Domainとは主にデータセットのことを指し、学習データを Source Domain、観測データを Target Domainと呼ぶ。DAは Source Domainと Target Domainの分布の差を軽減することを目指す手法であるため、本研究においては学習データセットを実際の観測データに近づけるために DAを応用する。

図-4において、横断歩道上での検出率の低下が確認できることから、物体検出は背景による影響が大きいと考えられる。また、中島ら<sup>10)</sup>はプラスチック部品の外観検査において物体検出アルゴリズムを使用し、不良品の検出方法について検討を行った。一般的な欠陥パターンのデータ拡張を行い実際の背景に合成して学習したところ、精度の向上が確認された。このように本研究においても、背景画像を観測地点とする合成画像学習データを作成し、SourceとTargetのDomainの分布を近づける。

#### (2) 合成画像を使用した実験

##### a) 学習画像生成の手順

学習画像の作成に関しては、Mask-R-CNNというFaster-R-CNNをもとにしたインスタンスセグメンテーションモデルを使用する。インスタンスセグメンテーションは、物体検出に加えて物体の形状まで抽出し、画像をピクセル単位で分類を行う。



図-5 Mask-R-CNN を用いた人物画像の生成手順

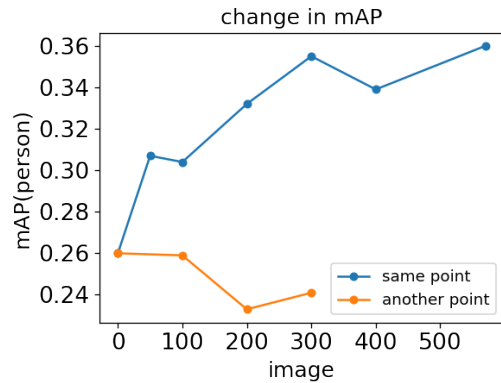


図-7 背景合成画像による学習結果



図-6 作成した合成画像 (観測地点 (左) 観測外地点 (右))

はじめに、YOLOv5 を用いて物体検出を行い、歩行者のみを切り抜いた画像を収集する。その後、Mask-R-CNN を用いて歩行者の輪郭情報を取得する。輪郭情報を用いて、歩行者を切り抜き、背景に合成する。今回は、観測対象地点である横断歩道を背景にした画像と、比較のために一般道路を背景にして合成した画像を作成した。

#### b) 実験方法と結果

転移学習時に使用したデータに、地点 1 が背景の合成画像を加えたデータセット、一般道路が背景の合成画像を加えたデータセットを作成し、学習データとする。加える合成画像は 100 枚ずつ増加させた。評価データは地点 1 のアノテーションデータのみとし、各データセット同士の mAP を比較する。なお、観測対象地点が same-point、比較地点は another-point とする。

実験結果は図-7 のように示され、観測地点と同一の背景を用いた場合において mAP の増加傾向が確認された。0-300 枚の間で mAP の上昇率は大きく、300 枚以上については大きな変化が見られなかった。一方で、一般道路が背景である合成画像を加えた場合 mAP の増加は見られず、合成画像を加えていない場合と比較して減少することが示された。どちらのデータセットにおいても学習に加えた歩行者の人数は同様であるため、背景による影響が大きいと考えられる。よって、観測地点を背景画像にした合成画像による学習は精度向上において有効な手段だといえる。

このような結果を踏まえると、交通量計測 AI の精度を上昇させる手段として、データの地点数を増加させるだけでなく、各クラスごとのデータのバリエーションを増加させ、観測地点ごとに背景を用いた自己学習を行うことが現実的な手段だと考えられる。

## 7. おわりに

本研究では、二輪車および歩行者の交通量を正確に算出する多目的交通計測 AI の開発のために、背景画像に注目した合成画像の作成手法の提案、深層学習ベースの物体検出モデルと物体追跡モデルを用いた新たな速度推定手法の提案を行った。

結果として、観測地点の画像を背景にした合成画像を学習させることで、対象観測地点において精度の向上が確認できた。速度推定手法に関しては、YOLOv5 と StrongSORT を組み合わせることで、位置情報と時間情報の取得が可能になり、射影変換によって距離情報を取得することで実現可能となった。

しかしながら、本研究に取り組んだ過程において明らかになった課題もある。一番大きな問題は物体検出モデルの精度である。本研究において、2000 枚程のアノテーションを作成したが、自転車や自動二輪のデータが少なくデータに偏りが生じてしまった上に、出現率が低いクラスのバリエーションを持たせることができなかった。そのため、データを増やすだけでなくモデルの学習を効率的に進められるように戦略的なデータ収集の必要性がある。また、今回提案した速度推定手法については観測地点ごとに手作業で射影変化を行う必要があるため局地的な使用に留まってしまう。よって、全国規模で使用するためには各地点において自動で距離情報を得られる手法についても検討する必要がある。

今後の展望として、地点ごとの状況に合わせて自動で学習する自己学習の機能を搭載した AI の開発が求められる。観測地点の画像を背景とした合成画像を学習することで精度が向上したように、地点別や観測特徴に合わせた AI の開発の行うことで更なる精度の向上が期待できる。速度推定手法については課題でも述べたように、射影変換の自動化が必須だ。画像の特徴量を用いて自動でマッチングを行う手法を開発し、衛生画像とカメラ画像を組み合わせることで全国の観測地点で速度の推定が可能になるだろう。

## 謝辞

本研究は国土交通省新道路技術会議「カメラ画像および複数の観測データを融合した次世代交通計測手法に関する研究開発」の助成を受けたものです。ここに記して感謝致します。

## 参考文献

- 1) 国土交通省道路局企画課道路経済調査室, 国土交通省国土技術政策総合研究所道路研究室: CCTV カメラ (AI 解析) の制度に関する報告, 2021.
- 2) 片岡源宗, 岡宏一, 熊谷靖彦, 岩佐隆: CCTV を用いた安価な道路監視・計測システムの開発, *土木計画学研究・講演集 (CD-ROM)*, Vol.35, 2007.
- 3) Caballo, Amie Rosarie and Aliac, Chris Jordan: Yolo-based tricycle detection from traffic video, *Proceedings of the 2020 3rd International Conference on Image and Graphics Processing*, pp. 12–16, 2020.
- 4) Jia, Wei and Xu, Shiquan and Liang, Zhen and Zhao, Yang and Min, Hai and Li, Shujie and Yu, Ye: Real-time automatic helmet detection of motorcyclists in urban traffic using improved yolov5 detector, *IET Image Processing*, Vol.15, No.14, pp.3623–3637, 2021.
- 5) Kurniawan, A. and Ramadlan, A. and Yuniarno, E. M.: Speed monitoring for multiple vehicle using closed circuit television (cctv) camera, *2018 International Conference on Computer Engineering, Network and Intelligent Multimedia (CENIM)*, pp. 88–93, 2018.
- 6) Wang, Chen and Musae, Aibek: Preliminary research on vehicle speed detection using traffic cameras, *2019 IEEE International Conference on Big Data (Big Data)*, pp. 3820–3823, 2019.
- 7) ultralytics: Yolov5, <https://github.com/ultralytics/yolov5>.
- 8) Du, Yunhao and Song, Yang and Yang, Bo and Zhao, Yanyun: Strongsort: Make deepsort great again, *arXiv preprint arXiv:2202.13514*, 2022.
- 9) Waters, Robert L and Lunsford, Brenda Rae and Perry, Jacquelin and Byrd, Richard: Energy-speed relationship of walking: standard tables, *Journal of Orthopaedic Research*, Vol.6, No.2, pp.215–222, 1988.
- 10) 中島彩奈, 西谷一希, 茂木和弘, 田中雄太, 白石洋一: 機械学習にもとづくプラスチック製品の外観検査における欠陥と背景の画像合成に関する検討, *エレクトロニクス実装学会誌*, Vol.22, No.6, pp.559–567, 2019.

(2023. 4. 13)

# DEVELOPMENT OF MULTI-PURPOSE TRAFFIC MEASUREMENT AI FOR BICYCLES AND PEDESTRIANS AND PROPOSAL OF A POINT-BY-POINT SELF-LEARNING METHOD

Kentaro OBARA, Hideki YAGINUMA, Shintaro TERABE, Haruka UNO  
and Yu SUZUKI

In recent years, AI-based image analysis methods have been introduced in automobile traffic volume measurement, and a system for automatic observation has been established. However, the measurement accuracy for bicycles and pedestrians is low, and there are high expectations for the development of multi-purpose traffic measurement AI that can acquire not only traffic volume but also movement speed, movement trajectory, and other information. This study aims to develop a multi-purpose traffic measurement AI that can estimate traffic volume and vehicle speed for bicycles and pedestrians. In particular, we trained a convolutional neural network (CNN) using new training labels and utilized information obtained from object tracking to enable traffic volume and speed estimation.

In addition, we proposed a self-learning method for each location to resolve the difference in measurement accuracy among locations, and confirmed the effectiveness of learning synthetic images using the background images of the observation locations.