深層学習を用いた人物追跡手法によるバス乗降者数の推定

Estimating the Number of Bus Passengers Using a Person Tracking Method Based on Deep Learning

西浦 翼[1*]　横山 想一郎[2]　山下 倫央[2]　川村 秀憲[2]
佐藤 好美[3]　長谷川 怜[3]　平澤 幸[3]

Tsubasa NISHIURA[1], Soichiro YOKOYAMA[2], Tomohisa YAMASHITA[2], Hidenori KAWAMURA[2],
Yoshimi SATO[3], Rei HASEGAWA[3], Miyuki HIRASAWA[3]

[1] 北海道大学 大学院情報科学院
[1] Graduate School of Information Science and Technology, Hokkaido University
[2] 北海道大学 大学院情報科学研究院
[2] Faculty of Information Science and Technology, Hokkaido University
[3] 株式会社シーズ・ラボ
[3] C's Lab Co., Ltd.

Abstract: In this research, an algorithm for estimating the number of bus passengers using tracking for highly accurate estimation of bus passenger OD data was proposed and validated. To verify the practicality of the estimation algorithm, digital cameras were installed at the boarding and alighting gates of a bus route and collected video data of passengers like the actual operation. To estimate the number of bus riders, we proposed a method to track passengers and estimate the number of riders by applying tracking technology. Our proposed estimation algorithm was applied to the collected video data and verified the estimation accuracy. As a result of our experiment, we confirmed that the number of passengers can be estimated with high accuracy when a wide interval is maintained between passengers. Based on the result that the estimation accuracy was not high, we confirmed the problems in the object detection, Re-ID, and the judgment of boarding and alighting, and discussed the improvement plan of the estimation accuracy of the number of bus passengers.

1　はじめに

バス事業者の経常収支率は平成 27 年以降悪化傾向にあり，令和 2 年 11 月に国土交通省が発表した令和元年度乗合バス事業の収支状況 [1] によると，黒字事業者は調査対象事業者全体の 27.7 %となっている．特に地方都市では人口減少と共にバス利用者も減少しているが，路線バスは地域住民にとって重要な移動手段である．また，2019 年 4 月に東京池袋で発生した高齢運転者による母子死亡事故をきっかけに，運転免許の自主返納が社会的に促進されており，公共交通機関である路線バスが注目されている．

このような背景に基づき，バス事業存続のために経営状況の改善が求められており，効率的なダイヤや路線の策定が求められている．効率的なダイヤ・経路探索の分野では近年人工知能技術や IoT 技術の活用が始まっており，バス利用者の乗降情報を集計したバス乗客 OD(Origin-Destination) データの分析による収支改善が望まれている．実際，国土交通省はビッグデータの活用等による地方路線バス事業の経済革新支援やビッグデータの活用によるデータ収集・分析ツールの発展高度化を目指している [2].

バス乗客 OD データには次の 2 種類がある．各バス停の乗降者数を記録した「バス停別乗降者数」と，各乗客が乗車したバス停と降車したバス停を記録した「乗車区間別乗降者数」である．バス停別乗降者数からバス停や時間帯毎の利用者の多寡を分析できれば，そのバス停付近の路線の追加・削除やその時間帯の増便・減便が可能になる．また，乗降区間別乗降者数から人流を分析できれば，利用区間毎の運賃の見直し等にも活用することができる [3].

バス乗客 OD データの収集方法として，IC カードが

*連絡先：北海道大学 大学院情報科学院
〒 060-0814 札幌市北区北 14 条西 9 丁目
E-mail: tsubasanishiura@ist.hokudai.ac.jp

挙げられるが，多くの場合バス事業者と IC カードを発行・管理する事業者が異なるため，IC カードの利用情報の管理は IC カードを発行・管理する事業者が行っており，バス事業者側が触れることはない．また，IC カードの普及が進んでいない場合には，調査員がバスに乗り込んで目視で乗降者数を調査する方法もあるが，人的コストが高く，利用者の傾向を把握できるほど多くのデータを収集することは困難である．

平成 28 年度に国土交通政策研究所が自治体に対して行ったアンケートによって，バス乗客 OD データの必要性がを裏付けられている [4]．民間バス事業者の持つデータについて，交通行政の充実化のために地方公共団体として必要なものの回答 684 件の中で 56.4 %にあたる 386 件が「各バス停間の OD」であった．その一方で「対象路線の各バス停間の OD」を民間バス事業者から収集できているのは回答 581 件の中でわずか 11.2 %にあたる 65 件であった．このようにバス乗客 OD データはバス事業者が必要であると考えていながら収集できていない．そのため，低コストで長期間にわたりバス乗客 OD データを収集するシステムの開発が求められている．

このような背景に基づいて，これまで我々はバス乗客 OD 推定システムの開発を進めてきた．本研究では，バス乗客 OD 推定におけるバス停別乗降者数の推定に対してトラッキング技術を導入して，精度向上を図る．実際の路線バスを用いた模擬的な乗降客の動画データセットを用いて，トラッキング技術を用いたバス停別乗降者数推定アルゴリズムの効果を検証する．

本稿の構成を以下に示す．2 章では，従来研究および我々の取り組んできたバス乗客 OD 推定に関する研究について概説する．また，新たに導入するトラッキングに関する研究についても概説する．3 章では，本研究で利用する乗降客の動画データセットを説明し，提案するバス停別乗降者数推定アルゴリズムを説明する．4 章では，提案アルゴリズムの有効性を確認するための実験設定とその結果について説明して，実験結果を考察する．最後に 5 章で本論文をまとめる．

2　関連研究

本章では, バス車内モニタリングの従来研究を概観し，これまでに我々が開発してきたバス乗客 OD データの推定システムの概要を説明する．さらに，本研究でシステムに導入するトラッキングの概要について説明する．

2.1　バス乗客 OD データ推定システム

近年バス車内の状況をモニタリングするシステムは様々な種類が開発されている．バスの天井に二酸化炭素濃度センサーを設置してバス車内の二酸化炭素濃度を計測して，おおよその混雑度を把握し，温湿度計の数値と合わせて乗り心地の数値化や混雑状況の配信を行うシステム [5] や，赤外線センサーと GPS アンテナによって乗降情報・座標・日付・速度・進行方向を蓄積するシステム [6] がある．

カメラを用いたシステムとして，バス車内の乗客の移動を検知して乗客に注意喚起のアナウンスを行い，座っていない乗客を検知してドライバーに注意喚起のアナウンスを行うシステム [7] や，乗車口と降車口の天井に設置したカメラによって乗降人数をカウントするシステムがある．これらのシステムを用いることでバス停別乗降者数はある程度測定できるが，乗客個人の特徴を抽出していないため，バス乗客の個人単位の乗降データである乗車区間別乗降者数は測定できない．

乗車区間別乗降者数を推定する研究としては，バス車内に設置した 1 つのカメラから乗客の顔を検出し，バス停間で車内に存在する人物の差分から推定する研究 [8] がある．しかし，近年のマスク着用率の増加に伴う顔認識アルゴリズムの精度低下の報告 [9] や，手すりや乗客同士のオクルージョンの問題などから高精度の推定は難しい．

バス停別乗降者数と乗車区間別乗降者数の両方を推定するため，我々が開発してきたバス乗客 OD データ推定システムにおける処理の概要を図 1 に示す．

このシステムを用いたバス停別乗降者数の推定では，まず，乗車口と降車口に取り付けたカメラで取得した画像に物体検出モデル適用して，バス乗客が映っている矩形を抽出する．次に，抽出した矩形の画像に対して，似た特徴を持つ人物矩形が近く，異なる特徴を持つ人物矩形が遠くなるように学習させたメトリックラーニングを用いた個人識別器を適用して，似た特徴を持つ人物矩形群を同一人物としてバス停別乗降者数を推定する．最後に，乗車人物と降車人物の矩形群にメトリックラーニングを用いた個人識別器を適用して，近いクラスターを同一人物として 乗客 OD を推定する．

2.2　トラッキングの導入

本研究では，バス停別乗降者数の推定に，乗車人物・降車人物のトラッキング技術を導入する．トラッキングを導入する理由としては，単純なメトリックラーニングと比べて，トラッキングでは乗客の映っている画像の矩形の位置情報や時系列情報を利用するため，精度向上が期待できることが挙げられる．

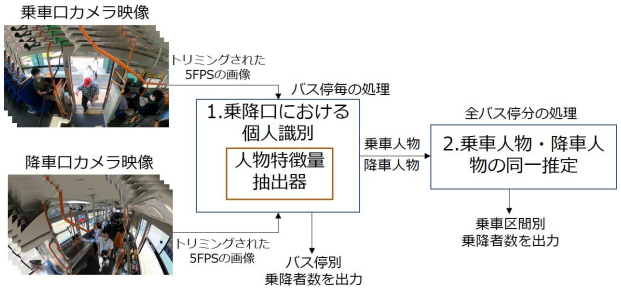


図 1: バス乗客 OD データ推定システムの処理の流れ

トラッキングは主に画像内に映っている物体や人物を検出するタスク（Object Detection）と，複数の画像で検出した物体や人物をマッチングするタスク（Re-Identification）が組み合わさったタスクである．Object Detection と Re-Identification のそれぞれに多くの課題が存在する．また，評価指標もトラッキングを行う目的に応じて変化する．

また，トラッキングでは，追跡の対象となるデータによって難易度が大きく変わる．まず，トラッキングのタスクの難易度は追跡する対象の多寡によって大きく変わる．リアルタイムで対象を追跡するオンラインのタスクでは，現在までのフレームしか使えない．それに対して，全区間のフレームが利用可能なオフラインのタスクでは，ある時点からの未来にあたるフレームも利用することができる．対象を追跡する時間が長い場合は同一対象に同一の ID を振り続けるために，対象が画像内に映っていないオクルージョンの対策が必要となる．さらに，対象の細かい状態や動作も追跡することが求められるのであれば，対象を矩形で検出するのではなく，セグメンテーションとすること必要となる．

本研究では，バス停別乗降者数の推定と乗降区間別乗降者数の推定に向けて，乗車口と降車口のそれぞれで同一人物群の生成を目的としている．また，リアルタイムにバス停別乗降人数や乗降区間別乗降者数が必要とされてはいないため，追跡の対象人物が数人で，オフラインでの使用を前提としたモデルが適用可能である．

3 バス停別乗降者数推定アルゴリズムの提案

本章ではバス停別乗降者数の推定の対象となるバス乗客乗降動画データを説明して，トラッキングを用いたバス停別乗降者数推定アルゴリズムを提案する．

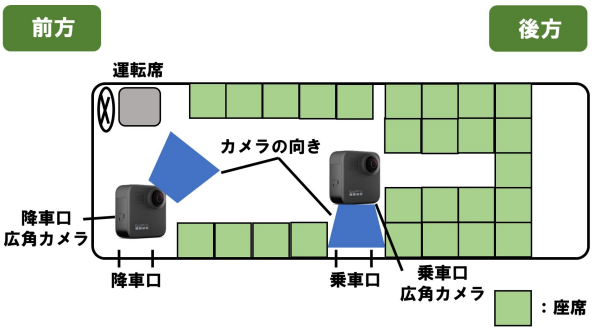


図 2: カメラの設置箇所

3.1 バス乗客の乗降動画データ

バス停別乗降者数の推定アルゴリズムの検証を行うために，実際の旅客運送業務に利用されているバスを用いて，市街地を実際に周回しながら乗客を模擬した人物が乗降するデータを 2020 年 9 月 9 日，2021 年 9 月 28 日，10 月 22 日の 3 日間で収集した．

図 2 に示されるようにバス車内の乗車口・降車口にカメラを設置して，4fps で撮影した．乗客が簡単に触れることができない天井近くにカメラを設置した．撮影日ごとに使用した機器が異なっており，動画の解像度が異なっている．撮影はドアが開く 10 秒前から撮影が開始され，ドアが閉まってから 10 秒後に撮影を終了している．各バス停における乗車数と降車数・カバンの有無・マスクの有無・上着の有無を事前に設定しておき，それに従って乗客がバス停で乗降を繰り返した．

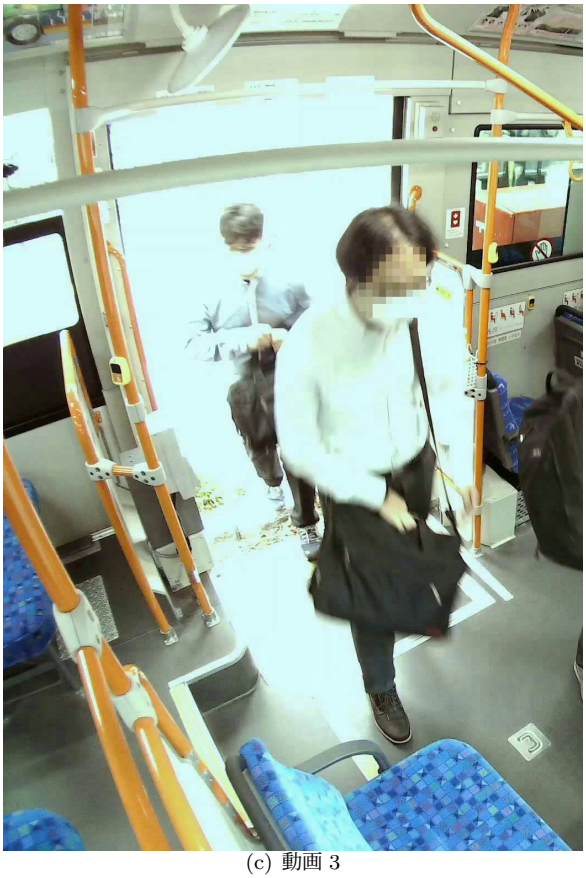
本研究では，収集したデータセットから表 1 に示される 5 つの動画データをバス停別乗降者数の推定アルゴリズムの検証に用いる．

図 3(a)，(b)，(c) に示される動画 1，2，3 の乗車口の動画データはいずれも乗降時間が 30 秒前後である．ただし，動画 2 は動画 1 と動画 3 の乗客 4 人よりも多い乗客 10 人が乗車してくるためオクルージョンが多い．

図 3(d) に示される動画 4 の降車口の動画データでは，乗降時間が約 1 分で 6 人しか降車せず，乗客が広い間隔を保って移動している．図 3(e) に示される動画 5 の降車口の動画データでは，26 秒で 6 人が降車するため，降車する乗客間の間隔が狭い．また，バス中央の通路の手前側には両端に降車しない乗客が写っているため，全体的に乗客間の間隔が狭い．動画 4 と動画 5 を比較すると，動画 4 を撮影したカメラは設置位置が高いため，オクルージョンが少ない．動画 5 はカメラの設置位置が低く，オクルージョンが多い．

(a) 動画 1

(b) 動画 2



(c) 動画 3

(d) 動画 4

(e) 動画 5

図 3: 乗降口で撮影した動画のスクリーンショット

表 1: アルゴリズムの検証に用いるバス乗客乗降動画データ

	カメラ位置	撮影日	乗降者数	動画時間	乗降時間	画像枚数	幅	高さ
動画 1	乗車口	2020/09/09	4	0:43	0:23	173	1,920	1,080
動画 2	乗車口	2021/10/22	10	0:54	0:34	217	1,410	1,320
動画 3	乗車口	2021/09/28	4	0:46	0:26	184	1,050	1,536
動画 4	降車口	2020/09/09	6	1:21	1:01	325	1,920	1,080
動画 5	降車口	2021/09/28	6	0:46	0:26	184	1,410	1,536

3.2 バス停別乗降者数推定アルゴリズム

バス停別乗降者数の推定アルゴリズムの概要を図 4 に示す. まず，物体検出モデルを適用して，バスの乗車口・降車口を撮影した画像から乗客を抽出する．次に，トラッキングモデルを適用して，複数フレームで得られた乗客矩形画像を入力として，乗客の軌跡を推定する．最後に，乗車口で得られた軌跡に対応している乗客が乗車しているか，降車口で得られた軌跡に対応している乗客が降車しているかをルールベースで判定し，バス停別乗降者数を推定する．

物体検出の従来研究において，高い認識精度を示すモデルとしては, FasterR-CNN[10] や tracktor++[11], Scaled-YOLOv4[12] や EfficientDet[13] が挙げられる．本研究では，それらのモデルよりも高い精度で高速な推論を行うことができる YOLOR[14] を使用する．

バス停別乗降者数の推定はオフラインで行われることを想定しているため, 既存研究におけるオフラインのトラッキングモデルとして MPNTrack を用いる．MPNTrack は追跡途中で対象人物の軌跡が入れ替わってしまう ID スイッチの少ないモデルで [15], 2020 年に行われたマルチオブジェクトトラッキングのベンチマークとなっている MOTChallenge[1]の MOT20 において最も ID スイッチの発生回数が少なかった．現在の研究においては，トラッキングのモデルとしては Siamese CNN[16] をベースとしたものが多いが，MPNTrack はマルチオブジェクトトラッキングの定式化に Message Passing Network [17] を応用したモデルである．

MPNTrack は入力される人物矩形画像をノードとして見なして，異なるフレーム間の人物矩形画像のノードをグラフで接続する．次に人物矩形画像を CNN を用いて特徴量を埋め込み，neural message passing を通して情報を伝播させて，アクティブなエッジとノンアクティブなエッジに分ける．アクティブなエッジを結ぶノードの矩形群を乗客 1 人分の軌跡とする．

乗降の判定に関しては，それぞれ以下の判定方法を適用する．乗車口の画像では，ドアが開く前から映っている人物は既にバスに乗っている乗客である．また，

[1]https://motchallenge.net/

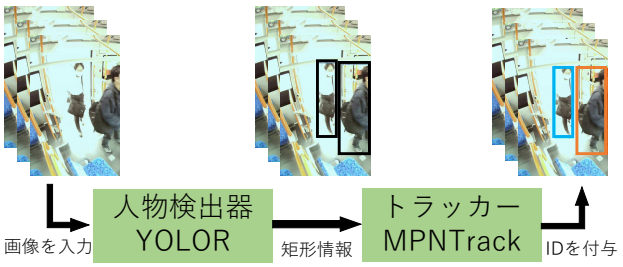


図 4: トラッキングのアルゴリズム

降車しようとするバス後方にいる乗客はバス後方 (画像の左) から現れて，バス前方 (画像の右) に移動する．そのため，乗車の判定に関しては，ある乗客の軌跡が，i) ドアが開いた時間以降に現れている，ii) 最初のフレームの矩形の中心座標が乗車口の範囲に入っている，という 2 つの条件を満たしている場合に乗車と判定する．

降車の判定に関しては，ある乗客の軌跡が，i) ドアが閉まる時間以前に軌跡が消失している，ii) 最後のフレームの矩形の中心座標が降車口の範囲に入っている，という 2 つの条件を満たしている場合に降車と判定する．

4 実験と考察

第 3 章で示したバスの乗降口の映像データにバス停別乗降者数推定アルゴリズムを適用し，乗降者数の推定結果から有効性を検証する．また，人物のトラッキングついての評価指標を算出し，乗降者数のより正確な推定に向けたアルゴリズムの課題点について考察する．

4.1 実験設定

人物検出器として, MSCOCO で学習済みの YOLOR を使用する．YOLOR により「person」クラスの確信度が 30 %以上として推定された矩形を人物矩形とみなしてトラッキングモデルに入力する．ただし，予備実

験において，乗車口のドア付近でオクルージョンにより体の一部しか検出できなかった人物矩形やドアの奥に写る別のバスを待って並んでいる人物や歩行者の人物矩形などが乗降者数推定に悪影響を与えることが確認されたため，こうした人物矩形を取り除く目的で，縦ピクセル数と横ピクセル数の積が 30,000 ピクセル以下の人物矩形は人物検出器の推定結果から除外した．同様に，推定結果の評価指標を計算する際のアノテーションデータからも 30,000 ピクセル以下の人物矩形を除外した．

トラッカーとして，MPNTrack を用いる．MPNTrack のネットワーク構造に含まれる特徴量抽出器には [15] と同様に ImageNet[18] で事前学習させた ResNet50[19] を Market1501[20], CUHK03[21], DukeMTMC[22] の 3 種類のデータセットで再学習させたものを用いる．乗車口および降車口の範囲は予備実験により良好な結果が得られるものを経験的に指定した．その一例を図 5 に示す．検証用データセットの動画ごとにカメラの設置位置や画角が異なるため，乗車口および降車口の範囲は動画ごとに独立して決定した．

トラッカーから得られる各人物の移動軌跡から乗降者数をルールベースで推定する際に，降車口付近で立っているだけの人物がトラッキングの誤推定のために誤って降車したとみなされる状況がみられた．このため，人物矩形の中心座標の x 軸の移動範囲を x_c，y 軸の移動範囲を y_c としたときに移動面積 $x_c{}^*y_c$ を計算し，移動面積が 500 以下であれば降車人物から除いた．データセットに含まれる動画に対して乗降者数推定アルゴリズムを直接適用し，乗降口のドア付近のみのトリミングといった前処理は実施しない．これは，動画のフレームレートが 4fps と決して高くはなく，予めバス車内にいた人物がドアの前を横切る際にドアと重ならない位置から映り込まなければ乗車人物と区別ができないためである．

乗降者数推定アルゴリズムの有効性検証のため，検証用データセットに含まれる各動画に対し，乗降者数の推定結果と真値との比較を行う．また，アルゴリズムの改良に向けた課題点の把握のため，トラッキングタスクの評価指標としてよく用いられる CLEAR MOT を用いて人物検出器およびトラッカーの精度を評価する．CLEAR MOT では，人物検出器により推定される人物矩形の座標値と，人物矩形に対しトラッカーにより推定された人物 ID の両方を用いて精度が評価される．前フレームで正しく推定されていた人物 ID に対し，推定された人物矩形とアノテーションされた人物矩形の IoU が 0.5 以上であれば，推定結果を正解とみなす．推定対象のフレームで新しく出現した人物や，前フレームで正しい推定が行われなかった人物に関しては，そのほかの推定された人物矩形との照合が行われ，IoU が 0.5 以上の人物矩形が推定されている場合は，新しく対応付けが行われる．動画の全フレームを通して対応づけが行われなかった人物矩形の数や，同一の人物の矩形に対して異なる人物 ID の推定結果が対応付けられた数をもとに，以下の評価指標が計算される．

FN アノテーションされた人物矩形に対して，対応付けられる人物矩形の推定結果が存在しない数の全フレーム間の合計

FP 人物矩形の推定結果に対して，対応付けられるアノテーションされた人物矩形が存在しない数の全フレーム間の合計

IDSW 同一人物がアノテーションされた人物矩形に対して，対応付けられた人物矩形の推定人物 ID が変化した回数の合計

また，これらの指標を総合したトラッキングタスクの評価指標 MOTA は，アノテーションされた人物矩形の総数 GT を用いて次の通り計算される．

$$MOTA = 1 - \frac{FN + FP + IDSW}{GT} \in (-\infty, 1]$$

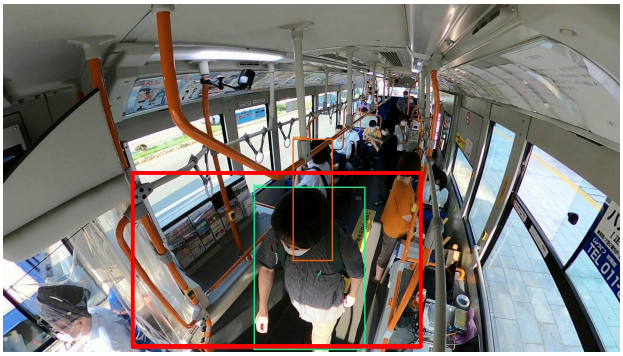
一般的な人物トラッキングに CLEAR MOT を適用する場合と異なり，本稿ではトラッキングを乗車人物または降車人物のみを推定の対象として上記の評価指標を計算する．このため，バスの中にとどまり続ける人物は，人物検出器やトラッカーによる推定が行われるものの，CLEAR MOT の評価指標には反映されない．

4.2 実験結果と考察

各動画にトラッキングを適用し，各種の評価指標を計算した結果を表 2 に示す．乗降者数の推定結果について，動画 4 において正しい人数が推定できているものの，残る 4 つの動画に対しては 30%を超える誤差が生じ実用に足る精度が得られていない．これは，動画 4 は乗者数に対して降車時間が長く乗客が間隔を空けてバスを降りたために，トラッキングの難易度が低かったものの，その他の動画では十分な精度でトラッキング行えなかったためと考えられる．以下では CLEAR MOT の評価指標をもとに精度を悪化させる要因とその対処方法を考察する．

動画 1〜3 の乗車口の動画では FP と FN がどちらも大きな値をとり，MOTA の値を下げている．FP が大きくなる原因に関して，MPNTrack の想定するトラッキングのアノテーションと本研究のデータ作成の際に付与されたアノテーションの相違が挙げられる．MPNTrack は一時的にオクルージョンが発生した場合でも，動画の時系列上で前後に映る同一人物の位置を手掛かりに，

(a) 乗車判定の範囲



(b) 降車判定の範囲

図 5: 乗降判定の範囲．赤線で囲った部分が該当範囲となる．

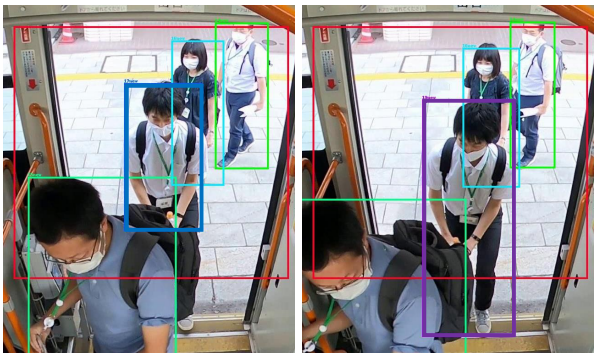


図 6: 乗車口における人物検出器の推定する矩形サイズの変化とトラッカーによる人物 ID の誤推定の例

図 7: 降車口付近で誤検出されることにより乗降者数の推定精度が悪化した人物の例

視認できない人物の領域に対してアノテーションが付与されることを想定している．このために，途中のフレームで人物を検出できなかった際に前後のフレームから矩形を補正するアルゴリズムを備える．一方で，本研究ではこうした一時的なオクルージョンが発生した場合にアノテーションを付与していないため，こうした人物領域に対する推定が偽陽性と判定された．本研究の目的である乗降者数の推定には悪影響を与えていないものの，より正確な評価のためにアノテーションの方針を見直すことが必要である．FN が大きくなる原因としてオクルージョンによるトラッキングの誤推定が挙げられる．図 6 の青枠および紫枠で示された人物矩形の推定結果が示すように，異なるフレームに写る同一人物が人物検出器により推定されるとき，全身が推定される場合とオクルージョンにより上半身の一部のみが推定される場合が混在することがみられた．このとき，トラッカーによりこれらの推定結果に対して異なる人物 ID を割り振られ，人物 ID の切り替わりが発生するケースが多い．乗降口の範囲外でこうした人物 ID の切り替わりが発生した場合，本来は新しく乗車した人物がバス車内に以前から乗っていた人物とみなされ，FN の増加の要因となる．対策としては，上半身など人物の一部分のみが検出された場合にも，同一人物の全身画像と同一の人物 ID が割り振られるよう，トラッカーの学習データにこうした人物画像の組を追加することが挙げられる．

降車口の動画に関して，動画 4 はカメラ位置が高く，人物の移動も疎であるため全ての人物をトラッキングすることができた．動画 5 では FP が大きな値をとり，本来の降車数に比べ過大な降車数が推定されたことがわかる．これは，図 7 の橙枠に示す通路の手前でとどまり降車しなかった人物に関する誤推定により発生した．この人物に対し人物検出器が推定する矩形サイズがオクルージョンにより大きく変化した際に，トラッカーにより推定される人物 ID の切り替わりが発生し，多くの異なる人物が降車したと判定された．こうした問題に対しては，トラッカーの出力結果に対する降車判定のルールの修正や，降車口カメラの取り付け位置を調整してより降車口の扉に近い領域を撮影し，降車口の範囲をバス内にとどまる人物が滞留し得ない領域に設定することによる対処が考えられる．

表 2: 乗降者数推定の結果

	カメラ位置	乗降者数	推定乗降者数	IDSW	MOTA	FP	FN
動画 1	乗車口	4	6	2	0.227	53	54
動画 2	乗車口	10	15	5	0.365	181	71
動画 3	乗車口	4	6	2	0.539	24	87
動画 4	降車口	6	6	0	0.903	2	19
動画 5	降車口	6	11	2	-0.289	270	0

5 おわりに

本研究ではバス乗客 OD データの高精度な推定に向けて，トラッキングを用いたバス乗降者数の推定アルゴリズムを提案し，その検証を行った．バス乗降者数の推定アルゴリズムの実用性を検証するために，路線バスの乗車口と降車口にカメラを設置して実際の運行に近い形で乗降者の動画データを収集した．バス乗降者数の推定のために，トラッキングを適用して乗客を追跡し，乗降者数を推定する手法を提案した．収集した動画データに対してバス乗降者数の推定アルゴリズムを適用して，推定精度を検証した．乗客が広い間隔を保っている場合には，高精度でバス乗降者数を推定可能であることを確認した．また，高い精度で推定が行えなかった結果に基づいて，物体検出，ReID, 乗車・降車判定における課題を確認し，バス乗降者数の推定精度の改善案を検討した．

参考文献

[1] 国土交通省． 令和元年度乗合バス事業の収支状況について. https://www.mlit.go.jp/report/press/jidosha03_hh_000326.html, 2020.

[2] 国土交通省． 「ビッグデータの活用等による地方路線バス事業の経済革新支援」(平成 27 年度) 「ビッグデータの活用によるデータ収集・分析ツールの発展高度化」(平成 29 年) について． https://www.mlit.go.jp/sogoseisaku/jouhouka/sosei_jouhouka_fr1_000014.html, 2017.

[3] 福本雅之. 公共交通利用実態データの調べ方と活用法あなたのまちの公共交通、把握・分析、しっかりできていますか？〜地域公共交通セミナー（秋期）〜. https://wwwtb.mlit.go.jp/chubu/aichi/kikaku/seminafukumotosennsei.pdf, 2019.

[4] 国土交通省国土交通政策研究所． 地方公共交通事業における収益性と効率性に関する調査研究． https://www.mlit.go.jp/pri/houkoku/gaiyou/kkk138.html, 2016.

[5] エコモット株式会社． 混雑状況モニタリング. https://www.ecomott.co.jp/solution/840/.

[6] ITS アライアンス株式会社. 簡易 乗降数カウントシステム あしあとカウンター. http://www.its-alliance.jp/counter.html.

[7] BOLDLY 株式会社. Daily by disptcher. https://www.softbank.jp/drive/service/daily/.

[8] 山田遊馬, 廣森聡仁, 山口弘純, 東野輝夫. 路線バスにおけるカメラ画像を用いた od 計測システムの提案. 研究報告モバイルコンピューティングとパーベイシブシステム（MBL）, 第 93 巻, pp. 1–8, 2019.

[9] Mei L Ngan, Patrick J Grother, and Kayee K Hanaoka. Ongoing face recognition vendor test (frvt) part 6a: Face recognition accuracy with masks using pre-covid-19 algorithms. 2020.

[10] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, Vol. 28. Curran Associates, Inc., 2015.

[11] Philipp Bergmann, Tim Meinhardt, and Laura Leal-Taixé. Tracking without bells and whistles. In *The IEEE International Conference on Computer Vision (ICCV)*, October 2019.

[12] Chien-Yao Wang, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. Scaled-YOLOv4: Scaling cross stage partial network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 13029–13038, June 2021.

[13] Mingxing Tan, Ruoming Pang, and Quoc V. Le. Efficientdet: Scalable and efficient object detection. *2020 IEEE/CVF Conference on Com-*

puter Vision and Pattern Recognition (CVPR), pp. 10778–10787, 2020.

[14] Chien-Yao Wang, I-Hau Yeh, and Hong-Yuan Mark Liao. You only learn one representation: Unified network for multiple tasks. *arXiv preprint arXiv:2105.04206*, 2021.

[15] Guillem Brasó and Laura Leal-Taixé. Learning a neural solver for multiple object tracking. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[16] Gregory Koch, Richard Zemel, and Ruslan Salakhutdinov. Siamese neural networks for one-shot image recognition. 2015.

[17] Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. Neural message passing for quantum chemistry. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Conference on Machine Learning*, Vol. 70 of *Proceedings of Machine Learning Research*, pp. 1263–1272. PMLR, 06–11 Aug 2017.

[18] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pp. 248–255. Ieee, 2009.

[19] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778. IEEE, jun 2016.

[20] Liang Zheng, Liyue Shen, Lu Tian, Shengjin Wang, Jingdong Wang, and Qi Tian. Scalable person re-identification: A benchmark. In *Computer Vision, IEEE International Conference on*, 2015.

[21] Wei Li, Rui Zhao, Tong Xiao, and Xiaogang Wang. Deepreid: Deep filter pairing neural network for person re-identification. In *CVPR*, 2014.

[22] Ergys Ristani, Francesco Solera, Roger S. Zou, R. Cucchiara, and Carlo Tomasi. Performance measures and a data set for multi-target, multi-camera tracking. In *ECCV Workshops*, 2016.