

深層学習を用いた路線バスにおける 乗客 OD および混雑状況の推定手法の開発

山下 倫央
北海道大学

tomohisa@ist.hokudai.ac.jp

久保 田遼裕
北海道大学

kubota@ist.hokudai.ac.jp

川村 秀憲
北海道大学

kawamura@ist.hokudai.ac.jp

西浦 翼
北海道大学

tsubasanishiura@ist.hokudai.ac.jp

横山 想一郎
北海道大学

yokoyama@ist.hokudai.ac.jp

弓崎 潔

株式会社 シーズ・ラボ
yumisaki@cslab.co.jp

佐藤 好美

株式会社 シーズ・ラボ
yoshimi@cslab.co.jp

要旨

本稿では、路線バスを対象としたバス車内モニタリングシステムの開発をおこない、路線バスを安全で利便性の高い公共交通機関とすることを目的とする。バス車内モニタリングシステムの実用に向けて、深層学習を用いた乗客の OD データの推定手法やバス車内の混雑状況の推定手法を提案する。これらの提案手法の有効性を検証するため、実際の路線バスを利用して検証用データを作成した。検証用データを用いた実験をおこない、バス乗客の OD データの推定やバス車内の混雑状況の推定の実用可能性を確認した。

1 はじめに

近年、車両を運転するドライバーを計測して、居眠りやわき見の検知やふらつき運転への警告する運転指導が注目を浴びている。また、車内を撮影するカメラでバス乗客を対象とした車内モニタリングがバスの効率的な運行やバス乗客の満足度の向上の観点から検討されはじめている。これまでのバス事業者に対するヒアリングの結果、乗客の乗降したバス停 (OD : Origin and Destination) の情報やバス車内の混雑状況に対して強いニーズがある

ことが分かってきた。

バス乗客 OD の推定に関しては、バスの乗客の利用状況を集計したバス乗客 OD データの分析による効率的な運営への改善が期待されており、実際に国土交通省はビッグデータの活用した経営革新支援や、データ分析・収集ツールの発展・高度化を図っている [1]。バス料金支払い用の交通系 IC カードの普及で把握が容易になったと思われがちだが、実際は IC カードの利用情報は IC カードの発行・管理をおこなう事業者によって管理されている。IC カードの利用状況の分析結果を利用するためにはバス事業者の追加料金の支払いが必要である。そのため、収益の低い路線バス事業者には利用が困難であるが収益改善には必要であるため、簡易かつ安価に乗客の OD データを収集するシステムが求められている。

また、バス車内の混雑状況の推定に関しては、バス車内動画を解析することで、乗客数や混雑具合を把握して、バスの運行の管理や運行の円滑化に貢献する。さらに混雑状況をリアルタイムに利用者に配信することで、バス利用者がいつ、どのバスに乗るかという判断する手助けになる。店舗・施設や電車では混雑度情報の発信が順調に普及している一方で、路線バスにおいては状況が異なる。路線バスでは鉄道駅のように改札を設置して人数を数えることが難しく、出入口に赤外線センサーを設置し

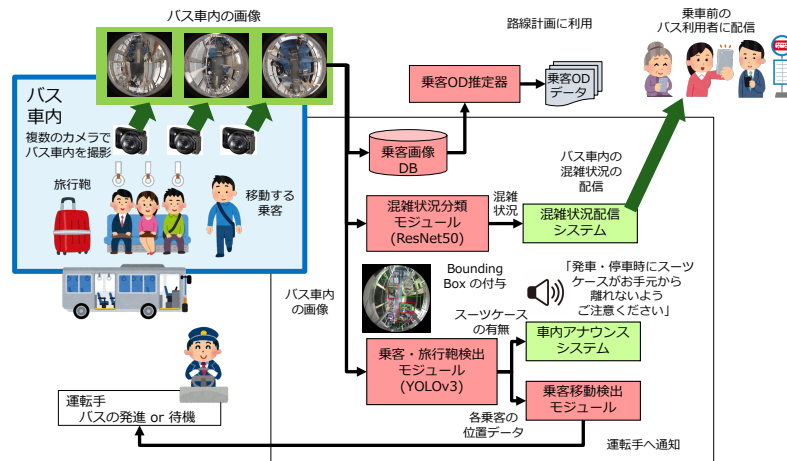


図 1. バス車内モニタリングシステムの構成の概要

でも人や荷物の重なりによって正確に数えることが難しい。西鉄グループのように一部 IC カードの利用実績から混雑度を配信している企業もあるが、前述のバス乗客 OD の推定と同様に、多くの路線バスではバス事業者と IC カード発行・管理する事業者が異なるため、バス事業者が IC カードの利用情報を参照することは容易ではない。そのため、バス事業者が主体となって混雑度を測定することが求められる。

このような状況を踏まえて、路線バスにおいて効率的なバス運行やバス乗客の満足度を向上させるため、バス乗客 OD データを収集し、バス車内の混雑状況を推定するバス車内モニタリングシステムの開発を目指す (図 1)。本稿では、バス乗客の乗降時の画像から OD データを推定する手法とバス車内の混雑状況を推定する手法を提案する。さらに、実際のバスを用いて収集したデータを用いてそれぞれの手法の有効性を検証する。

本稿の位置付けとしては、従来手法を統合して提案するバス乗客 OD 推定手法や混雑状況推定手法に関して、大規模な実データの収集や製品化への本格的な開発をおこなう前に、小規模なバス車内で撮影したデータで提案手法の有効性を確認することである。

本稿の構成を以下に示す。2 章ではバス乗客 OD 推定手法を提案し、3 章で提案手法の有効性を検証する。4 章ではバス車内の混雑状況推定手法を提案し、5 章で提案手法の有効性を検証する。最後に、6 章で本稿の結論と今後の展望を記す。

2 バス乗客 OD 推定手法の提案

本章では、図 2 に示されるように乗降口における個人識別と乗車人物・降車人物の同一推定から構成されるバス乗客 OD の推定手法を提案し、提案手法の実装方法について述べる。

2.1 バス乗客 OD 推定手法の概要

バス乗客 OD データにはバス停毎における乗降人数を記録したバス停別乗降人数と、乗客単位での乗車したバス停と降車したバス停を集計した乗車区間別乗降者数がある。これらのデータを分析することで収益改善のための効率的なダイヤへの見直しや路線の増便・減便が可能となる。さらに、乗車区間別乗降者数を把握することで運賃の見直しにも活用することができる [2]。

バス乗客 OD の推定においては、まず、乗降口における個人識別で乗降口と降車口で人物を識別して、識別された人数をバス停別乗降者数として出力する。乗降口における個人識別は、人物特徴量抽出器で画像から人物画像の抽出と、人物画像の特徴量抽出をする前処理をおこなう。次に乗車人物・降車人物の同一推定をおこない乗車・降車ペアの作成をおこない、乗車・降車ペアを合計して乗車区間別乗降者数として出力する。

2.2 メトリックラーニングの適用

バス車内における同一人物判定においては、訓練データには存在しない未知の乗客の画像を同一人物か否かを

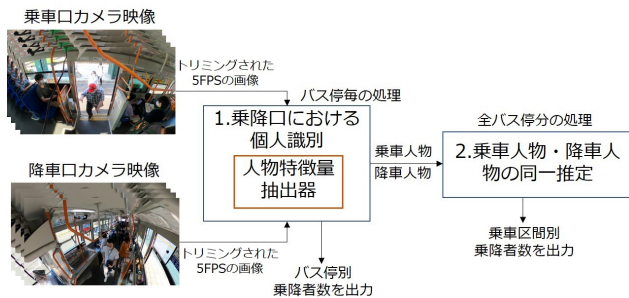


図 2. バス乗客 OD データ推定手法の概要

判定する必要がある。畳み込みニューラルネットワーク (Convolutional Neural Network; CNN) を用いたクラス分類は事前に学習したクラスの分類には有効だが、乗客の OD 推定で求められている未知の2つのデータに対して類似度を計算する構造を持っていないため、事前に学習したクラス以外の分類において高い性能を発揮することは期待できない。そのため、ある未知のデータが類似しているか否かを学習する手法としてメトリックラーニングが開発されている。メトリックラーニングでは、似た特徴を持つデータ間の類似度を近く、異なる特徴を持つデータ間の類似度を遠くなるよう学習がおこなわれ、車両の認証 [3] や顔の認証 [4] といった分野において優れた結果を残してきた。メトリックラーニングは Person Re-Identification という人物を認証するタスクに対しても適用されている [5]。

メトリックラーニングを用いた Person Re-Identification においては、基準となるアンカーに対して異なるクラスであるネガティブなデータよりも同クラスであるポジティブなデータとの距離をより近くなるような学習をおこなう triplet loss を用いた手法 [6] がある。Person Re-Identification において優れた結果を残している深層学習モデル BoT は、画像分類の深層学習モデル ResNet50 と triplet loss を用いた標準的なベースラインに、メトリック損失と分類損失を2つの異なる特徴空間に分離する BNNeck という手法、学習率の Warmup, Random Erasing Augmentation [7], Label Smoothing [8], Last Stride [9] といった複数の手法を適用し、Center Loss [10] と、triplet loss と人物 ID の分類損失である ID loss の3つを損失を組み合わせることで、他手法に比べて高い精度を得られている。そのため、本提案手法では BoT を用いた人物画像からの特徴量の抽出を適用する。

2.3 バス乗客 OD 推定手法の実装

2.3.1 乗降口における個人識別

乗降口における個人識別の処理の流れを図 3 に示す。ここでは、次章で説明する路線バスで収集したバス停別乗降者データの画像サイズ (ピクセルサイズ 1920×1080) である画像を対象とする。

人物画像からの特徴量の算出 路線バスの乗降口を撮影した画像から人物画像を抽出するために、物体検出モデル YOLOv4 を適用して、確信度 60 % 以上で「person」と認識された矩形部分を元画像から抽出する。

まず、乗降中の客ではない人物も多く抽出されてしまう問題がある。例えば、乗車口画像では別の路線のバス待ちで並ぶ人物や歩行者が写りこむことがある。また、降車口画像では乗車後に座席に座ったりつり革に掴まって立っている人物が写りこむことがある。さらに、乗降中の客の中でも身体の一部のみが検出されたり、人物同士が大きく重複する画像は人物画像としてノイズとなる。これらの問題を避けるために、画像サイズが乗車口画像では 650×900、降車口画像では 820×650 にトリミングして、この画像に対して人物画像を抽出する。

次に、人物同士が重複した矩形が画像が抽出されないように、人物が検出されたときに矩形面積を計算し、その面積に 25 % 以上重複する他に検出された物体の矩形がある場合には、その画像を利用しない条件を設定した。また、身体の一部のみが写った画像が抽出されないように矩形サイズの条件と乗降中の乗客以外の人物や車外の人物が抽出されないように矩形の中心位置による条件を設定した。矩形サイズの条件は、縦ピクセル数と横ピクセル数の積が 30,000 以上であることとし、矩形の中心座標が設定した範囲の座標内にある場合に抽出した。

このように、乗降車口の画像から抽出した人物画像から BoT を用いて特徴量を算出する。BoT には人物画像を 128×256 にリサイズして入力し、出力として 2048 次元の特徴量が得られる。

人物画像の特徴量を用いた個人識別 人物画像に対して付与されている出現時刻、検出座標、特徴量を用いて乗降口における個人識別をおこなう。

乗降口における個人識別はバス停毎に乗車口カメラと降車口カメラで別々におこなわれる。個人識別として乗車人物群と降車人物群を作成し、それぞれの人物群の

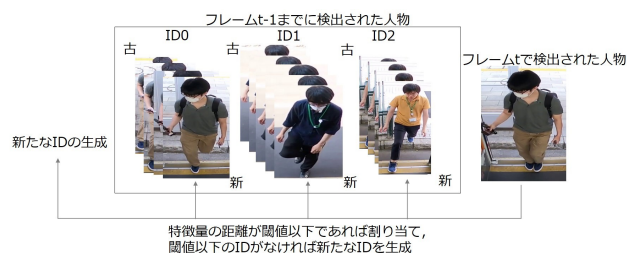


図 3. 乗降口における個人識別の流れ

数を乗車人数、降車人数として出力する。一連の処理は [11] における同一人物のグルーピング手法を参考に、人物画像間の特徴量の距離を算出して比較をおこない、ハンガリアン法により ID を割り当てる。¹ 全ての人物画像間で特徴量を比較するのではなく、図 3 に示されるように、ある人物画像の出現前の時刻までに検出された人物との特徴量とユークリッド距離の比較をおこなう。

ある人物画像と別の人物画像の映っている人物を同一人物と判定するユークリッド距離の閾値を設定し、閾値以下の中でユークリッド距離が最小の人物と同じ ID を割り当てる。閾値以下の人物が存在しない場合は新たな人物として、新たな ID を割り当てる。

各バス停において乗車口と降車口で抽出された人物画像の出現時刻順に ID を割り振っていき、あるバス停 k における乗車人物群 P_{on}^k と降車人物群 P_{off}^k を作成する。

2.3.2 乗車人物・降車人物の同一推定

乗降口における個人識別によって作成された乗車人物群 P_{on} と降車人物群 P_{off} から同一人物と思われる人物のペアを作成するための乗車人物・降車人物の同一推定手法について説明する。本稿での乗車人物・降車人物の同一推定は、全バス停におけるデータ収集が終わった後に一括しておこなう処理方式を想定している。

まず、乗降口における個人識別で全てのバス停における乗車人物群 P_{on} と降車人物群 P_{off} で特徴量のユークリッド距離行列を作成する。このユークリッド距離行列は降車人物群 P_{off} の要素数を i 、バス車内人物群 P_{on} の要素数を j としたとき、 $i \times j$ の行列となる。ただし、この距離行列に対して、ペアを作成するためのハンガリアン法をそのまま適用すると、乗車したバス停より前に

¹ハンガリアン法は `scipy 0.18.1` の `optimize.linear_sum_assignment` を用いて実装する。Scipy.org, <https://docs.scipy.org/doc/scipy/reference/release.0.18.1.html>.

降車した人物とのあり得ないペアが同一人物とされる可能性がある。そのため、ハンガリアン法を適用する前にユークリッド距離行列に対して簡単なバス停ルールを適用する。バス停ルールは、バス乗客はあるバス停で乗車してから、必ずそれ以降のバス停で降車するという仮定が成り立つとして、(i) ある人物の乗車画像に関して、それ以前のバス停での降車時に撮影された人物画像は候補としない、(ii) ある人物の降車画像に関して、それ以降のバス停での乗車時に撮影された人物画像は候補としない、として、該当するユークリッド距離行列の要素を無限大に設定する。

3 バス乗客 OD 推定手法の検証

本章では、バス乗客 OD データ推定手法の検証用のデータ収集について概説し、提案手法により算出した乗降口における個人識別によりバス停別の乗降者数の精度を確認する。

3.1 バス停別乗降者データの収集

バス停別乗降者数と乗車区間別乗降者数の推定手法の検証用のデータを確保するため、路線バスをレンタルして、研究グループのメンバーを含めた 46 名が乗客役となり、撮影をおこなった。画像のピクセルサイズ 1920×1080 の広角で撮影をおこなうデジタルカメラ (GoProMax) を乗車口と降車口の 2 箇所に設置した。カメラの設置位置に関しては、乗客が簡単に触れることができない地上高の高い位置に設置するという条件のもと、実際に運用することを考慮した位置に設置している。

乗車口カメラと降車口カメラで撮影した例を図 4, 5 に示す。乗降者や車内の乗客が増えることによってオクルージョンが頻繁に発生し、画像に乗客の全身が写るタイミングは限られる。

仮のバス停と設定した三地点を周回し、バス停毎に乗降者数を変えながら計 24 回の乗車映像、計 28 回の降車映像を撮影した。

3.2 検証データの作成

特徴抽出に用いる BoT には既存の学習済みモデルがあり、本研究では、さらに Market-1501, DukeMTMC-ReID の学習済みモデルを用いる。しかしそれらのデータセットと比べると、バスに設置したカメラは検出され



図 4. 乗車口を撮影した画像例



図 5. 降車口を撮影した画像例

た人物画像は、上から見下ろした画角が多くなる。予備実験ではバスで収集したデータで訓練データを作成し、新たなモデルを作成する。次に提案したバス乗客 OD 推定手法の検証のため、乗降データを組み合わせることで、計 22 人が 11 箇所のバス停で 7 回ずつ乗降する 22 人のバス停別乗降画像データと、訓練データを含まない計 11 人が 9 箇所のバス停で 5 乗降する 11 人のバス停別乗降画像データを作成した。各乗客は乗車・降車を各 1 度ずつのみおこなった。

3.3 バス停別乗降者数の推定結果

乗降口における個人識別によりバス停別の乗降者数の推定精度の検証をおこなう。乗降口における個人識別された乗降人数と、実際の乗降人数の絶対誤差と絶対誤差率を確認する。11 人のバス停別乗降画像データでは各 5 回の乗降を対象に作成した 2 つの学習済みモデルと作成した 3 つのモデルを合わせた 5 つのモデルで、22 人のバス停別乗降画像データでは各 7 回の乗降を対象に 3 つのモデルで実験をおこなった。

11 人のバス停別乗降画像データの実験結果を表 1, 2,

22 人のバス停別乗降画像データの実験結果を表 3, 4 に示す。いずれのバス停データにおいてもバス停別の推定結果は、推定誤差は 0 または 1 人であり、学習済みモデルに対してファインチューニングモデルの方が同等または少ない誤差で人数の推定を行うことができた。しかし、データによっては複数のモデルで共通して誤差が発生している箇所あり、図 6 のように人物がカメラに写る角度によって、実際は同一人物の画像が別人として識別されたり、逆に別人同士が同一人物として認識された。



図 6. 個人識別の誤判定パターン

また、車内を移動する人物が降車口の人物画像に写りこんでしまい、降車人物として認識されてしまう例もあった。

表 1. 11 人の乗車人数推定誤差

実際の人数	2	2	4	1	2	計
Market-1501 (Pre-trained)	0	1	0	0	1	2
Market-1501 (Fine-tuned)	0	0	0	0	0	0
DukeMTMC-ReID (Pre-trained)	0	0	0	0	1	1
DukeMTMC-ReID (Fine-tuned)	1	0	0	0	0	1
バスデータ モデル	1	0	0	0	2	3

3.4 乗車区間別乗降者数の推定結果

乗車人物・降車人物の同一推定により作成された乗車人物と降車人物のペアに対して、乗車区間別乗降者数を

表 2. 11 人の降車人数推定誤差

実際の人数	2	1	5	2	1	計
Market-1501						
(Pre-trained)	0	0	2	0	0	2
Market-1501						
(Fine-tuned)	0	0	1	0	0	1
DukeMTMC-ReID						
(Pre-trained)	0	0	1	0	0	1
DukeMTMC-ReID						
(Fine-tuned)	0	0	0	0	0	0
バスデータ						
モデル	0	0	1	0	0	1

表 3. 22 人の乗車人数の推定誤差

実際の人数	2	2	7	2	1	4	4	計
Market-1501								
(Pre-trained)	1	0	0	0	0	0	0	1
DukeMTMC-ReID								
(Pre-trained)	0	0	3	0	0	1	0	5

表 4. 22 人の降車人数の推定誤差

実際の人数	2	7	1	4	5	2	1	計
Market-1501								
(Pre-trained)	0	0	0	1	2	0	0	3
DukeMTMC-ReID								
(Pre-trained)	0	0	0	0	1	0	0	1

用いた推定精度を検証をおこなう。作成される乗車人物と降車人物のペアの数は実際のペアと同じ数とは限らないので、評価には再現率、適合率、F 値を用いる。

11 人のバス停別乗降画像データの実験結果を表 5、22 人のバス停別乗降画像データの実験結果を表 6 に示す。

11 人のバス停別乗降画像データでは、DukeMTMC-ReID 学習済みモデルにおいて F 値が 1.000 となった。全般的に高い推定精度が得られたのは、推定の対象となるデータの規模が 11 人と小規模であったため、推定が比較的容易でどのモデルでも推定精度が高く、モデル間で差がつかなかったと考えられる。

22 人のバス停別乗降画像データでは、Market-1501 学習済みモデルに比べ、DukeMTMC-ReID 学習済みモデル

の方が精度が約 20 % 高くなったが、どちらのモデルも 11 人のバス停映像データと比べると精度が下がっている。誤判定の原因としては、学習済みモデルでは姿勢の変化で同一人物であっても距離が大きくなってしまい、別人と判定されてしまうことがあったため、バス停別乗降画像データに大きな姿勢変化を含むデータが少ないことが原因と考えられる。

表 5. 乗車区間別乗降者数の推定精度の比較 (11 人)

モデル	再現率	適合率	F 値
Market-1501			
(Pre-trained)	0.909	0.909	0.909
Market1501			
(Fine-tuned)	0.818	0.900	0.857
DukeMTMC-ReID			
(Pre-trained)	1.000	1.000	1.000
DukeMTMC-ReID			
(Fine-tuned)	0.909	0.909	0.909
バスデータ			
モデル	1.000	0.917	0.957

表 6. 乗車区間別乗降者数の推定精度の比較 (22 人)

モデル	再現率	適合率	F 値
Market-1501			
(Pre-trained)	0.682	0.682	0.682
DukeMTMC-ReID			
(Pre-trained)	0.909	0.870	0.889

3.5 考察

前章で提案したバス乗客 OD 推定手法に対して、バス停別乗降者数と乗車区間別乗降者数の推定をおこない、精度の検証をおこなった。

バス停別乗降者数の推定では、どのバス停でも 0 から 1 人の誤差に抑えて推定することができた。そのため、本稿で提案した乗降口における個人識別手法の有用性を確認することができた。本稿では同一人物かどうかの判定には特徴量距離のみを用いたが、今後は画像内の座標を用いた手法 [11] や、人物画像の撮影時間の時系列情報

を扱うことで、乗降口における個人識別の精度だけでなく、乗車人物・降車人物の同一推定の精度も向上することが期待される。

乗車区間別乗降者数においても小規模であれば、高精度での推定をおこなえることを確認した。また、ファインチューニングにより同一人物・別人での特徴量距離の差は大きくなったが、乗客の姿勢の変化の影響を受けやすくなることを確認した。今回の実験で得られたバス停別乗降者数と乗車区間別乗降者数の推定精度であれば、学習データの数や多様性を増加させることで精度向上が期待できる。そのため、今後は実運用に向けて訓練データに路線バスで実際の乗客が乗降する状況のデータを取得して、バス停別乗降画像データを充実させることを目指す。

4 バス車内混雑度推定手法の提案

本章では、画像分類モデルと物体検出モデルを用いた二つのバス車内の混雑度推定手法を提案する。

4.1 画像分類による混雑状況の分類

画像分類モデルである CNN を用いて、バス車内の画像を入力として、バス車内が空いている、普通、混んでいるの 3 クラス分類をおこなう。乗車人数が 9 名までを空いているとして混雑度 0、乗車人数が 10～25 名の普通として混雑度 1、乗客数が 26 名以上を混んでいるとして混雑度 2 とする。混雑度 1 の上限はバス車内の座席数とした。CNN のモデルとして、ResNet50 を用いる。

4.2 物体検出によるバス乗客のカウント

物体検出モデルを用いて、バス車内の画像を入力として、画像内の人物検出をおこなう。検出された人数によりバス車内の人数をカウントして、混雑度を推定する。物体検出モデルには、Microsoft COCO のデータセット [12] で学習済みモデルの YOLOv3 を用いる。

5 バス車内混雑度推定手法の検証

本章では画像分類モデル ResNet50 を用いた 3 クラス分類と物体検出モデル YOLOv3 を用いた人数カウントの 2 つの手法の精度の検証をおこなう。画像分類による混雑度の 3 クラス分類ではバス車内全体が写っている前

方広角カメラと中央魚眼カメラの画像を用いる。物体検出によるバス乗客のカウントでは、物体検出が適用可能な前方広角カメラと後方広角カメラの画像を用いる。

5.1 バス車内画像の収集

今回実験に用いるデータセットの収集のために、3.1 節で記述されたバス停別乗降者データの収集と並行して、混雑度推定用の検証データも収集した。

混雑度推定用のバス車内画像の収集のために、バス車内の前方に広角カメラ 2 台、バス中央に魚眼カメラ 1 台、バス後方に広角カメラ 1 台を設置した。動画は全て 30fps で撮影されており、サイズは 1,920*1,080 ピクセルである。また、撮影に際して、実際の使用状況に近い撮影をするため、マスク・帽子・上着・鞆の有無、乗降時に座席の変更をおこなった。乗車人数に関しては、3, 4, 6, 7, 8, 24, 25, 38 名の状況を撮影した。バス停からバス停を 1 つのシナリオとして撮影した動画を分割し、30fps の動画を全て画像として切り出した。

予備実験の結果、バス車内全体を撮影できて、手すり等の干渉を受けずに適切な画像分類により混雑度の推定をおこなえる画像を撮影していたのは、前方広角カメラ 1 台と中央魚眼カメラ 1 台であることが確認された。そのため、画像分類による混雑状況の分類には、バス前方広角カメラと中央魚眼カメラで得られた画像データを対象に推定をおこなう。前方広角カメラに関して、73,024 枚の訓練データをバッチサイズ 128, 30 エポックで学習させた。中央魚眼カメラに関して、52,891 枚の訓練データを同様にバッチサイズ 128, 30 エポックで学習させた。

また、中央魚眼カメラの画像は人物の上下が揃っておらず、また端に行けば行くほどゆがみが大きくなり、人物を検出することができなかった。そのため物体検出によるバス乗客のカウントは、前方広角カメラと後方広角カメラで得られた画像データを対象に推定をおこなう。前方広角カメラの画像では後部座席の面積が小さく、座っている人物の多くが頭部しか写っていないため、後部座席最奥中央しか人物の検出をすることができない。そのため前方広角カメラの画像の後部座席部分をマスクして、後方広角カメラと合わせて人数カウントをおこなう。

5.2 画像分類による混雑状況の 3 クラス分類結果

画像分類による混雑状況の 3 クラス分類の結果として、前方広角カメラのテストデータ 22,967 枚と中央魚眼カ

表 7.3 クラス分類の正解率

混雑度	前方広角カメラ	中央魚眼カメラ
0	0.976	0.955
1	0.993	0.940
2	1.00	1.00
全体	0.990	0.958

表 8. 混雑度推定結果

混雑度	推定人数 (前+後)	正解人数 (前+後)
0	4.93 (1.63+3.30)	6.43 (2.93+3.50)
1	16.2 (5.02+11.2)	19.3 (7.84+11.5)
2	19.8 (6.16+13.6)	22.9 (7.88+15.0)
全体	13.9 (4.35+9.59)	16.6 (6.51+10.1)

メラのテストデータ 17,344 枚に対する正解率を表 7 に示す。いずれの混雑度においても前方広角カメラの方が中央魚眼カメラより正解率が高かった。この原因としては、中央魚眼カメラでは通路上のオクルージョンが激しいこと、バスの前方・後方の奥側がゆがんで潰れていることなどが考えられる。

5.3 物体検出によるバス乗客のカウントの結果

物体検出によるバス乗客のカウントの結果として、

前方広角カメラと後方広角カメラの両方の画像に対して YOLOv3 で人物検出をおこなう。データセットとして、前方広角カメラと後方広角カメラで撮影した画像 110 枚を用いる。この 110 枚の画像は各シナリオから 5 枚ずつ抽出した、混雑度 0 の画像 30 枚、混雑度 1 の画像 55 枚、混雑度 2 の画像 25 枚から構成される。

各カメラ画像に対する性能評価を表 9 と表 10 に示す。後方広角カメラの画像では、適合率も再現率もある程度の性能を発揮しているのに対して、前方広角カメラの画像の性能は大きく劣っている。特に再現率が 0.509 と極めて低い値となっている。この原因として、座っている乗客で、立っている乗客やバスの手すりとのオクルージョンが大きいときに検出できないケースが多かったためである。

各混雑度ごとの混雑度推定結果を表 8 に示す。いずれのクラスに関しても正解よりも少ない人数を検出する結果となった。正解との差は後方広角カメラよりも前方広角カメラが大きかった。オクルージョン等により十分に

表 9. 前方広角カメラの人物検出性能

混雑度	適合率	再現率	F 値
0	0.956	0.522	0.653
1	0.780	0.496	0.598
2	0.673	0.523	0.580
全体	0.803	0.509	0.609

表 10. 後方広角カメラの人物検出性能

混雑度	適合率	再現率	F 値
0	0.983	0.861	0.911
1	0.914	0.854	0.882
2	0.892	0.744	0.809
全体	0.928	0.831	0.873

人物を検出できなかった結果であると言える。

各混雑度の画像に対して、物体検出をおこない、人物を検出した例を図 7 ～ 図 10 に示す。バス車内画像に YOLOv3 が人物として出力した矩形を重畳している。

前方広角カメラに関して、図 7 に示されるように、混雑度 0 では画像左側に座っている人物は矩形が抽出できているが、画像右側に座っている人物はバス天井に取り付けられている手すりとのオクルージョンがあり、検出できなかった。後部座席の人物は面積が小さく、また頭部しか写っていないため検出されなかった。図 8 に示されるように、混雑度 1 では画像の左側が依然として検出できているのに加えて、右側で立っている人物も検出できた。しかし、立っている人物と座っている人物とのオクルージョンがあると座っている人物を検出できないケースが多かった。図 9 に示されるように、混雑度 2 では立っていて体が十分写っている人物を検出できたが、座っている人物の頭部や体が一部しか映らないケースが多く、車内の人数を十分に把握することができなかった。

後方広角カメラに関して、図 10 に示されるように、少ない人数の時はもちろん、満席に近い状態でも十分に人物を検出することができた。カメラの角度と撮影する面積が適切であったため、十分な性能を発揮できたと考えられる。画像右奥で一部、乗客の位置によってはオクルージョンが起き、1,2 人検出できないケースがあった。

5.4 考察

前章で提案したバス車内混雑度推定手法に対して、バス車内画像を用いて精度の検証をおこなった。

画像分類による混雑状況の分類に関しては、全体で 0.99 を超える正解率は数字だけを見れば十分実用に足る性能である。しかし一方で、今回用いたデータセットが設定するタスクとして簡単であった。3, 4, 6, 7, 8, 24, 25, 38 名の状況に対して、混雑度 0 は 9 名以下、混雑度 1 が 10 名から 25 名まで、混雑度 2 は 26 名以上となっている 3 クラスを設定しているため、検証用のデータのクラスごとの差が明確であったため、高い正解率が得られた。そのため、乗客数に関してはより網羅的なデータを収集して、各クラスの境界付近の人数である 10 余人や 30 人前後の画像を含めて学習することが必要である。

また、物体検出によるバス乗客のカウントに関しては、YOLOv3 を用いた人物検出では実用に足る精度での推定が困難であり、その原因としてオクルージョンの問題が挙げられた。改善方法として、人物の全身を抽出する検出方法に加えて、頭部検出などを用いて人物の一部でも写っていれば乗客として検出するモデルを適用することが挙げられる。今回のデータ収集では、バス車内の状況をモニタリングするカメラとして想定していたが、乗降口を写すカメラで乗車数、降車数をカウントして、現在の乗客数を算出する手法を適用することも考えられる。

6 おわりに

路線バスにおいて効率的なバス運行やバス乗客の満足度を向上させるため、バス車内モニタリングシステムにはバス乗客 OD データやバス車内の混雑状況を推定する機能が求められている。このような背景を踏まえて、本稿では、バス乗客 OD データの推定手法とバス車内の混雑状況手法の提案と検証をおこなった。

バス乗客 OD データの推定手法に関しては、メトリッククラusteringを適用して、乗降口における個人識別をおこない、乗降人物・降車人物の同一推定をおこなう手法を提案した。提案したバス乗客 OD データの推定手法の実用性を検証するために、実際の路線バスを利用したバス停別乗降者データの収集をおこない、異なるデータセットで学習した推定器のバス停別乗降者数と乗車区間別乗降者数の推定精度を比較して、今後のデータ収集と改善による実用性を確認した。

バス車内の混雑状況手法においては、画像分類モデル ResNet50 を用いて混雑度を推定する手法と、物体検出モデル YOLOv3 を用いて車内の人数を推定する手法を提案した。提案したバス車内の混雑状況手法の実用性を検証するために実際の路線バスを利用したバス車内データの収集をおこない、それぞれの混雑状況推定手法の推定精度を確認した。画像分類による混雑度の推定に関しては、実験においては実用性を示したが、検証用のデータが不足しているため、今後のデータの収集の必要性を確認した。また、推定精度が低かった物体検出によるバス乗客のカウントに関しては、改善案を検討し、この手法の有効な状況を検討した。

謝辞

本研究の一部はノーステック財団「研究開発助成事業」の助成を受け実施されたものです。ここに感謝の意を表します。

参考文献

- [1] 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.
- [2] 福本雅之. 公共交通利用実態データの調べ方と活用方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://www.mlit.go.jp/chubu/aichi/kikaku/seminafukumotosennsei.pdf>, 2019.
- [3] Shuting He, Hao Luo, Weihua Chen, Miao Zhang, Yuqi Zhang, Fan Wang, Hao Li, and Wei Jiang. Multi-domain learning and identity mining for vehicle re-identification. In *Proc. CVPR Workshops*, 2020.
- [4] Hao Wang, Yitong Wang, Zheng Zhou, Xing Ji, Dihong Gong, Jingchao Zhou, Zhifeng Li, and Wei Liu. Cosface: Large margin cosine loss for deep

face recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

- [5] Nicolai Wojke and Alex Bewley. Deep cosine metric learning for person re-identification. In *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 748–756. IEEE, 2018.
- [6] Alexander Hermans, Lucas Beyer, and Bastian Leibe. In defense of the triplet loss for person re-identification. *CoRR*, Vol. abs/1703.07737, , 2017.
- [7] Zhun Zhong, Liang Zheng, Guoliang Kang, Shaozi Li, and Yi Yang. Random erasing data augmentation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020.
- [8] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the inception architecture for computer vision. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2818–2826, 2016.
- [9] Yifan Sun, Liang Zheng, Yi Yang, Qi Tian, and Shengjin Wang. Beyond part models: Person retrieval with refined part pooling (and a strong convolutional baseline). In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.
- [10] Yandong Wen, Kaipeng Zhang, Zhifeng Li, and Yu Qiao. A discriminative feature learning approach for deep face recognition. In Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling, editors, *Computer Vision – ECCV 2016*, pp. 499–515, 2016.
- [11] 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スンシステム (MBL) , 第93巻, pp. 1–8, 2019.
- [12] Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollar. Microsoft coco: Common objects in context, 2015.

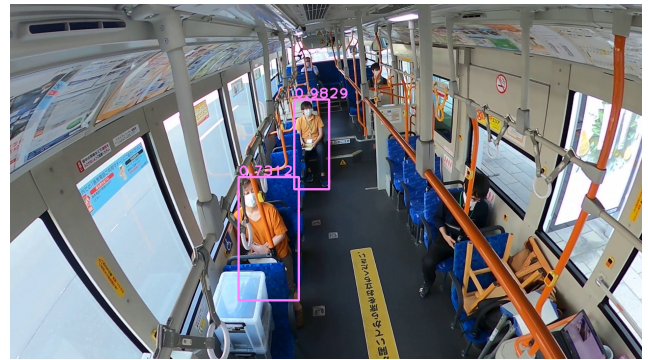


図 7. 人物検出した前方広角カメラの画像例 (乗客数 6 人の場合)

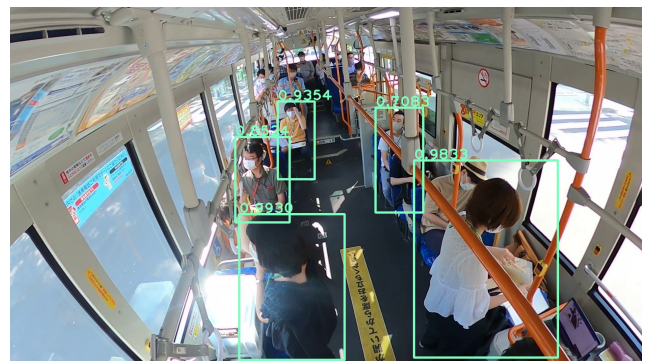


図 8. 人物検出した前方広角カメラの画像例 (乗客数 24 人の場合)

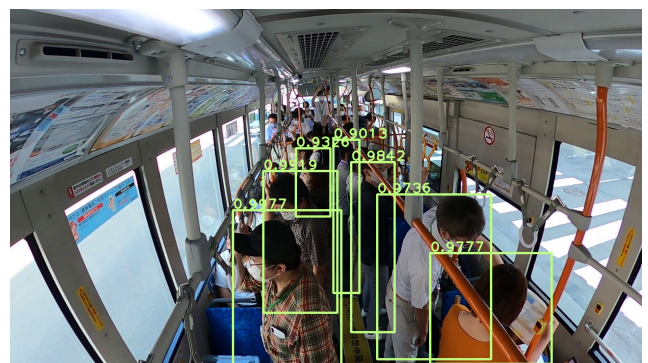


図 9. 人物検出した前方広角カメラの画像例 (乗客数 38 人の場合)

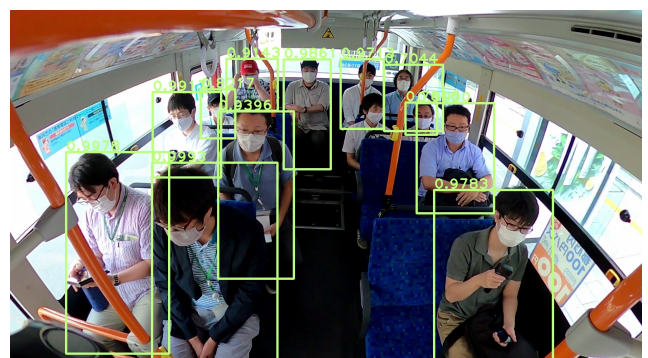


図 10. 人物検出した後方広角カメラの画像例