

# 車載カメラの画像を用いた車両数検出手法の検討

菅原 大夢<sup>†</sup> 横山 大作<sup>†</sup>

<sup>†</sup> 明治大学理工学部情報科学科 〒 214-8571 神奈川県川崎市多摩区東三田 1-1-1

E-mail: <sup>†</sup>{ee67088,dyokoyama}@meiji.ac.jp

あらまし 道路の交通量や渋滞状況の情報は、ドライバーや道路管理者にとって重要である。これらの情報は、道路上に設置されているカメラや車両感知器から得ることができ、これらの情報を用いて道路の混雑度などの交通指標を算出する。しかし、これらの設備は主要道路に限られているなどの問題がある。本論文では、カーナビ等の簡易な車載カメラ画像に写っている車両数が具体的な交通量の指標になり得ると考え、車載カメラ画像中の車両数を検出手法を提案する。本手法では、車載カメラの画像にセマンティックセグメンテーションを行い、車両同士の境目には視覚的な形状特徴が表れることを利用して、車両領域の  $x$  座標ごとの長さを確認して検出する。実験には CityScapes データセットの test データ 100 枚を用いて、人間の目視で設定した正解台数と提案手法による予測台数の差異を求めた。結果は、二乗誤差平均が 1.82、決定係数が 0.45 となった。

キーワード 交通量、車載カメラ、セマンティックセグメンテーション

## 1 はじめに

近年、日本各地の地方都市で、人口の集中が起こっている。[1] それに伴い、地方における公共交通機関の脆弱さが露呈している。その影響で、日本の多くの地方中核都市では、自動車の生活必需品化が進み、自動車の保有台数が増加している。この影響により地方中核都市では、東京・大阪・名古屋の 3 都市からなる、3 大都市と呼ばれる地域ほど深刻ではないものの、各地で道路混雑が問題となっている。また、そのような地域に居住している交通弱者についての対応も問題となっている。このような問題がある為、各地で公共交通機関を活性化させる交通対策が試みられている。[2]

車両の渋滞情報や事故情報を得るための手段の一つとして、主要な道路に設置されている定点カメラや車両感知器がある。特に、車両の交通量を測定する目的でトラフィックカウンターと呼ばれるものが用いられている。これらの設備は予め決められた場所を監視し、車両の通行量、速度などを記録する。道路交通において、これらの情報を用いて、道路がどの程度混雑しているかを判断する為によく用いられる代表的な交通指標として、混雑度がある。混雑度は自動車通行可能な道路の最大容量に対する実際の通行量の比となっており、道路における混雑を数値化したものとなっている。[3] しかし、これらの設備は都市部や高速道路などの主要道路に設置が限られており、それ以外の道路などのデータが不十分であると言える。また、設備間で何か局所的なことが起きているとき、どこで起きているかわからない、画像での車両の検出において解像度が低い、といった問題がある。これ以外の問題点として、これらの設備に対しての定期的なメンテナンスコストがかかることなどが挙げられる。

近年、カメラの小型化や低価格化によって、ドライブレコーダーを搭載する車両が増え、車載カメラを使用した自動車運転

支援システムが開発されてきた。また、長距離配送などに用いるトラックには、速度・位置情報・エンジン回転数・急加速・急減速などの情報を記録する運行記録計の取り付けが義務化されている。多くのドライブレコーダーには、位置情報による追跡機能が搭載されている。混雑情報や事故情報の取得をこういった車載カメラで実現できれば、これまで捕捉できなかった特定の地域における道路上の混雑の状況などを発見できると考えられる。また、こうした試みによって路面や道路周辺の変化を発見する為の広範囲における長時間の通常時のデータを取得することにも繋がり、これらの地域における交通リスクや特定の日時との差異を発見することができる。

本論文では、道路の交通量をより具体的に知ることを目的として、車載カメラのデータを用いて画像中の車両数を検出手法を提案する。具体的な検出手法は各節で説明するが、概要としては、画像中の物体ごとにセグメンテーションを行い、車両の上部の輪郭の視覚的な形状特徴とその部分の  $x$  座標ごとの車両領域の長さの増減に注目し、画像中の車両数をカウントする。

実験は CityScapes データセットの test データを用いて行い、予測データと正解データを比較して提案手法を評価した。結果、二乗誤差平均が 1.82、決定係数が 0.45 となった。このことから、実測値と予測値には約 2 台の差が存在し、実測値と予測値には弱い相関があることがわかった。

以下第 2. 節で関連研究、第 3. 節で車両の輪郭と車両領域の面積を用いた車両数検出手法の提案、第 4. 節で CityScapes データセットを用いた実験、第 5. 節で考察、第 6. 節でまとめと今後の課題について述べる。

## 2 関連研究

混雑度における関連研究として、小塚らは [4] 平均旅行速度と混雑度に相関があることを示し、また、この相関は一定間隔

ごとの信号密度によって、薄れる傾向があるとしている。

車両検出における関連研究として、藤岡ら [5] は車両の水平エッジをヒストグラムにした際に山なりになることを利用して車載カメラ画像から車両を検出している。Ronan ら [6] は夜間において、車両をテールランプとブレーキランプの色情報で検出している。

車両数検出における関連研究として、Hyeok ら [7] はビデオを用いてフレーム間の画像の減算によって前景と後景を抜き出すことで車両と車両数の検出を行なっている。また、Erhan ら [8] はビデオを用いて、フレーム間の後景の減算とカメラからの推定距離に応じてバウンディングボックスのサイズを適応させることで車両を検出している。

### 3 車両数検出手法

本研究では、車両の上部の輪郭に視覚的な形状特徴が表れることを利用して、そのような形状特徴の範囲内における  $x$  座標ごとの車両領域の長さの増減を確認することで、車載カメラ画像中の車両数を検出する。処理は、大まかに以下の流れで行う。

- 画像にセマンティックセグメンテーションを行い、車両の領域の最上部の輪郭と、 $x$  座標ごとの長さを取り出し、それぞれの配列に格納する。
- 決められた大きさのウィンドウを  $x$  座標に沿ってウィンドウの右端が配列の右端に到達するまで 1 ずつスライドさせる。
- 1 回のスライドごとにウィンドウ範囲内の配列を参照して特定の操作をする。

ここからは、それぞれの工程について述べる。

#### 3.1 データセットの詳細

本研究では、CityScapes データセット [9] を用いる。このデータセットは、50 都市における違う季節毎の車載カメラ画像で構成されている。それぞれの画像サイズは  $1024 \times 512$  で、画像はピクセル単位で領域にラベル付けされており、画像中のそれぞれの物体には 8 つのグループに属する 30 個のクラスのラベルが付いている。(例: 人間は human グループ、車両は vehicle グループに属する)

#### 3.2 画像のセグメンテーション

画像中の車両数を検出するにあたって、画像中の車両とそれ以外の部分とを分ける為にセマンティックセグメンテーションを用いる。セマンティックセグメンテーションは、画像中の各ピクセルごとに物体をカテゴリ分類する手法である。

本研究では、モデルに ESPNet [10] を用いる。ESPNet はリアルタイム性を求めた軽量な CNN のモデルで、低メモリ、省電力、高速といった特徴がある。本研究で使用するモデルは CityScapes データセットの train データ 2975 枚で学習済みで、入力画像を車、人間、街路樹など 20 種類に分類する。

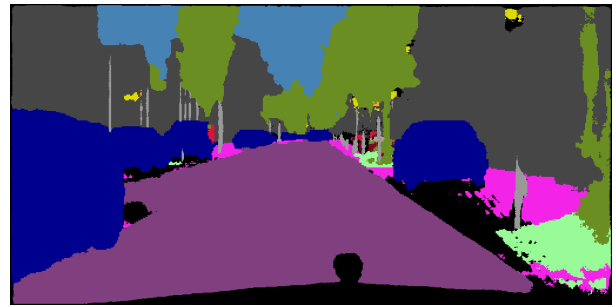


図 1 セマンティックセグメンテーションを行なった画像

#### 3.3 視覚的な形状特徴の検出

本工程では、車両領域の最上部の輪郭から視覚的な形状特徴を検出する。この形状特徴には大まかに分けて以下の 2 つのパターンがある。

- 車両とそれ以外の物体との境界
- 車両同士の境界

車両とそれ以外の物体との境界を検出する場合には、セグメンテーション後の画像において、車両が存在しない領域には車両の上部輪郭が存在しないことを利用して、車両の上部輪郭を取り出した際に、ある程度以上の落差が生じ、車両領域が一定以上の長さで存在しなかった場合に、その部分を車両とそれ以外の物体との境界とみなす。車両同士の境界を検出する場合には、車両領域の輪郭から発見できる視覚的な特徴を検出し、その部分を車両同士の境界として仮決めし、次の工程で確定させる。

図 1 から上部輪郭と  $x$  座標ごとの車両領域の長さを配列に格納し、可視化すると図 2 のようになる。ここからは、図 2 の円で囲まれた部分について、ウィンドウ内での操作を述べる。

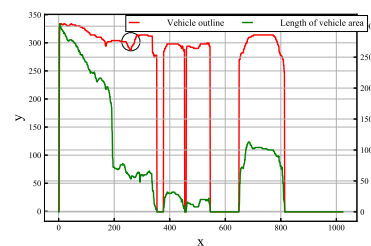


図 2 車両領域の上部の輪郭と  $x$  座標ごとの車両領域の長さ。

図 3 は、図 2 の円で囲まれた部分の拡大図である。この画像中の灰色に着色された範囲がウィンドウとする。形状特徴の検出は以下の処理によって行われる。

- ウィンドウの中で、左端の点ではない部分で極小値を発見する。
- 極小値よりも右の範囲から極大値を発見する。ここでの極大値は輪郭の傾きが正から負に変わる部分のことを示す。
- 極小値から極大値までの  $x$  座標での長さが一定以上だった場合、ウィンドウの左端から極大値までの範囲を形状特徴としてマークする。

以上の処理が完了し、この部分が形状特徴としてマークされた場合、この形状特徴の右端から配列の走査を再開する。画像

全体について調べ終わったら、次の工程に移る。

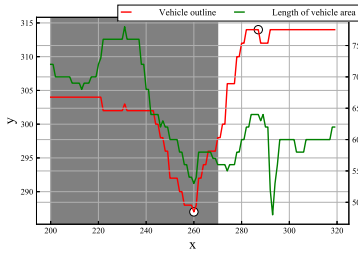


図 3 図 2 の円で囲った部分の拡大図。左から 1 つ目の円が極小値、2 つ目の円が極大値を表す。

### 3.4 形状特徴の範囲における x 座標ごとの車両領域の長さの増減の利用

セマンティックセグメンテーションを行った画像は、同手法による誤差で、特定の微小な部分が欠けていたり盛り上がっていることがある。第 3.3 節までの手法では、図 4 のような画像において、画面中央上部の凹凸の激しい部分で誤検出が発生する。このような誤検出を防ぐために、第 3.3 節で検出したそれぞれの形状特徴の範囲における車両領域の x 座標ごとの長さの増減を確認する。



図 4 誤検出が発生しやすいセマンティックセグメンテーション後の画像に元の画像を透過させた画像。

セマンティックセグメンテーション後の画像において、車両領域の上部の輪郭に車両同士の境界がある時、下部の輪郭も同様に境界があり、窪みのような形状特徴が見られる場合が多い。つまり、上部では下に凸な窪みがあり、下部では上に凸な窪みが存在している状態になる。この時、車両の境界部分の x 座標ごとの車両領域の長さに大きな増減が発生する。これを検出する為に、この範囲の x 座標ごとの車両領域の長さの最大値と最小値の差を求め、この値がこの範囲における中央値の 2 倍以上ならば、この部分を車両同士の境界としてカウントする。

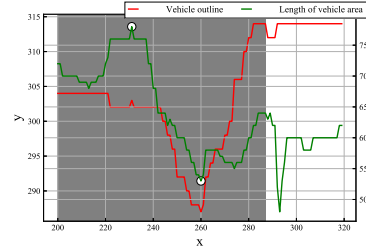


図 5 図 3 に形状特徴の範囲をマークした図。x 座標ごとの車両領域の長さの最小値と最大値の部分の円で囲ってある。

## 4 実験

提案手法によって十分な車両数の検出が可能であるか、また、その予測結果が妥当なものであるかを検証する為、以下の実験設定で車両数検出実験を行った。

### 4.1 実験設定

#### 4.1.1 実験に用いるデータセット

本実験には、CityScapes データセットの test データを用いる。この中から昼間の晴天時の画像のうち車両が 1 台以上明瞭に写っているものを 100 枚取り出した。それぞれの画像に正解台数を目視で設定しており、台数の合計は 499 台となった。

#### 4.1.2 評価指標

本実験では、正解データと提案手法による予測結果との差異を求める。提案手法の評価には以下の 2 つの指標を用いる。画像  $i$  に対する本手法の予測台数を  $\hat{y}_i$ 、正解台数を  $y_i$ 、正解台数の平均値を  $\bar{y}$ 、画像データの個数を  $N$  として、

二乗誤差平均 ( $RMSE$ )

$$\sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (1)$$

決定係数 ( $R^2$ )

$$1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (2)$$

以上の 2 つを用いる。

$RMSE$  は実測値と予測値との距離の絶対値である。これは値が小さければ実測値と予測値の距離が近いことを表す。よって、値が小さければ予測が正確であることを意味する。

$R^2$  は相関係数の 2 乗で、実測値が大きい時には大きい値を、小さい時には小さい値を予測できたかを表し、値が大きければ大きいほど予測が正確であることを意味する。

この 2 つの指標により、実測値と予測値にはどの程度の差異が生じているのか、予測値は妥当な値であるかを評価する。

#### 4.1.3 実験手法

実験は、手法の詳細を変えて 2 種類行う。

①... 視覚的な形状特徴から検出

②... 特徴範囲内の x 座標ごとの車両領域の長さの確認

- 手法 (1):①のみ

- 手法 (2):①+②

これら 2 つに分けて、提案手法の性能を評価する。

## 4.2 実験結果

実験結果を表 1、正解台数ごとに何枚の画像があるかを表 2、予測台数の正解台数ごとの分布を図 6 に示す。箱ひげ図の分布より、写っている車両数が 9 台以上の時、ほぼ全ての予測台数と正解台数に大きな差異が生じていることがわかる。特に、10、13 台以上の場合においては予測結果が正解台数を大幅に下回っている。また、写っている車両数が 1~4 台までにおいては正解台数周辺に予測データが集中し、精度よく検出できていることがわかる。これらの特徴は手法 (1)、手法 (2) 共に見られた。

表 1 から、手法 (1) と手法 (2) を比較すると、検出した車両数の合計が減少していることから、セマンティックセグメンテーションの誤差による誤検出を防ぐことができていると考えられる。また、 $RMSE$ 、 $R^2$  共に向上が見られる。

表 1 各手法による車両数検出結果

	①のみ	①+②
全車両数	555	481
$RMSE$	2.09	1.82
$R^2$	0.29	0.45

表 2 正解台数ごとの画像数

正解車両数 (台)	画像枚数 (枚)
1	7
2	8
3	14
4	15
5	16
6	18
7	10
8	5
9	2
10	2
11	0
12	1
13	2

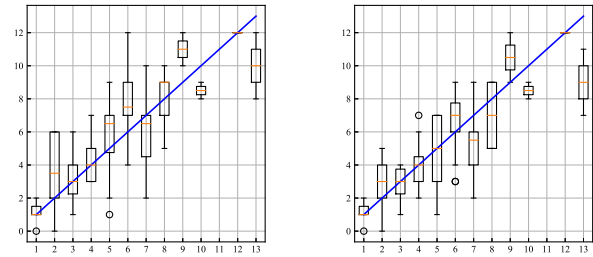


図 6 横軸に正解台数、縦軸に予測台数を表す。

左：①のみによる予測台数の分布

右：①+②による予測台数の分布

## 5 考 察

本研究では、画像中の車両領域の最上部の輪郭の視覚的な形状特徴によって車両数を検出する手法を提案した。また、セマンティックセグメンテーションの誤差による誤検出を防ぐために、形状特徴の範囲における x 座標ごとの車両領域の長さの増減を確認することで予測精度を上げを試みた。実験結果は、各実験手法を比較すると、検出した全車両数が大幅に減少し、二乗誤差平均、決定係数共に向上したことから、予測精度は上がったと言える。(表 1) 全予測車両数の減少から、x 座標ごとの車両領域の長さの増減を調べることでセマンティックセグメンテーションによる誤検出を減らすことができたと推測できる。また、正解台数が 5 台以上の画像において、正解台数が 5 台より少ない画像と比べると誤差が大きくなっている。画像中に写っている車両がある程度以上多い画像や、画像を撮影している車両から距離が遠い車両において、セマンティックセグメンテーションを行った後の形状特徴や x 座標ごとの車両領域の長さの変化が微小なものとなる。このような場合に検出できていないのがこれらの誤差の原因と考えられる。

実測値が 10 台以上の画像では、予測値がほぼ全て実測値を下回る結果となっている。本手法では、車両の領域の最上部の輪郭を用いて視覚的な形状特徴を検出しているが、例えば図 7 のような画像では、画像中央の車両の最上部の輪郭は、より上部に存在する車両の輪郭で隠れてしまう。よってこの車両については検出できなくなる。このような状況が正解台数が 10 台以上のほぼ全ての画像において発生したことにより、全体的に予測値が減少しているのだと考えられる。

また、本研究ではまだ試していないが、車載カメラ画像中に写っている車両数が 0 台の時、本手法では、2 種類の出力が考えられる。1 つ目は、セマンティックセグメンテーションによって領域分けされた中に車両領域が存在しない場合で、この場合は本手法は 0 台であると検出する。2 つ目は、セマンティックセグメンテーションによって領域分けされた中に車両領域が存在してしまった場合で、この場合は少なくとも 1 台以上の誤検出が生じると考えられる。

本研究の目的は道路上の具体的な混雑状況を知る為の指標として、車載カメラ画像中の車両数を用いた。結果としては、実測値との差異が約 2 台分生じているが、表 1 の決定係数と図 6



の右側の箱ひげ図の分布より、実測値の車両の増加に対応して予測値もある程度増加していることがわかる。このことから、本手法によって、特定の地域における通常時の混雑状況を把握することが可能であると考えられる。また、実験結果から、混雑時における車両数の増加に対応することについても可能である為、通常時と混雑時の混雑状況の比較が可能であると考えられる。よって、本研究における車両数の検出は、具体的な混雑状況を知るための指標として、実験結果における予測値の分布から、1~10 台までの検出においては有効であると考えられる。



図 7 セグメンテーション結果を test データに重ねた画像。正面奥の車両が最奥の車両群で覆われている。

## 6 まとめと今後の課題

本研究では、実際の混雑状況をより具体的に知る為の指標として、車載カメラの画像中から車両数を自動的に求めて利用する方法を提案した。車両数を得る為の手法として、車載カメラの画像中の車両の輪郭における視覚的な特徴と特徴付近の x 座標ごとの車両領域の長さから車両数を検出することを提案した。実験の結果、実測値と予測値との二乗誤差平均が 1.82、決定係数が 0.45 となった。また、実測値の台数の合計が 499 台で、予測値の台数の合計は 481 台となった。4 台以下の場合では精度良く検出できたが、それ以上の台数では比較的大きく、10 台以上については、実測値よりも予測値が小さくなることが判明した。現状、本研究に残されている今後の課題としては、以下の 4 つがある。

- 車両の前に障害物があり、車両が 2 つに別れて写っている画像などでは誤検出が多くなる。
- ある特定の車両の上部輪郭が他の車両領域によって完全に覆われた場合、検出できなくなる。
- 本研究で使った ESPNet の学習済みモデルは外国で撮影された車載カメラ画像のデータセットを用いて学習している為、日本での使用には適さない。本手法を日本で実際に試す場合は、日本の車載カメラ画像を用いて再学習をする必要がある。
- 形状特徴の範囲の x 座標ごとの車両領域の長さの増減を確認する際の閾値には、同範囲内の値の中央値を 2 倍したものをを用いている。この閾値に関してはより精度が上がる値があると考えられる。

従来の物体認識手法では、検知する物体が限りなく近い場合であったり、多くなった場合などの複雑な状況下では、認識す

べき物体をうまく認識できない。本研究ではセマンティックセグメンテーションを用いたが、車両領域の面積などを定量的に見ることで、道路が空いている状態なのか混雑している状態なのか、などの状況を認識できると考えられる。しかし、上記に挙げた問題点の通り、本研究では未だ限定的な状況における車両数の検出となっているため、今後は精度を上げるだけではなく、混雑そのものを把握するシステムを開発することを目標として、検討を進めたい。

## 謝 辞

本研究の一部は国土交通省「道路施策の質の向上に資する技術研究開発」の研究テーマ「交通流理論と AI 学習による非日常の発見とアラート発信」によって実施された。

## 文 献

- [1] 岡田豊, “今後都心集中が地方圏でも加速”, 2019
- [2] 青木亮, “地方中核都市における公共交通対策”, 三田商学研究第 43 巻第 3 号, 2000
- [3] 日本道路協会:「道路の交通容量」, 1984.9
- [4] 小塚清, 上坂克己, 青木亮二, “道路交通センサデータからみた混雑度と旅行速度の関係に関する一考察”, 2012, 11
- [5] 藤岡明紘, 中村靖, “水平エッジヒストグラムマッチングを用いた単眼車載カメラ画像中の前方車両検出”, 広島工業大学紀要研究編 第 46 巻, 2012, 181 – 186
- [6] Ronan O’Malley, Martin Glavin, Edward Jones, “Vehicle Detection at Night Based on Tail-Light Detection”, 1st International ICST Symposium on Vehicular Computing System.
- [7] Hyeok Jang, In-Su Won, and Dong-Seok Jeong, “Automatic Vehicle Detection and Counting Algorithm”, IJCSNS International Journal of Computer Science and Network Security, VOL. 14 No. 9, pp. 99-102, September 2014.
- [8] E. Bas, A. M. Tekalp and F. S. Salman, “Automatic Vehicle Counting from Video for Traffic Flow Analysis”, 2007 IEEE Intelligent Symposium, Istanbul, 2007, pp. 392-397.
- [9] Cordts et al, “The cityscapes dataset for semantic urban scene understanding” CVPR, 2016
- [10] Sachin Mehta<sup>1</sup>, Mohammad Rastegari<sup>2</sup>, Anat Caspi<sup>1</sup>, Linda Shapiro<sup>1</sup>, and Hannaneh Hajishirzi<sup>1</sup>, “ESPNet: Efficient Spatial Pyramid of Dilated Convolutions for Semantic Segmentation”, European Conference on Computer Vision(ECCV), 2018, pp. 552-568.