

エッジデバイスによるリアルタイム路上駐停車判定システム

松田 明大^{1,a)} 松井 智一¹ 松田 裕貴^{1,2} 諏訪 博彦^{1,2} 安本 慶一^{1,2}

概要: 近年、禁止区域での路上駐停車が社会問題の一つになっている。特に大都市圏や観光地では、多くの路上駐停車が存在する。また、路上駐停車は渋滞や事故を引き起こす要因にもなり得るため、リアルタイムな検出による迅速な対応が求められる。先行研究では、主に定点カメラが用いられていたが、検出エリアが限定されていることが大きな問題としてある。そこで本研究では、幅広く普及しているドライブレコーダ動画を用いて学習モデルを作成し、リアルタイムに路上駐停車の認識を行うシステムを提案する。エッジデバイスとして車両に搭載可能なデバイスである Jetson TX2 を用いて、車両の認識・処理を行い、路上駐停車の判定を行う。認識・処理速度を確認したところ、YOLOv3-tiny モデルにおいて平均約 7[fps] であった。この結果からエッジデバイスにおいて、リアルタイムな路上駐停車判定が可能であると確認された。

1. はじめに

近年、禁止区域での路上駐停車が社会問題の一つとなっている。2019 年に警察庁交通局により行われた路上駐停車に関する調査では、東京都特別区の禁止区域における瞬間路上駐停車台数は約 5 万 2700 台であり、多くの違法駐停車車両が存在する [1]。違法な路上駐停車はそれ自体が交通上の問題でありながら、渋滞や人身事故といった様々な問題を引き起こす要因にもなり得る。

これまで、路上駐停車に起因する問題に対応するため、駐車規制区間の延長や時間制限区間の実施の検討が行われてきた。しかし、これらは根本的な解決とはならず、既存の問題が複雑化することも考えられる。さらに、国や行政などの大規模な協力が必要であり、莫大なコストと時間が必要である。関連研究としては路上駐停車に対し、街中に設置されている定点カメラを複数台用いて、短期的な駐停車車両を検出するシステムが開発されている [2]。このシステムでは、街の中に設置されている定点カメラを複数台用いて路上駐停車の検出を行なっている。しかし、定点カメラでは画角に収まる範囲の限定的な検出であることや、設置場所が限られているため、特定のエリアのみの検出であることが問題として挙げられる。

そこで本研究では、一般車両への搭載が増加しているドライブレコーダの動画を用いて、定点カメラよりも効率的

でリアルタイムな路上駐停車を判定するシステムの実現を目指す [3]。ドライブレコーダは街のあらゆる状況を録画しており、情報量が圧倒的に多い。また、エリアの制限を受けないため、従来の手法よりも効率的に情報を収集することが可能である。一方で、収集した動画を全てクラウドにアップロードし解析を行うのは、通信網への負荷や通信コスト、リアルタイム性の観点から問題である。そこで、通信負荷や通信コストを削減するために、車両に搭載されたエッジデバイスで路上駐停車の判定をリアルタイムで行い、結果のみをクラウドにアップロードすることを目的とする。本論文では、ドライブレコーダの動画を用いて学習したモデルを用いて、車両に搭載したエッジデバイスによって認識・処理を行う路上駐停車判定手法を提案する。

本論文の構成は以下の通りである。第 2 章では、既存の路上駐停車の判定に関する研究を取り上げ、既存手法における課題を整理する。第 3 章では、リアルタイムに路上駐停車を判定するためのシステムに必要な要件について述べる。第 4 章では、提案するエッジデバイスによる路上駐停車判定手法について述べる。第 5 章では、エッジデバイス上における認識・処理速度と精度について述べる。最後に第 6 章では、本論文の結論及び今後の展望について述べる。

2. 関連研究

本章では、路上駐停車における関連研究と既存手法について述べる。路上駐停車に関する研究は多くなされているが、路上駐停車を判定する研究は少ない [4][5][6][7]。本論文では、路上駐停車の判定に関する関連研究について述べる。

¹ 奈良先端科学技術大学院大学

Nara Institute of Science and Technology

² 理化学研究所 革新知能統合研究センター

RIKEN, Center for Advanced Intelligence Project

a) matsuda.akihiko.lr2@is.naist.jp

Real-time Street Parking Detection System using the Proposed Method

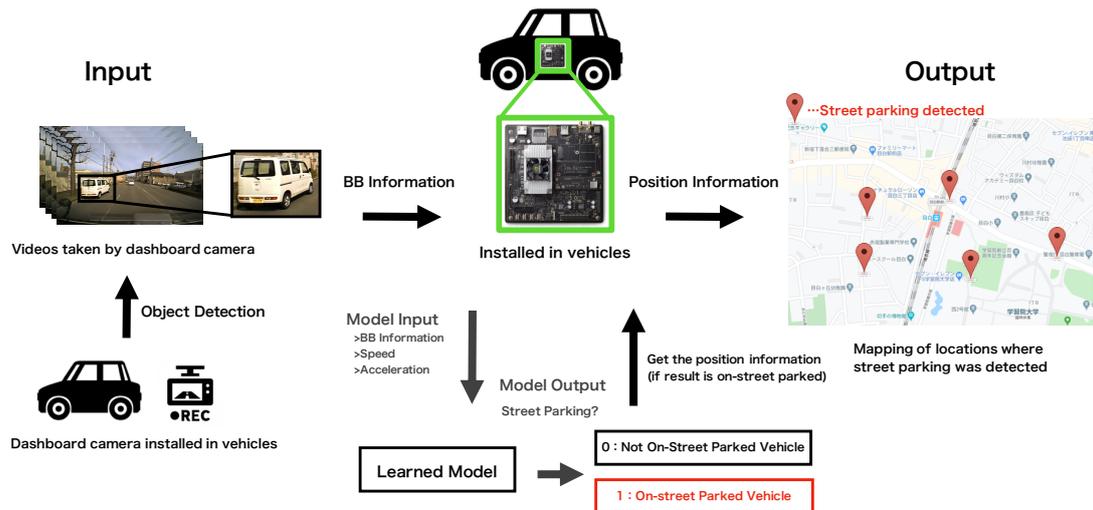


図 1 提案システムの全体構成

久徳らは、過去の車載カメラ映像との時空間差分による不特定障害物の検出手法について提案している [8]。現在と過去の対応づけられたフレームの中で、撮影視点の違いによるずれを考慮し、検出領域を平面であると仮定した道路面に限定している。差分の指標として、特徴量は輝度変化に頑健な RGB ベクトル間の角度が用いられることにより、検出精度を向上させた。しかし、道路面の位置合わせにおける手間や精度の問題、限定的な検出領域や走行車線などの問題がある。また、過去の車載カメラ映像との差分を検証する際に、過去映像に路上駐停車をはじめとする障害物が存在する場合には、差分を検証することが難しい。

Wen らは、Mobile Laser Scanning (MLS) の点群から車両認識や位置検出、分類を行うシステムを提案している [9]。レーザスキャンの点群から、車両の認識や位置検出、分類を行う。点群によって認識対象をセグメント化し、各対象に車両認識モデルを適合させる。取得した位置や向きなどの情報を、異なる時刻に取得されたデータと照らし合わせ、特徴の一致度を持続時間とともに推定する。車両の認識と分類、変化の検出において高い精度であるが、MLS の届く範囲での認識であるため路上駐停車の側道を通り、スキャンを行う必要がある。また、異なる時刻のデータが必要であるため、リアルタイムな認識が難しい。

Xuemei らは、SSD (Single Shot MultiBox Detector) アルゴリズムをベースとした違法駐車検出システムを提案している [10]。ある一定の角度から撮影された固定カメラの映像に対して、SSD を用いて車両の検出を行い、ROI (Region Of Interest) 内の車両をトラッキングし、停車した対象のカウントを行なっている。しかし、固定カメラの映像を対象とした手法あり、エリアが限定されている。また、停車した車両全てに対してカウントが開始されるため、検出車両が信号待ちなのか判定することが難しい。

本研究の位置付けとして、本節で述べた既存研究の問題点を解決するため、路上駐停車をエッジデバイスを用いてリアルタイムに効率的かつ高精度で判定するシステムの構築を目指す。

3. システム要件

本章では、エッジデバイスによるリアルタイム路上駐停車判定システムの実現のための必要要件について述べる。提案システムに求められる主な要件 3 つを以下に示す。

- 要件 1 車両に搭載可能な簡易エッジデバイスであること
- 要件 2 認識から駐車位置取得までのリアルタイム性を担保できること
- 要件 3 路上駐停車を広範囲かつ高精度で認識可能であること

要件 1 について、車両に搭載しオンデバイスによる処理を行うため、一般車両に搭載可能な簡易的なデバイスであることが求められる。既存の研究では、大掛かりな装置を車両に搭載することが必要であるなどの問題があった。本論文では要件 1 を満たすため、一般に販売される車両に搭載可能な NVIDIA 社製のデバイスである Jetson TX2 を用いる。そして、カメラやセンサモジュールを接続することで、撮影や加速度、GPS 情報を取得可能なドライブレコーダを模倣した簡易的なデバイスを作成する。

要件 2 について、路上駐停車に起因する問題を未然に防ぐためにリアルタイム認識し、迅速に駐車位置を把握することが求められる。既存の研究では、あらかじめ撮影しておいた画像との比較など、リアルタイム性に問題があった。本論文では要件 2 を満たすため、オンデバイス認識・処理を行う。ドライブレコーダを模したデバイスを用いて撮影を行いながら軽量な認識モデルと学習済み判定モデルによって、路上駐停車位置の結果のみを取得する。GPU を

表 1 デバイスの比較表

item	Jetson TX2	Jetson AGX Xavier	Raspberry Pi 4 model B
Size	87×50mm, 170×170mm(base board)	105×105×65mm	85(w)×56(d)×17(h)mm
Weight	85grams(core)	140grams	47grams
OS	Ubuntu	Ubuntu	Raspbian
CPU	Dual-Core NVIDIA Denver 2 64-Bit CPU Quad-Core ARM Cortex-A57 MPCore	8-Core ARM v8.2 64-Bit CPU 8 MB L2 + 4 MB L3	Quad-core ARM Cortex-A72
GPU	256-core NVIDIA Pascal™ GPU with 256 NVIDIA CUDA cores	512-Core Volta GPU with Tensor Cores	VideoCore IV
Memory	8GB 128-bit LPDDR4 Memory ³	16 GB 256-Bit LPDDR4x	4GB LPDDR4
Connector	USB, HDMI, SD, Ethernet	USB, HDMI, SD, Etherne	USB, HDMI, microSD
Bluetooth	Bluetooth	—————	Bluetooth 5.0 (BLE)
Wireless LAN	IEEE802.11ac	—————	IEEE802.11b IEEE802.11g IEEE802.11n IEEE802.11ac
Power	5.5-19.6V	9-20V	DC 5V
Consumption	7.5W 15W	10W 15W 30W	3.4W 7.6W

搭載したデバイスにより処理速度を向上させ、画像データをテキストデータに変換し、判定モデルに入力することで処理負荷を軽減する。また、路上駐停車と判定された車両の位置情報結果のみの取得によっても負荷の軽減を行う。

要件 3 について、搭載車両が移動する広範囲で複雑な状況下においても、高精度で車両を判定することが求められる。既存の研究では、限定的な範囲での認識であることや認識可能な車両が限定されているなどの問題があった。本論文では要件 3 を満たすため、タクシー会社から提供される多くのドライブレコーダ動画を解析し、様々な状況における車両データを学習した判定モデルを用いることで、高精度な判定を行う。また、今後は物体認識モデルを転移学習し、認識精度を向上させる。

4. 路上駐停車判定システム

本章では、提案するエッジデバイスによる路上駐停車判定システムの全体的な構成および、それぞれの詳細について述べる。

4.1 システム構成

提案システムの全体構成を図 1 に示す。デバイスにカメラやセンサモジュールを取り付け、撮影や様々な情報の取得を行う。そして、物体認識モデルを用いてリアルタイムに車両の認識を行う。認識の際に車両に付与される Bounding Box(以下, BB) の座標情報を、各車両ごとにテキストデータとして時系列的に取得する。また、撮影側であるデバイスの速度や加速度を BB 情報と時系列的に結合する。その情報を入力として、路上駐停車判定の学習済みモデルにより判定を行う。学習済みモデルにより、車両が路上駐停車であると判定された場合、その位置情報を取得し、マッピングのために情報を保持する。また、その際のフレームを切り出し、画像とともに路上駐停車を可視化する。

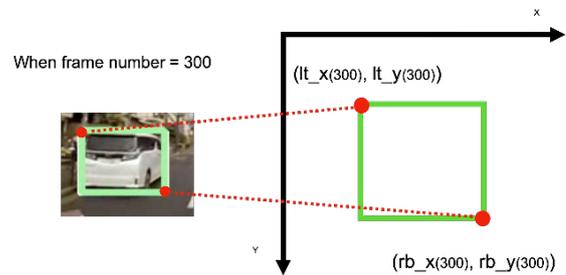


図 2 取得する座標値

る。以上の認識から処理までのフローを全てエッジデバイスで行う。

4.2 使用するデバイス

デバイスには、NVIDIA 社製の組み込みシングルボードコンピュータである Jetson TX2^{*1}を用いる。Jetson TX2 と他デバイスとの比較として、それぞれスペックを表 1 に示す。Jetson TX2 は AI コンピューティングデバイスとして、GPU を搭載したエッジ IoT デバイスでありながら、持ち運び可能なサイズと重さであることから車両に搭載するデバイスとして選定した。また、jetson tx2 専用の Development Tool Kit を使用することで、USB や LAN, Bluetooth などのコネクタに対応可能であるため、様々なモジュールを取り付けることを想定した車載デバイスにおいて、拡張性の面でも選定理由として挙げられる。

4.3 路上駐停車の判定手法

路上駐停車の判定には、物体認識モデルを用いて BB の座標値とその際の撮影車側 (デバイス) の加速度等の情報を取得した後に、学習済みモデルにより判定を行う [11]。次節から学習済みモデルの作成について述べる。

^{*1} Jetson : <https://www.nvidia.com/ja-jp/autonomous-machines/embedded-systems/>

表 2 使用した特徴量一覧

Domain	Features name
Frame	Number
Bounding Box	Top left x-coordinate
	Top left y-coordinate
	Bottom right x-coordinate
	Bottom right y-coordinate
	Width
	Height
Recording car	Speed
	Acceleration(x,y,z direction)

表 3 認識・処理速度結果

Model \ Device	PC	Jetson TX2
YOLOv3	1.73[fps]*	1.54[fps]*
	1.79[fps]**	1.67[fps]**
	1.89[fps]***	1.90[fps]***
YOLOv3-tiny	8.14[fps]*	3.84[fps]*
	9.10[fps]**	4.89[fps]**
	10.89[fps]***	6.99[fps]***
YOLOv4	0.63[fps]*	0.81[fps]*
	0.63[fps]**	0.81[fps]**
	0.66[fps]***	0.90[fps]***

* 1280×720[px]

** 1024×576[px]

*** 512×288[px]

4.3.1 対象データ

学習済みモデルには、ドライブレコーダ動画を12本(約24分)を解析し、計1765台分の車両の結果を用いた。動画の元サイズは1280×720[px]であり、解析負荷の軽減のため、16:9のアスペクト比を保ったまま512×288[px]にリサイズする。動画の撮影エリアは主に京都府内の観光地であり、天候条件は晴れまたは曇りのデータを対象とした。内訳は路上駐停車以外の車両(以下、一般車)1709台に対して、路上駐停車56台である。ただし、この比率は時期や時間帯によって変動するものである。

4.3.2 特徴量

本論文で用いた特徴量を表2に示す。特徴量には、認識された車両に付与されるBBの座標情報(BB)と撮影車側(デバイス)の速度や加速度などの自車データを用いる。扱うBBの座標値を図2に示す。画像に対してBBの左上及び右下の x 座標と y 座標の4つの座標値を取得し、それぞれの幅や高さ、面積などを特徴量として作成する。そして、座標情報と撮影車情報に対して最大最小、平均、分散を求め45個の特徴量を作成する。これらの車両ごとに保持した特徴量を用いて、学習モデルを作成する。

4.3.3 物体認識モデル

ドライブレコーダ動画解析のため、物体認識モデルには



図 3 YOLOv3-tiny による認識

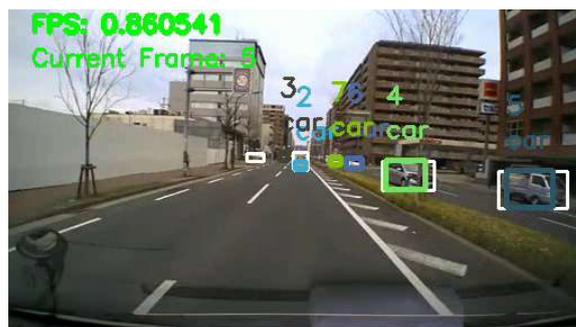


図 4 YOLOv4 による認識

YOLOv3(You Only Look Once)を選定した[12]。YOLOは、認識精度が高く計算量の少ない低負荷なアルゴリズムとして、物体認識の分野で幅広く用いられている。学習済みデータセットとしてCOCOデータセットを用いた[13]。また、認識した物体のトラッキングにはDeepSORTを用いた[14]。

4.3.4 学習モデルの作成

解析したデータに対して、均衡化を行うため一般車に対してランダムサンプリングを行った。一般車と路上駐停車の均衡化し、ランダムフォレストにより学習しモデルを作成した。様々な学習手法と評価方法で検証し、ランダムフォレストのホールドアウト法で93%のF値を確認したため、このモデルを選定した。

5. 解析・処理速度及び認識精度

本章では、学習モデル作成の際の動画解析速度と、学習モデルを用いて路上駐停車判定を行った際の処理速度について述べる。また、それぞれにおける認識精度についても述べる。

5.1 解析速度

動画解析には、Google Colaboratory(以下、Colab)のGPU環境を使用した。Colabを用いた環境では、動画に対して平均15[fps]であった。Colabを用いないPC環境のみで動画を解析したところ、処理時間は1.5[fps]程度であった。PCにはMacbook Pro2018のメモリ8GBモデルを用いた。この速度は、今後のリアルタイム解析において指標

となる。

5.2 処理速度

学習モデルを用いた認識・処理に PC 環境と Jetson TX2 環境において、YOLOv3, YOLOv3-tiny, YOLOv4 の 3 つの物体認識モデルを用いて FPS (Frames Per Second) を確認した [15]。PC 環境は前節と同じである。対象動画は大通りを走行し、複数台の車両が確認できるものとした。また、動画は元サイズである 1280×720[px]* とそれぞれリサイズした 1024×576[px]**, 512×288[px]*** である。これらの結果を表 3 に示す。FPS は動画に対して、10, 20, 30 フレーム目の平均を算出した。結果に着目すると、PC, Jetson TX2 ともにもっとも軽量の YOLOv3-tiny がリアルタイムに近い認識速度であると確認された。他 2 つのモデルにおいては、環境によらず速度に大きな差は確認されなかった。また、動画の解像度を下げることによって、特に YOLOv3-tiny においては一定の速度向上が確認できた。動画上では Jetson TX2 にて確認した約 7[fps] の速度で路上駐停車の判定がされたことから、十分にリアルタイムな認識・処理が可能であると考えられる。これにより、速度の面ではエッジデバイス上で認識・処理を行うために YOLOv3-tiny を選定し、取得する動画の解像度を下げることにより、リアルタイムな認識を行う。

5.3 認識精度

認識精度については、認識・処理速度とトレードオフの関係にある。比較として Jetson TX2 上での、512×288[px] サイズの動画の 5 フレーム目における YOLOv3-tiny と YOLOv4 での認識状況を図 3 と図 4 に示す。YOLOv3-tiny はもっとも高速であるが、車両のトラッキングが定まらず、インスタンス認識などに問題がある。また、撮影車から離れた車両を認識することが難しいため距離の制限がある。それに対して、特に YOLOv4 では高精度な認識が可能であり、大通りの反対車線を走る車両も全て認識・トラッキングが可能である。YOLOv3-tiny においても認識精度を保持し高速処理を実現するために、転移学習を用いて路上駐停車判定に特化した認識を行う。

6. おわりに

本論文では、エッジデバイスによるリアルタイム路上駐停車判定システムを提案した。リアルタイムな認識・処理のために速度を計測し、エッジ処理に最適なモデルを選定した。エッジデバイス上で路上駐停車の認識・処理において、約 7[fps] の速度を確認したため、リアルタイムな判定の有効性を示した。

今後の展望として、選定したモデルの認識精度を向上させるために、転移学習を行う。車両の認識のために、様々な状況の車両によって転移学習を行うことで、汎用的なモ

デルとなる。また、認識・処理には路上駐停車判定に特化した認識モデルを作成することで、精度と速度を両立した判定システムになると考える。そして、路上駐停車であると判定された車両の位置情報を取得し、マップ上への可視化を行う。このシステムを用いて、警察や行政などと連携し、違法な路上駐停車そのものを減らすことや渋滞、事故などの路上駐停車によって引き起こされる間接的な問題の解決を図る。

謝辞 本研究を実施するにあたり、株式会社デンソー様よりドライブレコーダ動画の提供をいただいた。ここに記し感謝申し上げる。

参考文献

- [1] 全日本駐車協会東京駐車協会. 駐車対策の現状, 2019. <https://www.npa.go.jp/bureau/traffic/seibi2/kisei/tyuusyua/parking.pdf>.
- [2] Edgar A Bernal, Zhigang Fan, Yao Rong Wang, Robert P Loce, Norman W Zeck, and Graham S Pennington. Video-based detector and notifier for short-term parking violation enforcement. October 10 2013. US Patent App. 13/441,294.
- [3] 電子情報技術産業協会. ドライブレコーダー統計出荷実績, 2018. <https://www.jeita.or.jp/japanese/stat/drive/>.
- [4] 小野晋太郎, 平原清隆, 影沢政隆, 池内克史. 車載レンジセンサを利用した路上駐車車両の自動検出. 電子情報通信学会論文誌 A, Vol. 88, No. 2, pp. 247–256, 2005.
- [5] Bo Li, Tianfu Wu, and Song-Chun Zhu. Integrating context and occlusion for car detection by hierarchical and-or model. In *European Conference on Computer Vision*, pp. 652–667. Springer, 2014.
- [6] Xavier Sevillano, Elena Marmol, and Virginia Fernandez-Arguedas. Towards smart traffic management systems: Vacant on-street parking spot detection based on video analytics. In *17th International Conference on Information Fusion (FUSION)*, pp. 1–8. IEEE, 2014.
- [7] Krzysztof Malecki. A computer simulation of traffic flow with on-street parking and drivers' behaviour based on cellular automata and a multi-agent system. *Journal of computational science*, Vol. 28, pp. 32–42, 2018.
- [8] 久徳遙矢, 出口大輔, 高橋友和, 目加田慶人, 井手一郎, 村瀬洋. 過去の車載カメラ映像との時空間差分による不特定障害物検出. 2012.
- [9] Wen Xiao, Bruno Vallet, Konrad Schindler, and Nicolas Paparoditis. Street-side vehicle detection, classification and change detection using mobile laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 114, pp. 166–178, 2016.
- [10] Xuemei Xie, Chenye Wang, Shu Chen, Guangming Shi, and Zhifu Zhao. Real-time illegal parking detection system based on deep learning. In *Proceedings of the 2017 International Conference on Deep Learning Technologies*, pp. 23–27, 2017.
- [11] 松田明大, 松井智一, 松田裕貴, 諏訪博彦, 安本慶一. ドライブレコーダ動画を用いた路上駐停車判定手法の提案. 情報処理学会 マルチメディア、分散、強調とモバイル (DICOMO2020) シンポジウム, pp. 1102–1110, 2020.
- [12] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement, 2018. <https://arxiv.org/abs/1804.02767>.

- [13] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pp. 740–755. Springer, 2014.
- [14] Alex Bewley, Zongyuan Ge, Lionel Ott, Fabio Ramos, and Ben Uppcroft. Simple online and realtime tracking. In *2016 IEEE International Conference on Image Processing (ICIP)*, pp. 3464–3468. IEEE, 2016.
- [15] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. Yolov4: Optimal speed and accuracy of object detection, 2020. <https://arxiv.org/abs/2004.10934>.