

走行履歴情報からの状況適応型走路の自動獲得

右島 琢也^{1,a)} 久徳 遙矢² 出口 大輔¹ 川西 康友¹ 井手 一郎¹ 村瀬 洋¹

概要

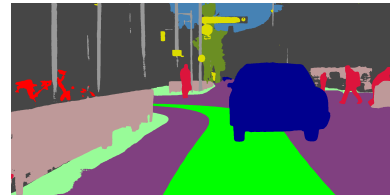
近年、自動運転技術が注目を集めており、その実現に向けた研究開発が盛んである。自動運転の実現のため、安全に走行可能な道路領域を推定する走路推定技術が求められている。走路の推定には周囲環境認識技術が不可欠であり、セマンティックセグメンテーションがその中で重要な役割を担うと期待されている。これまでに、セマンティックセグメンテーションを走路推定に用いる試みは行われているが、道路上に存在する他の車両や歩行者との接触は考慮されていない。そこで我々は、接触を回避した走路を得るために、車速の変化に着目して自車両の減速行動を学習データへ反映し、さらに複数フレームを用いて他物体の動きの学習を試みる。本発表では、自車速を用いて以降の自車両の走行履歴情報を画像上に投影することで、多数のラベル画像を自動的に獲得し、他物体との接触がない走路(状況適応型走路)を推定する手法を提案する。名古屋駅周辺の走行履歴情報を用いて提案手法の評価を行なった結果、自動獲得した画像から走路推定器を構築できることを確認した。

1. はじめに

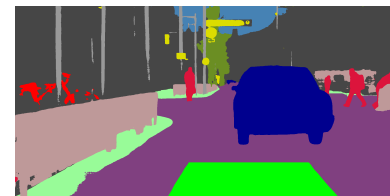
近年、交通事故対策方法の一つとして、自動運転技術への期待が高まっている。高速道路での自動運転は実現しつつあるが、周囲環境の多様性から一般道での自動運転は依然として困難である。安全に車両を自動走行させるためには、様々な物体が存在する環境においても走行できる走路を把握する必要があるため、車載カメラを用いた走路推定への需要が高まっている。

走路は車両が今後走行する経路を意味し、走路推定は自動運転の車両制御に影響を与える重要なタスクである。走路を推定するためには詳細に周囲環境を認識し、歩行者や他車両の状況を反映する必要がある。

この周囲環境認識技術の一つに、セマンティックセグメンテーションがある。セマンティックセグメンテーションは画像中の物体領域を推定する枠組みであり、入力画



(a) 通常の走路の例.



(b) 状況適応型走路の例.

図 1 通常の走路と状況適応型走路の違い.

像に対し画素単位でラベル付けされた画像が出力される。Barnes ら [1] や Zhou ら [2] はこの枠組みを拡張し、画像中に直接的な画像特徴としては存在しない走路の推定をセマンティックセグメンテーションを用いて実現する手法を提案している。これらの研究では、車両の移動経路を画像上に投影したものを学習データとすることでセマンティックセグメンテーションモデルを構築し、走路を推定している。しかし、これらの手法により推定される走路は、図 1(a) の緑色に示すような減速や一時的な停車などの先まで長期的に含めて車両が今後走行する領域であり、走路領域上の物体との接触は考慮していない。そのため、安全な自動運転の実現を考えた場合、前方車両との接触を回避する図 1(b) のような領域を走路とする必要がある。そこで我々は、このような走路を状況適応型走路と定義し、この推定を目的とする。

状況適応型走路を推定するためには、下記の二つの課題を解決する必要がある。第 1 の課題は、学習データの生成である。前述したように、状況適応型走路の推定にセマンティックセグメンテーションモデルを用いる場合、走路領域上の物体との接触を回避する走路ラベルを生成する必要がある。第 2 の課題は、前方物体の移動状態の学習である。

¹ 名古屋大学

² 豊田工業大学

^{a)} migishimat@murase.is.i.nagoya-u.ac.jp

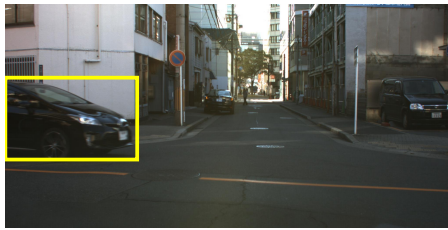


図 2 前方に車両が存在する例。

前方車両が走路決定に影響を及ぼす例を図 2 に示す。車両が存在する交差点を直進する際、黄枠で示した前方車両の動きによって状況適応型走路は変化する。そのため、前方物体の移動状態を考慮してモデルを構築する必要がある。

第 1 の課題を解決するため、我々は自車速を学習データ生成に用いる。走行中の自車両が減速した場合、その運転者が車両が減速させた原因が車両前方に存在すると考えられる。しかし、減速の原因となる物体は歩行者や他車両だけでなく、信号機、交通標識、交差点の有無など多岐にわたる。そこで我々はその前段階である、減速行動に着目して学習データ生成を行なうことで、前方物体との接触を回避する、状況適応型走路ラベルを生成する。

第 2 の課題を解決するため、我々は複数フレームの学習を行なう。そのための手段の一つに、RNN (Recursive Neural Network) の一種である LSTM (Long Short-Term Memory) がある。一般的に、LSTM は時系列の情報を学習可能である一方で、空間方向の情報を失ってしまう。この空間情報保持に関する問題を解決するため、LSTM 中の全結合層を畳み込み層に変更したネットワークである ConvLSTM [3] を用いる。セマンティックセグメンテーションにおける ConvLSTM の応用例として、Future prediction [4] などのタスクがあり、我々は車載カメラ画像上の他物体の移動を学習するため、ConvLSTM を走路推定モデルに組み込む。

上記の考え方に基づき実際の走行履歴情報から学習データを自動獲得し、状況適応型走路を推定する手法を提案する。本発表の貢献は下記の二つである。

- 自車両の速度情報を用いた、状況適応型走路推定器の学習用データの自動獲得
- ConvLSTM を用いた、他物体の移動を捉えた状況適応型走路の推定

2. 提案手法

図 3 に、状況適応型走路を推定する一連の処理手順を示す。提案手法の学習段階は、学習データの自動獲得部分と獲得した学習データを用いた推定器構築部分からなる。まず、実環境を走行して撮影した車載カメラ画像と、その車両の移動経路および速度情報を利用して学習データを自動獲得する。ここで、走路推定に必要な情報である道路や歩行者、他車両などのラベルをアノテーションするため、既

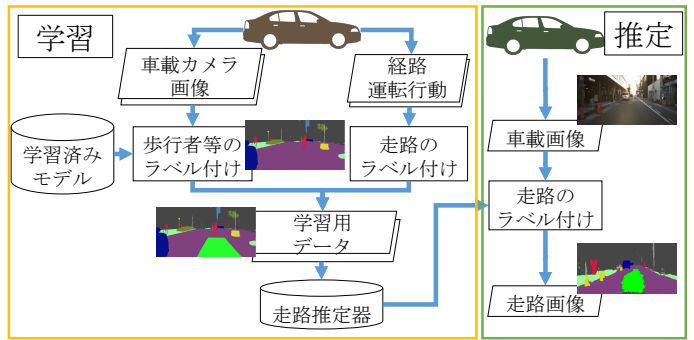


図 3 提案手法の処理手順。

存のセマンティックセグメンテーションモデルを活用する。同時に走路ラベルを画像上に付与するため、車両が実際に走行した経路を収集し、速度情報で制限を加えつつ画像上に投影する。そして、走路ラベルとそれ以外のラベルを統合することで、学習データのラベル画像とする。次に、作成した学習データを用いてセマンティックセグメンテーションモデルを構築し、走路推定器とする。この走路推定器は画像を入力とするものであり、推定時、すなわち実利用時には車載カメラ画像のみを用いて走路を推定する。

以下、2.1 節で学習データの自動獲得部分について述べ、2.2 節で走路推定器について説明する。

2.1 学習用データの自動獲得

セマンティックセグメンテーションモデルの構築には、車載カメラ画像とその画像に対応するラベル画像が必要となる。ラベル画像に必要な情報は走路推定の真値となる走路領域のラベル、および走路決定に影響する歩行者・他車両・道路などのラベルである。そのうち、走路以外の歩行者・他車両・道路などのラベルについては既存のセマンティックセグメンテーションモデルを活用して付与する。また、走路ラベルの付与には車両の自己位置推定結果による、車両が実際に走行した経路を用いる。任意の時刻 t における車両位置を $X_t = \{x_t, y_t, z_t\}$ とした際、この位置のある時刻 t' におけるセンサ座標系へ変換する行列を $F_{t'}$ と定義すると、時刻 t' のセンサ座標系における時刻 t の車両位置 X'_t は式 (1) で表される。

$$X'_t = F_{t'} X_t \quad (1)$$

ここで、この処理を以下の停止条件に合致するまで繰り返す。時刻 t' におけるセンサ座標系での各時刻の車両位置の集合 $\chi'_{t'}$ を獲得する。自車両より一定以上遠方に存在する走路は推定が困難であり、かつ実利用時には不要である。そこで、まず自車両の位置 $X_{t'}$ と各地点 X_t との距離を求め、 $X_{t'}$ からの移動距離が D より大きくなった時点で変換を停止する。また、1. で述べたように、走路の決定に影響を及ぼす前方物体の移動状態を学習するため、減速行動に

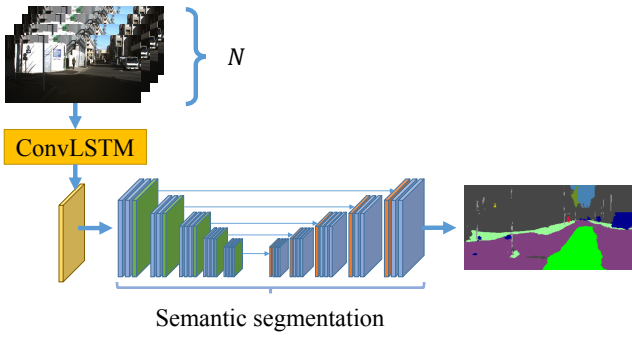


図 4 走路推定用ネットワーク.

着目して学習データを生成する. 減速し始めた瞬間に, 通常の走行を妨げる物体が車両前方に存在すると仮定し, その時点で車両位置の変換を停止する. 以上の条件に自車が停止した場合を加えた最終的な停止条件は式 (2) で示される. なお, α_t は時刻 t における加速度 $[\text{m/s}^2]$ を示し, v_t は時刻 t における速度 $[\text{m/s}]$ を示す.

$$\|X_t - X_{t'}\|_2 > D \quad \text{もしくは} \\ \alpha_t \leq -2 \quad \text{もしくは} \quad v_t = 0 \quad (2)$$

その後, 車両の走行経路, すなわち x'_t を画像上に投影する. 投影関数を P とすると, 画像上の画素位置 $\{x'_t, y'_t\}$ は式 (3) となる.

$$\begin{bmatrix} x'_t \\ y'_t \\ 1 \end{bmatrix} = P \begin{bmatrix} X'_t \\ 1 \end{bmatrix} \quad (3)$$

2.2 走路推定器の構築

上記の処理で獲得した学習用データを用いて走路推定器を構築する. ここで 1. で述べたように, 前方環境の移動状態を学習するため ConvLSTM を利用する. 構築したネットワークを図 4 に示す. このネットワークは, 前半の ConvLSTM で合計 N フレームの車載カメラ画像から時系列情報を抽出し, 後半のセマンティックセグメンテーションモデルで走路領域を含むラベル画像を推定する.

3. 評価実験

提案手法の有効性を確認するため, 名古屋駅周辺の走行履歴情報を用いて評価実験を行なった.

本実験では走路推定器に用いる車載カメラ画像のフレーム数 N を変更し, 走路推定精度を比較した. また, 推定結果とその真値の比較から求まる指標である, 走路ラベル領域の Intersection over Union (IoU) を評価に利用した. IoU は獲得したデータセットのラベル領域を \mathcal{A} , \mathcal{A} を利用して構築した走路推定器の推定結果を \mathcal{B} とした時, 式 (4) で表される値である.

$$\text{IoU} = \frac{\mathcal{A} \cap \mathcal{B}}{\mathcal{A} \cup \mathcal{B}} \quad (4)$$

表 1 走路ラベル推定結果

フレーム数 N	走路 IoU
1	65.3
3	64.7

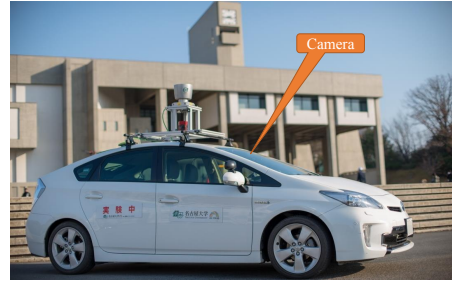


図 5 実験用車両.

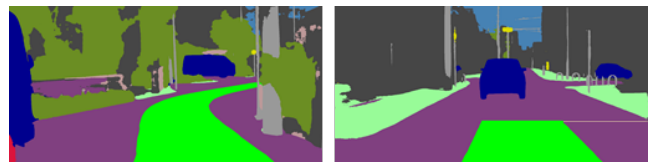


図 6 自動獲得したデータセットの例.

3.1 実験用データセットの作成

提案手法を評価するため, 図 5 に示す車両で名古屋駅周辺を走行して走行履歴情報を収集した.

走路以外のラベルを付与するため, Cityscapes データセット [5] で学習された DeepLab [6] モデルを利用した. Cityscapes データセット中のラベルは car, truck, bus のように車のラベルが細分化されている. この細分化は走路推定に対する影響が小さいと考えられるためこれらを統合し, 走路推定に必要な情報のみを学習に用いた.

提案手法によって走行履歴情報から自動獲得された学習用データセットの例を図 6 に示す. 獲得した学習用データセットのうち, 学習には 3,848 枚を用い, 評価には 1,925 枚を用いた.

3.2 実験結果および考察

図 7 に走路推定器の推定結果を示す. また, 走路ラベル領域の IoU を表 1 に示す.

図 6 より, 提案手法により車両前方に停車する車両との接触を回避する学習用データを獲得できたことを確認した. また $N = 3$ が ConvLSTM による他物体移動を考慮した場合であり, 図 7 iii) のように有効なケースがあることを確認した. しかし, 図 7 iv) のように推定に失敗するケースも存在した. また表 1 より, 走路推定精度も $N = 1$ の推定結果と比べ低下した. これは提案するネットワークにおいて, ConvLSTM をネットワークの前半部分に組み込んだため, 画像を直接 ConvLSTM に学習させたことから最適化が複雑になったためであると考えられる. そのため, 今後は適切なネットワーク構造の検討, および学習条

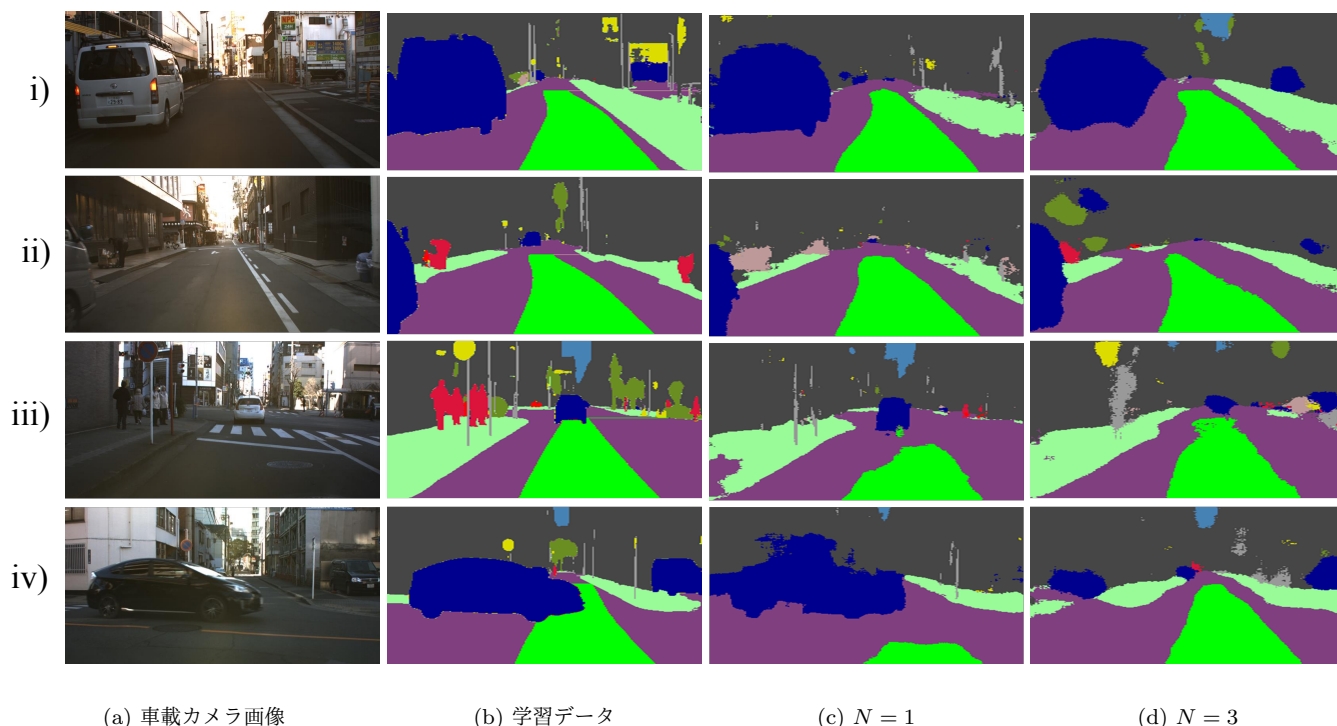


図 7 推定結果例.

件の再設定を行なう必要がある。

4. むすび

本発表では、状況適応型走路推定器の構築を目的とし、実際の走行履歴情報中の車載カメラ画像、自車速、位置情報を併用した学習データの自動獲得手法を提案した。具体的には、車両が実際に走行した経路を速度情報を元に画像へ投影し、学習データとした。そして獲得した学習データを用いてセマンティックセグメンテーションモデルを構築することで、状況適応型走路を推定した。

評価のため名古屋駅周辺の走行履歴情報に提案手法を適用した。その結果、5,773 枚のラベル画像を自動獲得でき、獲得した学習データを用いて状況適応型走路推定器を構築した。この走路推定器の推定精度を評価した結果、65.3%の精度で走路を推定できることを確認した。また、一部の画像について他物体の移動を学習することの効果を確認した。

本研究に関する今後の課題として、走路推定器構築におけるネットワーク構造や学習条件の検討が挙げられる。

謝辞 本研究の一部は科学研究費補助金による。

参考文献

[1] Barnes, D., Maddern, W. and Posner, I.: Find your own way: Weakly-supervised segmentation of path proposals for urban autonomy, *2017 IEEE International Conference on Robotics and Automation, ICRA 2017, Singapore, Singapore, May 29 - June 3, 2017*, pp. 203–210

(online), DOI: 10.1109/ICRA.2017.7989025 (2017).
 [2] Zhou, W., Worrall, S., Zyner, A. and Nebot, E. M.: Automated Process for Incorporating Drivable Path into Real-Time Semantic Segmentation, *2018 IEEE International Conference on Robotics and Automation, ICRA 2018, Brisbane, Australia, May 21–25, 2018*, pp. 1–6 (online), DOI: 10.1109/ICRA.2018.8460486 (2018).
 [3] Shi, X., Chen, Z., Wang, H., Yeung, D., Wong, W. and Woo, W.: Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting, *Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015, December 7–12, 2015, Montreal, Quebec, Canada*, pp. 802–810 (2015).
 [4] shahabeddin Nabavi, S., Rochan, M. and Wang, Y.: Future Semantic Segmentation with Convolutional LSTM, *British Machine Vision Conference 2018, BMVC 2018, Northumbria University, Newcastle, UK, September 3–6, 2018*, p. 137 (2018).
 [5] Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S. and Schiele, B.: The Cityscapes Dataset for Semantic Urban Scene Understanding, *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27–30, 2016*, pp. 3213–3223 (online), DOI: 10.1109/CVPR.2016.350 (2016).
 [6] Chen, L., Papandreou, G., Kokkinos, I., Murphy, K. and Yuille, A. L.: DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 40, No. 4, pp. 834–848 (online), DOI: 10.1109/TPAMI.2017.2699184 (2018).