

セマンティックセグメンテーションを用いた信号機認識

○正木 翔大[†], 後藤 圭汰[†], 平川 翼[†], 山下 隆義[†], 藤吉 弘亘[†]

[†]:中部大学

{masaki@mprg.cs., keita510@mprg.cs., hirakawa@mprg.cs., takayoshi@isc., fujiyoshi@isc.}@chubu.ac.jp

概要:自動運転支援において,車載画像からの信号機認識は重要なタスクの一つである.従来の信号機認識は,実画像に対して信号機の色や形状から識別を行う.しかし,車両のライトや太陽光などの影響により信号機とは関係のない領域まで抽出する可能性がある.そこで本研究では,セマンティックセグメンテーションを用いた信号機領域の抽出を行い,抽出結果から信号機の状態識別を行う手法を提案する.提案手法では,セマンティックセグメンテーションによる認識を行うことで信号機領域の抽出を行い,状態識別では,色抽出,Adaboost,CNNを用いた状態識別を行う.評価実験の結果,セマンティックセグメンテーションを用いることで高精度な信号機の認識を行うことができ,状態識別ではCNNによる識別がもっとも高い精度であることを確認した.

セマンティックセグメンテーション,信号機認識,自動運転支援

1. はじめに

自動運転支援システムにおいて,システムが信号機を認識し,信号機の状態を判断することが必要である.信号機の状態判断には,車載画像から信号機の灯火色を判断する必要がある.従来の信号機認識は,地図データを用いて信号機の認識を行う手法[1]や,実画像に対して信号機の色や形状から機械学習を用いて識別を行う手法[2]がある.しかし,全道路の地図データを用意することは,コスト的に困難である.また,車両の尾灯や太陽光などの影響により信号機とは関係のない領域を認識してしまう可能性がある.そこで本研究では,セマンティックセグメンテーションを用いた信号機領域の抽出を行い,抽出結果から信号機の状態識別を行う.セマンティックセグメンテーションは,画像内の物体をピクセル単位で識別するタスクである.そのため歩行者,道路,信号機など複数の物体を識別することができ,シーン認識に有効なタスクである.セマンティックセグメンテーションを活用することで,車両の尾灯や環境光などの影響を低減し誤認識を防ぎ,高精度な識別が期待できる.

2. 関連研究

DCNNを用いたセマンティックセグメンテーションの代表的な手法として,Fully Convolutional Network (FCN)[3],SegNet[4],DeepLab v3+[5]がある.FCNは,畳み込み層とプーリング層で構成されたネットワークであり,全結合層を畳み込み層

に置き換えることにより異なる入力サイズに対応できるという特徴がある.SegNetは,Encoder-Decoder構造を用いた手法である.Encoderでは入力画像を圧縮し,特徴マップのPooling時に座標情報を獲得する.Decoderでは,Encoderで獲得した特徴マップと座標情報を用いて入力サイズまで拡大する.この時,値が復元されない位置には0を補間する.これにより,正確な位置に情報を復元できる.DeepLabv3+は,DeepLab v3[6]を拡張したEncoder-Decoder構造のネットワークである.Decoderの改良によるオブジェクトの境界周りの認識精度向上とXception[7]モデルの採用により高精度なセマンティックセグメンテーションが可能になる.

3. 提案手法の流れ

本研究では,セマンティックセグメンテーションを用いて信号機の認識を行い,その結果から信号機の状態識別を行う.図1に提案手法の流れを示す.

3.1. セマンティックセグメンテーション

入力画像から,信号機の領域を識別するため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には,DeepLab v3+を用いる.DeepLab v3+はEncoder-Decoder構造をしており,Encoderでは,1つの特徴マップに対して異なる位置に畳み込みを行うAtrous Spatial Pyramid Pooling (ASPP)を導入している.これにより,ASPPはマルチスケールな特徴を獲得できる.

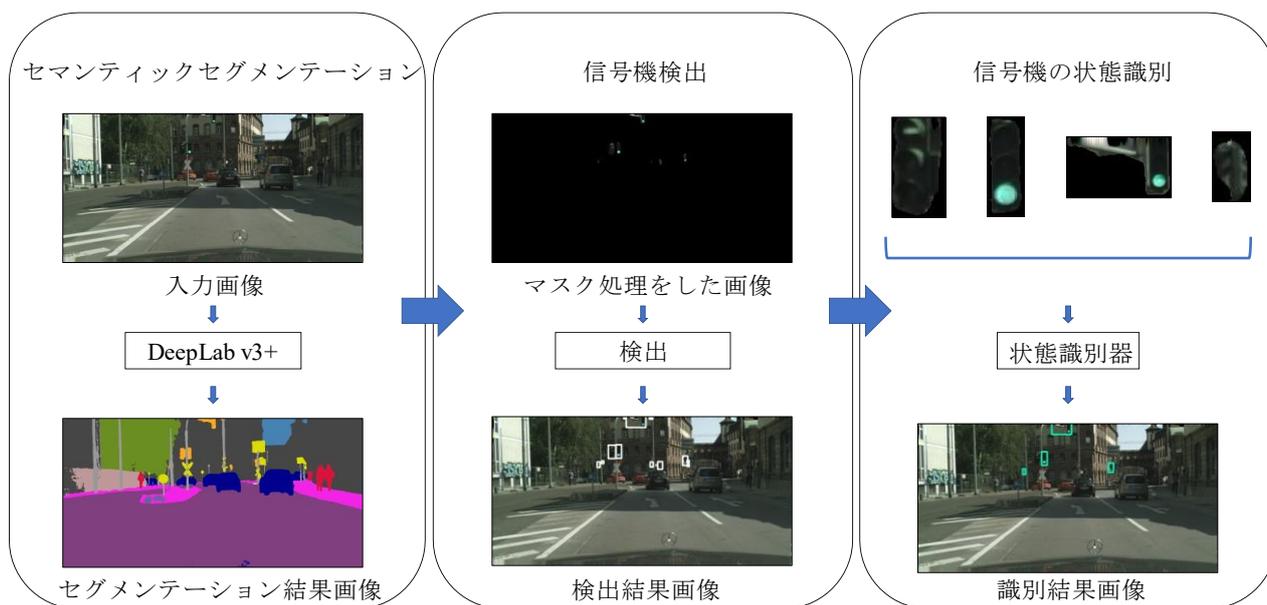


図 1：提案手法の流れ

また、Xception モデルの Max Pooling を Depthwise Convolution に置き換えることで、セグメンテーションの精度を保ちつつパラメータの削減を実現している。Decoder では、低次元の特徴を Encoder で獲得した特徴マップと結合することにより、オブジェクト境界周りの認識精度の向上が期待できる。

3.2. 信号機の検出

セマンティックセグメンテーションの出力結果から、信号機領域をピクセル単位で獲得できる。セグメンテーション結果を用いてマスク処理を行い、信号機領域のみを獲得した画像を獲得する。そして、獲得した信号機領域を外接矩形で囲い、検出結果とする。

3.3. 信号機の状態識別

信号機の状態識別は、色の抽出による識別、Haar-like 特徴と Adaboost によるカスケード型分類機、CNN の 3 種類を用いた識別を行う。色の抽出による状態識別は、HSV 色空間を使用して灯火色の赤色と青色を抽出し、円形のテンプレートを用いて色の識別を行う。円形のテンプレート内に赤色と青色の閾値のピクセルが 30% 以上ある場合灯火領域と判定する。灯火色の抽出に使用した HSV 色空間の閾値を表 1 に示す。

Haar-like 特徴と Adaboost によるカスケード型識別器を用いた状態識別では、Cityscapes[8]の学習用データから信号機を検出した画像を学習データに使用する。学習用データと検証用データは 30×60 のサイズに正規化処理する。赤信号 1357 枚、青信

表 1：HSV 色空間の閾値

	H	S	V
赤	$0 \leq H \leq 40$ or $160 \leq H \leq 180$	$70 \leq S \leq 200$	$100 \leq V \leq 255$
青	$55 \leq H \leq 110$	$70 \leq S \leq 200$	$100 \leq V \leq 255$

号 974 枚、裏向きや横を向いている信号機をその他クラスに分け学習させて信号機の状態識別を行う。弱識別器の数は 100 個に設定する。

CNN を用いた状態識別は、ネットワークに ResNet 18[9]を用いる。学習用データと検証用データは、 224×224 に正規化処理する。学習データには、Adaboost と同じ枚数の赤信号、青信号、その他の信号機のデータを学習させて状態識別を行う。

4. 評価実験

Cityscapes データセットを用いて、信号機認識の評価実験を行う。信号機領域の認識にセマンティックセグメンテーションを用いて、その結果から信号機の状態識別を行う。

4.1. 実験条件

DeepLab v3+は、事前学習済みモデルに、ImageNet[10]で事前学習を行なったモデルを使用する。識別クラスは 19 クラスである。学習時と評価時の画像サイズは 1024×2048 で行う。

状態識別に用いる ResNet 18 は、ImageNet で事前学習を行ったモデルを使用する。データのバリエーションを増加させるため、Data Augmentation を行う。信号機の状態識別の評価には、評価データに信号機の灯火色ラベルを付けたものを使用する。

表 2: セマンティックセグメンテーションの定量的評価[%]

Mean IoU	信号機の IoU
77.3	61.93

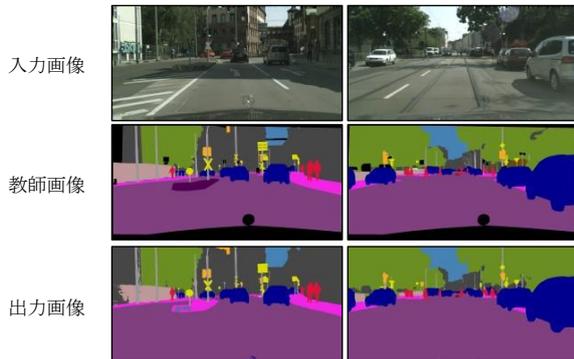


図 2: セグメンテーション結果画像

4.2. 評価方法

セマンティックセグメンテーションの評価には Intersection over Union (IoU)を用いる。IoU は正解領域と予測領域がどの程度重なっているかを示す指標である。また、クラスごとの IoU の平均をとることで Mean IoU が求められる。正解した領域を TP, 正解クラスを別のクラスと予測した領域を FP, 別のクラスを正解のクラスと予測した領域を FN とすると式(1)のように表すことができる。

$$IoU = \frac{TP}{TP + FP + FN} \quad (1)$$

信号機の検出率は、セグメンテーションの教師画像と結果画像で信号機の外接矩形をとり、教師画像の信号機のうち、いくつ検出できているかを評価する。正解の信号機に対して、検出した信号機の領域が 50%以上一致している場合を検出成功とする。状態識別の評価は、信号機領域の灯火部分にアノテーションを行った評価データを用いる。

4.3. 評価結果

信号機の認識、検出、状態識別を行った結果を示す。

4.4. セマンティックセグメンテーションの評価結果

図 2 に DeepLab v3+による結果画像を示す。教師画像と出力画像を比較すると、信号機領域を捉えることができていることがわかる。定量的評価を表 2 に示す。表 2 より、Mean IoU は 77.3%, 信号機の IoU は 61.93%となった。信号機は小さな物体のため認識が難しいが、提案手法では高い認識精度であることがわかる。

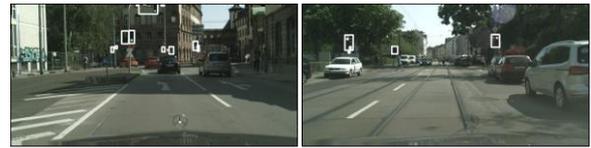


図 3: 信号機の検出結果画像

4.5. 検出結果

図 3 に信号機の検出結果画像を示す。信号機の検出率は、59.6%となった。図 3 より、遠方かつ後ろ向きの信号機の検出が出来ていないことが確認できる。しかし、信号機の状態識別において後ろ向きの信号機は認識する必要がない。そのため認識には問題がないと言える。次に、灯火領域にラベル付けした信号機の検出率を調べたところ 86.23%であった。認識が出来ていない信号機の多くは遠方のものが多いことがわかった。

4.6. 状態識別結果

図 4 に状態識別の結果画像を示す。また、表 3 に状態識別の定量的評価を示す。色抽出の結果画像を見ると、赤信号が未検出となっていることがわかる。未検出の信号機の多くは、色の抽出ができていないため未検出となっている。しかし、色を抽出する範囲を広げた場合、信号機の枠部分が多く残り誤検出の割合が増えてしまう。そのため色の抽出から状態識別を行うことは困難であると考ええる。

Adaboost を用いた識別器は、未検出、誤検出ともに最も多い結果となった。結果画像を見ると後ろ向きの信号機を赤信号と認識しているものや、認識できていない信号機があることがわかる。また、赤と青を間違えて認識するものも確認できた。原因として、学習データの不足などが考えられるが、同じデータを使用した CNN での識別結果のほうが精度が高いため Adaboost による識別は難しいと考える。

CNN を用いた状態識別の結果は、正答率が一番高いが、誤検出率は色抽出よりも多い結果となった。これは、セマンティックセグメンテーションの結果が信号機以外の領域を抽出したものを誤認識している。学習したデータには、信号機領域のみのデータを使用したため、信号機以外の領域が多い入力画像に対して誤認識が起りやすいと考えられる。そのため、学習データを増やすことで精度向上が可能であると考ええる。

5. おわりに

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信号機認識を行った。セグメンテーシ

表 3 : 状態識別の定量的評価[%]

識別器	灯火色	正答率	未検出	誤検出
色抽出	赤	51.82	38.46	9.71
	青	47.61	43.33	9.04
Adaboost	赤	17.74	58.36	23.89
	青	29.90	33.55	36.54
CNN	赤	60.00	16.89	23.10
	青	52.86	31.27	15.58

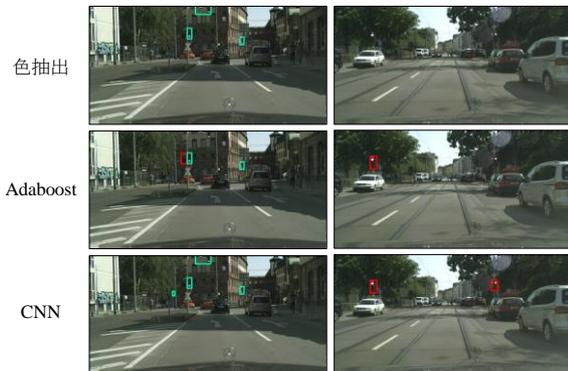


図 4 : 信号機の状態識別結果画像

ョン結果から高精度な信号機領域を獲得できることが確認できた。状態識別では、色抽出と Adaboost による状態識別は困難であることを確認した。また、CNN を用いることで高い識別を行うことができることを確認した。しかし、CNN は未知データがある領域に対して精度が悪くなるため学習データを増加させる必要があると考える。今後は、状態識別の精度を向上させる手法を検討する。

謝辞

本研究は、総合科学技術・イノベーション会議の戦略的イノベーション創造プログラム (SIP) 第 2 期/自動運転(システムとサービスの拡張)「自動運転技術(レベル 3, 4)に必要な認識技術等に関する研究」(管理法人: NEDO)によって実施されました。

参考文献

- [1] 米陀 佳祐, 菅沼 直樹, モハマド・アムロ・アッディバージャ, “複数の点灯物による誤検出を考慮した信号機認識” 自動車技術会論文集, Vol.48, No.2, p.575-580, 2017.
- [2] 松尾 治夫, 木村 久美子, “学習及び形状・色識別を用いた交通信号機認識” 情報処理学会研究報告, Vol.2014-CVIM-192, No34, p.1-7, 2014.
- [3] J. Long, E. Shelhamer, T. Darrell, “Fully convolutional networks for semantic segmentation,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.
- [4] V. Badrinarayanan, A. Kendall, R. Cipolla, “SegNet: A Deep Convolutional Encoder-Decoder Architecture for

Image Segmentation” IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 2017.

- [5] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam. “Encoder-decoder with atrous separable convolution for semantic image segmentation” European Conference on Computer Vision (ECCV), 2018.
- [6] L.-C. Chen, G. Papandreou, F. Schroff, H. Adam. “Rethinking Atrous Convolution for Semantic Image Segmentation”, European Conference on Computer Vision (ECCV), 2017.
- [7] Francois Chollet. “Xception: Deep Learning with Depthwise Separable Convolutions”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
- [8] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele. “The cityscapes dataset for semantic urban scene understanding” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- [9] H. Kaiming, X. Zhang, S. Ren, J. Sun” Deep Residual Learning for Image Recognition”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- [10] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. F. Fei. “ImageNet: A Large-Scale Hierarchical Image Database”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009.

正木翔大: 現在中部大学工学部情報工学科在学中, セマンティックセグメンテーションを用いた信号機認識の研究に従事。

後藤圭汰: 2018 年中部大学工学部情報工学科卒業。現在中部大学大学院修士課程在学中, 全周囲画像を用いたセマンティックセグメンテーションの研究に従事。

平川翼: 2013 年広島大学大学院博士課程前期修了, 2014 年広島大学大学院博士課程後期入学, 2017 年中部大学研究員 (~2019), 2017 年広島大学大学院博士後期課程修了。2019 年中部大学特任助教。コンピュータビジョン, パターン認識, 医用画像処理の研究に従事。

山下隆義: 2002 年奈良先端科学技術大学院大学博士前期課程修了。2002 年オムロン株式会社入社, 2009 年中部大学大学院博士後期課程修了(社会人ドクター), 2014 年中部大学講師, 2017 年より同大学准教授。人の理解に向けた動画像処理, パターン認識・機械学習の研究に従事。

藤吉弘亙: 1997 年中部大学大学院博士後期課程修了。1997 ~2000 年米カーネギーメロン大学ロボット工学研究所 Postdoctoral Fellow。2000 年中部大学講師。2004 年より同大学教授。2005 ~2006 年米カーネギーメロン大学ロボット工学研究所客員研究員, 計算機視覚, 動画像処理, パターン認識・理解の研究に従事。