

1U-04

構内自走ロボットのためのカメラ画像セグメンテーションによる周辺環境認識

橋本 幸樹† 佐藤 明生‡ 高橋 竜朗‡ 小宮 聖司‡ 脇田 敏裕‡ 宮崎 剛†
 † 神奈川工科大学情報学部情報工学科 ‡ 神奈川工科大学創造工学部自動車システム開発工学科

1 はじめに

近年、日本では少子高齢化が深刻化してきており、65歳以上の老年人口を少ない人数の若い世代で支えなくてはならない状況になってきている。生産年齢人口は1995年頃をピークに減少してきており、今後も減少の一途を辿るといって推計が出されている [1]。生産年齢人口が減ることにより、労働力の確保が今後ますます難しくなっていくであろう日本では、様々な業種においてその労働力をロボットで代替しようとしている。

本学（神奈川工科大学）でも「KAIT モビリティリサーチキャンパス」と銘打った、ロボットを用いた業務効率化計画が進められている [2]。これは、大学での生活や学びを豊かで楽しいものにするため、自ら考えて行動するロボットたちが大学構内（以降、構内とする）を動き回るといって計画である。この計画で実現させようとしている機能の1つに、構内巡回や構内清掃がある。構内を巡回して不審物や不審人物を発見したら警備員に通報したり、構内の清掃を行ったりすることで、警備員等の負担軽減やコスト削減が期待できる。

本研究で実現しようとしている自律走行ロボットは3D-LiDARとWebカメラを搭載し、3D-LiDARでロボット前方の障害物等を検出し、Webカメラでロボット近距離の障害物等を検出する。本稿ではWebカメラによる物体検出について述べる。検出対象を、落ち葉、縁石、排水溝、車、芝生、建物、背景の7クラスとし、芝生や落ち葉など対象物が集まって形作られる広がりのある不定形な領域の検出（セグメンテーション）方法を提案する。

2 セグメンテーションによる周辺環境認識

本研究で開発を進めている自律走行ロボットの外観を図1に、構成の概略を図2に示す。3D-LiDARとWebカメラはGPUを搭載した小型PC（Jetson AGX）に接続されており、Webカメラは地面から約1.1mの高さ



図1: 自律走行ロボットの外観と搭載装置

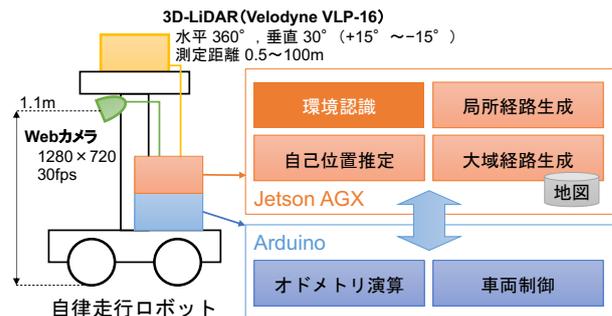


図2: 自律走行ロボットの構成（概略）

に取り付けられている。3D-LiDARの情報を元に人や障害物を回避する処理を行う。Webカメラ画像では、先述の通り、検出対象とする領域が不定形であるため、セマンティックセグメンテーションを用いて画素単位でクラス分類をする。画像のセグメンテーションにはU-Net[3]を用いる。

U-Netで学習するために、Webカメラで撮影した訓練画像に対してアノテーションを行い、教師データを作成する。ツール（VoTT）を使用して対象クラスの領域を多角形で囲み、多角形内の座標位置にクラスのラベルを付与する。図3にU-Netのデータ入出力と、出力結果から領域分割データを作成する流れを示す。出力データのサイズは $224 \times 224 \times 7$ であり、チャンネルの7に各クラスの確率が格納される。出力データにおける画素 (x, y) のクラス c の確率を $I(x, y, c)$ とすると、領域分割データ L は式(1)となる。訓練画像と画像化してクラス（芝生、建物、背景）で色分けした教師データの例を図4に示す。

Surrounding Environment Recognition by Camera Image Segmentation for Autonomous Running Robot
 †Koki HASHIMOTO ‡Haruki SATO ‡Tatsuro TAKAHASHI
 ‡Seiji KOMIYA ‡Toshihiro WAKITA †Tsuyoshi MIYAZAKI
 †Department of Information and Computer Sciences, Kanagawa Institute of Technology
 ‡Department of Vehicle System Engineering, Kanagawa Institute of Technology

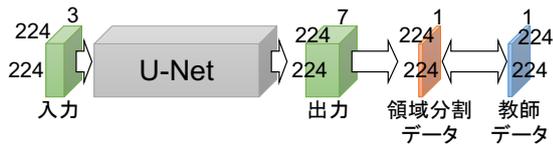


図 3: U-Net への入出力データサイズと領域分割データ

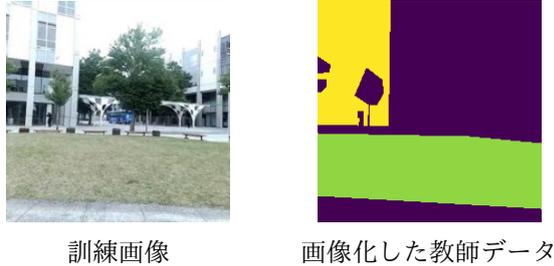


図 4: 学習用データセット例

$$L(x, y) = \arg \max_{c \in \{0, 1, \dots, 6\}} I(x, y, c) \quad (1)$$

3 実験・考察

本学構内で撮影した 450 枚の画像を使用して学習と評価を行った。画像内に含まれる各クラスの総数を表 1 に示す。表 1 に示すクラス以外の領域は背景とする。訓練用に画像 420 枚、テスト用に 30 枚使用した。U-Net のバックボーンには ResNet34 を使用し、最適化手法には Adam を、損失関数には categorical cross entropy を使用した。Google Colaboratory 上でバッチサイズ 10、エポック数 100 で学習を行い、学習済みファイルを Jetson AGX へ移して推論を行った。

表 1: データセットに含まれる各クラスの総数

落ち葉	縁石	排水溝	車	芝生	建物
263	383	229	165	23	49

テスト画像の領域分割結果の例を図 5 と図 6 に示す。なお、領域分割画像は領域分割データを画像化したものである。図 5 の画像では、落ち葉、縁石、排水溝、背景の 4 クラスの領域が正しく分割できている。図 6 の画像には落ち葉、縁石、車、背景の 4 クラスが含まれているが、それぞれ検出した領域に誤りがあった。テストデータのうち、図 6 に示すような日向と日陰が含まれるパターンを除けば、高精度で領域分割できることが確認できた。日向と日陰が含まれる明暗差のあるパターンで領域分割精度が低下した要因は、学習時のデータ数が少なかったことやデータに偏りがあったことが考えられる。今後は、明暗差のある画像も増やして学習させていく必要があると考える。

また、本稿では Web カメラ画像の領域分割手法の提案にとどまり、得られた結果をロボットの行動制御へつなげられていない。今後は、得られた情報をロボッ



図 5: 正しく領域分割できた例



図 6: 正しく領域分割できなかった例

トが持っている地図情報へマッピングさせる方法を検討し、実際のロボットの行動制御に適用させていきたいと考えている。

4 まとめ

本稿では、3D-LiDAR と Web カメラを搭載した自律走行ロボットの周辺環境認識手法を提案した。Web カメラ画像に写る、形の定まらない物体に対してセマンティックセグメンテーションを利用し、領域分割を行った。学習モデルには U-Net を使用し、本学構内で撮影した画像を使って背景を含む 7 クラスを学習した。テスト画像の領域分割では高い精度が得られたが、日向と日陰で明暗差の大きい一部の画像において、精度が悪くなることを確認した。今後は、対象クラスを増やすことや日陰や日向の場所による精度の低下を抑えることに加え、領域分割の結果をロボットの制御側へどのように渡すのかを考えていく必要がある。

参考文献

- [1] 国立社会保障・人口問題研究所：日本の将来推計人口 (平成 29 年推計), http://www.ipss.go.jp/pp-zenkoku/j/zenkoku2017/pp29_gaiyou.pdf (参照 2021-10-31).
- [2] 神奈川工科大学先進技術研究所:KAIT モビリティリサーチキャンパスの構築, <https://www.kait.jp/topics/atrc/report07.html> (参照 2021-10-31).
- [3] Ronneberger, O., Fischer, P. and Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Vol. 9351 of LNCS, pp. 234–241, Springer (2015).