

パーツを考慮した建機の検出と判断根拠の可視化

乙川 侑正[†]

† 横浜国立大学 理工学部

長尾 智晴^{††}

†† 横浜国立大学 大学院環境情報研究院

1. はじめに

建設現場において、建設機械(以下、建機)の自動運転や、その際の危険認知のために、周囲で作業する建機の種類と位置を正しく認識できることが求められており、近年ニューラルネットワークを用いた検出が行われている。しかし現状では、建機の種類と位置を正確に認識することが困難である。そこで、検出精度の低下要因を探り、その原因をもとにデータセットやモデルを修正することによって、検出精度の向上が期待できる。

本稿では、建機本体とそのパーツの検出を行い、それらの検出結果を分類根拠の説明に活用することで、建機のどのパーツに着目したかが説明しやすい手法を提案した。

2. 提案手法

図 1 に本手法で用いるモデル構造を示す。本モデルは Faster R-CNN[1]をベースにしたものとなっている。

Faster R-CNN は学習過程で、領域提案ネットワークによって物体領域候補(Region of Interest, RoI)が多数生成され、それらを全結合層に通すことでRoIのクラスとバウンディングボックスの座標を求めている。しかし、全結合層に入力する前に 1 次元ベクトルへ変換しているため、空間的な位置情報が失われてしまい、画像のどこに着目してクラスを判断したか分からない。そこで、RoI のクラス分類に Generative Contribution Mappings(GCM)[2]を用いることで、クラス分類の根拠を可視化できるようにする。

GCM はエンコーダ・デコーダ形式となっており、クラス分類を行いながら、画像のどの部分が分類根拠になったのかを可視化する手法である。本提案手法では、RoI のクラス分類を行う全結合層を、GCM のデコーダ部分に差し替えた。合計 9 層の畳み込み層・逆畳み込み層を用いて実装した。各バウンディングボックスに対して 112×112 の可視化マップが出力される。

3. 実験設定

本実験では、提供された建機 CG 画像 22000 枚と実写画像 921 枚、計 22921 枚を用いた。それぞれの画像には 4 種類の建機がランダムな位置、角度で 1~4 台写っている。また、建機本体 4 種類とパーツ 7 種類、合計 11 種類のラベルがアノテーションされている。

CG 画像のみで学習したのち、実写画像に本手法を適用して建機とそのパーツの検出、判断根拠の可視化を行った。

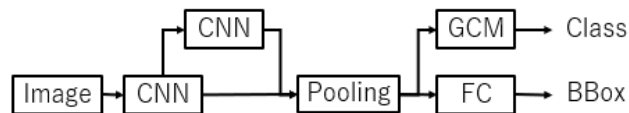


図1. モデルの構造

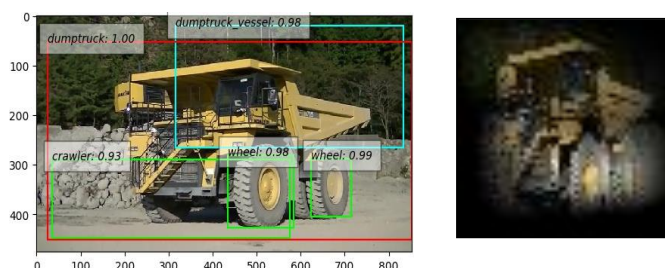


図2. 検出結果



図3. 可視化結果

4. 実験結果

図 2 に検出結果、図 3 に可視化結果を示す。図 2 では色の付いたボックス、ラベル名とスコアが示されている。本実験では CG 画像のみで学習を行ったが、実写画像に対しても建機の車種およびパーツを検出できている。また図 3 では、建機のどの部分に着目して検出結果を出力したかが示されている。図 2 に写っている建機をダンプトラックと判断しており、可視化マップを見るとタイヤの周辺に最も着目したということが分かる。このように GCM を用いて可視化したことで、建機のどこに着目して車種の判断がなされたのか確認できた。

5. まとめ

本稿では、パーツを考慮した建機検出とその判断根拠の可視化を行った。今後は、パーツごとの位置関係などを学習することで、さらに検出精度と説明性の高い手法を提案していく。

参考文献

- [1] Shaoqing Ren, et al. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.", In NIPS, 2015.
- [2] 荒井敏, 長尾智晴, "畳み込みニューラルネットワークを用いた画像分類タスクの直感的可視化方法", 研究報告数理モデル化と問題解決(MPS), vol.2016.10, pp.1-6, 2016.