

自動運転における深層学習を用いた信号認識

2020SC107 矢崎琢也 2020SC109 横山蒼太

指導教員：中島明

1 はじめに

近年、自動運転は著しく発展しており実用化への期待が年々高まっている。今後自動運転化が進むことで、交通事故の防止、ドライバーの運転負荷の軽減などの効果を得られる可能性がある。

自動車で行く際には、交通ルールを守り、障害物との接触を避けながら運転する必要がある。そのためには、刻一刻と変化する周囲の状況を認識し、その後の予測、適切な判断と操作を行うことが必要である [1]。自動運転でも同様のことが求められ、認識、判断、操作の3つの要素が組み合わさっている。本研究では、実際の信号機を用いて、正しく信号機の認識を行うことを目指している。

2 深層学習を用いた物体検出

2.1 深層学習とは

大量の学習データから規則性を自動的に学習する機械学習の手法の一つであり、多数の層からなる深層ニューラルネットワークを用いるものである [2]。

従来の機械学習は特徴抽出と分類ルールの学習を別々に行っていたが、深層学習ではそれらを同時に行うことができる。

2.2 物体検出の概要

深層学習を用いた物体検出モデルの1つとして YOLO があり、画像をニューラルネットワークに通すことで物体検出とクラス分類を同時に行える。これを用いるメリットとして短時間で正確に処理を行い、画像全体を見ての予測ができる [3]。

2.3 物体検出手法

YOLO の最新バージョンとして YOLOv5 があり、従来のモデルよりも処理が早く精度が高いという特徴がある [3]。我々はこの YOLOv5 を用いて、これから信号機の物体検出を行っていく。以下に実際に行っていく手法を簡単に説明する。

1. 教師データとなる画像を 2500 枚程度用意する
2. 1 で用意した画像の認識させたい信号機に位置とクラスのラベル付けを行う
3. 2 で作成した教師データを YOLOv5 に学習させる
4. 学習させた教師データを重みファイルとして物体検出を行う

5. 実行後、フレームとラベル (どのくらいの確率でなんというオブジェクトなのか記載したもの) 付きの画像ファイルが保存される

以上のような手順で物体検出を行っていく。

2.4 YOLOv5 を動作させた環境

YOLOv5 を動作させた環境を以下に示す。

使用 PC スペック

OS : Ubuntu 22.04.3 LTS

CPU : 11th Gen Intel(R) Core(TM) i7-11700F @ 2.50GHz

メモリ : 32GB

GPU : NVIDIA Corporation GA106 [GeForce RTX 3060]

python 環境

python 3.10.12

使用したカメラ

iPhone 11

2.5 学習画像の収集方法

信号機の認識を行うには、学習させるための画像が大量に必要となる。本研究では、実際の信号機を用いて、時間と距離を変えながら画像を撮影した。具体的な撮影方法を以下に示す。

- 枚数 : 310 枚
 - 時間 : 6 時から 24 時までの 2 時間おきに撮影
 - 距離 : 停止線から前方向 (マイナス) に 2m おきに 10m まで (青と赤 1 枚ずつ)
停止線から後ろ方向 (プラス) に 2m おきに 18m まで (青と赤 1 枚ずつ)
停止線で青, 赤, 黄 1 枚ずつ
- 撮影した信号機の例を図 1 に示す。



図 1 撮影した信号機の例 (左 : 8 時 -10m, 中 : 14 時 0m, 右 : 20 時 18m)

また、大量の画像を容易に得るために、コントラストの調整やノイズを付与するなどのデータ拡張を行った。

3 信号機認識の精度向上

本章では信号機認識の精度向上において、時間帯、距離で場合分けを行い、YOLOv5 の精度の評価や比較を行う。また、YOLOv5 の評価指標として Precision, Recall, AP(Average Precision), mAP(mean Average Precision) が用いられる。本章でも以上の評価指標を用いて、評価や比較を行う。また、今回は認識結果を定量的に判断するために、確率値を用いる。一般的に認識結果の確率値が 0.9 以上であれば認識ができているといえる。

3.1 認識を行った手順

実際に認識を行った手順を以下に示す。

- 朝 (6 時～10 時) の画像を学習させた場合
- 昼 (12 時～16 時) の画像を学習させた場合
- 夜 (18 時～24 時) の画像を学習させた場合
- 全種類 (6 時～24 時) の画像を学習させた場合

今回全てエポック数は 100, バッチサイズは 8 として認識を行った。

3.2 朝 (6 時から 10 時) の画像を学習させた場合

朝 (6 時から 10 時) に信号機を撮影し、学習、認識を行った。朝に撮影した場合において認識精度の評価、比較を行う。学習データと検証データの枚数以下に示す。

学習データの枚数 : 651 枚

検証データの枚数 : 310 枚

3.2.1 目的

朝の画像のみを学習させた場合に朝、昼、夜それぞれの認識精度にどのような影響があるかを知るため。

3.2.2 結果

朝に撮影した画像を教師データとした、YOLOv5 の学習結果を図 2 に示す。図 2 では、precision は 1 に収束しているが、recall と mAP は 1 に収束しなかった。結果としては、望ましい結果とはならなかった。

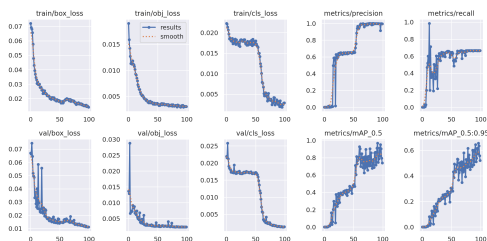


図 2 教師データが朝の場合の学習結果

実際に認識した結果を、表 1 と表 2 に示す。表 1 は、検証データが朝の平均、表 2 は昼の平均である。夜は全く認識できなかった。

表 1 教師データが朝の場合の認識結果 (検証データが朝の平均)

| 距離 | 青 | 赤 | 黄 |
|------|------|------|---|
| 0m | 0.92 | 0.84 | × |
| 2m | 0.90 | 0.89 | — |
| 4m | 0.92 | 0.81 | — |
| 6m | 0.87 | 0.84 | — |
| 8m | 0.89 | 0.92 | — |
| 10m | 0.88 | 0.85 | — |
| 12m | 0.87 | 0.74 | — |
| 14m | 0.90 | 0.79 | — |
| 16m | 0.81 | 0.87 | — |
| 18m | 0.82 | 0.73 | — |
| -2m | 0.9 | 0.89 | — |
| -4m | 0.93 | 0.94 | — |
| -6m | 0.91 | 0.93 | — |
| -8m | 0.93 | 0.93 | — |
| -10m | 0.94 | 0.98 | — |
| 平均 | 0.89 | 0.86 | — |

表 2 教師データが朝の場合の認識結果 (検証データが昼の平均)

| 距離 | 青 | 赤 | 黄 |
|------|------|------|---|
| 0m | 0.88 | 0.87 | × |
| 2m | 0.91 | 0.77 | — |
| 4m | 0.87 | 0.88 | — |
| 6m | 0.88 | 0.90 | — |
| 8m | 0.83 | 0.81 | — |
| 10m | 0.78 | 0.73 | — |
| 12m | 0.74 | 0.65 | — |
| 14m | 0.79 | 0.62 | — |
| 16m | 0.75 | 0.70 | — |
| 18m | 0.68 | 0.6 | — |
| -2m | 0.85 | 0.82 | — |
| -4m | 0.81 | 0.88 | — |
| -6m | 0.71 | 0.86 | — |
| -8m | 0.80 | 0.90 | — |
| -10m | 0.81 | 0.92 | — |
| 平均 | 0.81 | 0.79 | — |

3.2.3 評価

結果より夜は全く認識できなかった。また、表 1 と表 2 の平均を比較すると、青は表 1 の方が表 2 より確率値が高く、赤も表 1 の方が表 2 より高いことが分かる。黄色は、認識できないか違う色と認識し、正しく認識することができなかった。これらの結果から、太陽の有無 (明るさ) や太陽の方向の違いが認識精度に影響を及ぼすことが考えられる。黄色を正しく認識することができなかったのは、黄色の学習データの枚数が少なかったことが原因と考えられる。

3.3 昼 (12 時から 16 時) の画像を学習させた場合

昼 (12 時から 16 時) に信号機を撮影し、学習、認識を行った。昼に撮影した場合において認識精度の評価、比較を行う。学習データと検証データの枚数以下に示す。

表3 教師データが昼の場合の認識結果（検証データが朝の平均）

| 距離 | 青 | 赤 | 黄 |
|------|------|------|---|
| 0m | 0.39 | × | × |
| 2m | × | × | — |
| 4m | 0.32 | × | — |
| 6m | 0.37 | × | — |
| 8m | × | × | — |
| 10m | × | × | — |
| 12m | 0.27 | × | — |
| 14m | 0.32 | × | — |
| 16m | × | × | — |
| 18m | × | × | — |
| -2m | 0.57 | × | — |
| -4m | 0.32 | 0.37 | — |
| -6m | 0.66 | 0.40 | — |
| -8m | 0.42 | × | — |
| -10m | 0.66 | 0.92 | — |
| 平均 | 0.43 | 0.56 | — |

学習データの枚数：651枚

検証データの枚数：310枚

3.3.1 目的

昼の画像のみを学習させた場合に朝、昼、夜それぞれの認識精度にどのような影響があるかを知るため。

3.3.2 結果

昼に撮影した画像を教師データとした、YOLOv5の学習結果を示す。precisionは1に収束しているが、recallとmAPは1に収束しなかった。結果としては、望ましい結果とはならなかった。

実際に認識した結果を、表3と表4に認識結果を示す。表3は、検証データが朝の平均、表4は昼の平均である。夜は全く認識できなかった。

3.3.3 評価

結果から、学習データが朝の時と比べて全体的に認識の精度が低く黄色は認識できなかった。特に、朝のデータは信号と認識できたのも3割程度しかできず、認識できた精度もかなり低い。夜は、学習データが朝の場合と同様全く認識できなかった。これらの結果から、朝の時と比べて雲に隠れてしまう時間帯があったことで全体が暗くなってしまう、認識の精度が下がったと考えた。また、認識精度が全体的に低い原因の理由として学習データが朝の場合と同様、昼の学習データが少なかったことも理由の一つだと思う。

3.4 夜（18時から24時）の画像を学習させた場合

夜（18時から24時）に信号機を撮影し、学習、認識を行った。夜に撮影した場合において認識精度の評価、比較

表4 教師データが昼の場合の認識結果（検証データが昼の平均）

| 距離 | 青 | 赤 | 黄 |
|------|------|------|---|
| 0m | 0.87 | 0.82 | × |
| 2m | 0.86 | 0.77 | — |
| 4m | 0.86 | 0.79 | — |
| 6m | 0.85 | 0.81 | — |
| 8m | 0.78 | 0.85 | — |
| 10m | 0.85 | 0.75 | — |
| 12m | 0.88 | 0.88 | — |
| 14m | 0.82 | 0.84 | — |
| 16m | 0.78 | 0.86 | — |
| 18m | 0.78 | 0.72 | — |
| -2m | 0.88 | 0.86 | — |
| -4m | 0.88 | 0.91 | — |
| -6m | 0.87 | 0.92 | — |
| -8m | 0.88 | 0.91 | — |
| -10m | 0.93 | 0.96 | — |
| 平均 | 0.85 | 0.84 | — |

を行う。学習データと検証データの枚数以下に示す。

学習データの枚数：868枚

検証データの枚数：310枚

3.4.1 目的

夜の画像のみを学習させた場合に朝、昼、夜それぞれの認識精度にどのような影響があるかを知るため。

3.4.2 結果

教師データが夜の場合のYOLOv5の学習結果を図3に示す。precisionとrecall, mAPは1に収束している。結果としては、望ましい結果となった。認識結果を表5に示す。表5の認識結果は、検証データが夜の平均であり、朝と昼は全く認識できなかった。

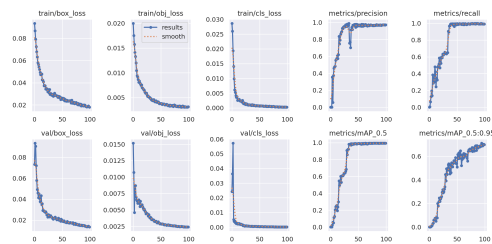


図3 教師データが夜の場合の学習結果

3.4.3 評価

結果より、朝と昼は全く認識できなかった。教師データが朝、昼の場合では認識できなかった黄色を初めて信号と認識することができた。これらの結果から、信号認識において夜の写真を学習するだけでは太陽が出ている日中の時間帯の信号を認識できないということが分かった。また、

表5 教師データが夜の場合の認識結果（検証データが夜の平均）

| 距離 | 青 | 赤 | 黄 |
|------|------|------|------|
| 0m | 0.79 | 0.86 | 0.89 |
| 2m | 0.82 | 0.89 | — |
| 4m | 0.87 | 0.89 | — |
| 6m | 0.88 | 0.90 | — |
| 8m | 0.87 | 0.88 | — |
| 10m | 0.85 | 0.90 | — |
| 12m | 0.86 | 0.89 | — |
| 14m | 0.85 | 0.84 | — |
| 16m | 0.80 | 0.77 | — |
| 18m | 0.80 | 0.73 | — |
| -2m | 0.89 | 0.85 | — |
| -4m | 0.89 | 0.92 | — |
| -6m | 0.90 | 0.90 | — |
| -8m | 0.89 | 0.90 | — |
| -10m | 0.87 | 0.89 | — |
| 平均 | 0.86 | 0.87 | — |

学習データが朝昼の時と同様学習データが少ないことで、認識精度が低いと考えられる。そこで、全種類の画像を用いて学習データを増やすことで信号認識の精度が高まるのか実験を行う。

3.5 全種類（6時から24時）の画像を学習させた場合

6時から24時に信号機を撮影し、学習、認識を行った。6時から24時に撮影した場合において認識精度の評価、比較を行う。学習データと検証データの枚数以下に示す。
 学習データの枚数：2170枚
 検証データの枚数：310枚

3.5.1 目的

全ての時間帯を学習させることによって、認識の精度にどのような影響が生じるのか知るため。

3.5.2 結果

全種類のデータを教師データとした、YOLOv5の学習結果を示す。precisionとrecall, mAPは1に収束した。結果としては、望ましい結果となった。

認識結果を表6に示す。表6の認識結果は、検証データが6時から24時の平均である。

3.5.3 評価

距離が-10mから-4mまでと-2mの青信号の時は確率値が全て0.9を超えた。それよりも距離が離れると確率値が0.9を超えることはなかった。また、夜間は確率値が明らかに低くなるがあった。これらの結果から、距離が近ければ近いほど確率値が高くなる傾向があることが分かる。夜間に確率値が明らかに低くなる原因として、前方の車のブレーキランプの影響を受けたことが原因の1つだと考えられる。

表6 教師データが全種類の場合の認識結果（検証データが6時から24時の平均）

| 距離 | 青 | 赤 | 黄 |
|------|------|------|------|
| 0m | 0.89 | 0.89 | 0.93 |
| 2m | 0.89 | 0.88 | — |
| 4m | 0.88 | 0.89 | — |
| 6m | 0.87 | 0.88 | — |
| 8m | 0.85 | 0.87 | — |
| 10m | 0.86 | 0.84 | — |
| 12m | 0.84 | 0.86 | — |
| 14m | 0.83 | 0.83 | — |
| 16m | 0.82 | 0.83 | — |
| 18m | 0.82 | 0.79 | — |
| -2m | 0.90 | 0.89 | — |
| -4m | 0.91 | 0.91 | — |
| -6m | 0.93 | 0.91 | — |
| -8m | 0.92 | 0.92 | — |
| -10m | 0.93 | 0.91 | — |
| 平均 | 0.88 | 0.87 | — |

3.6 本研究の反省点

本研究の反省点を以下に示す。

- 学習データが少なく精度が低かったので、学習データをもっと増やすべきだった。
- 信号機の撮影時、対象の信号機を拡大して撮影するべきだった。
- 周りの背景が写り分けにくかったので、ラベル付けを行う前に対象の信号機のみを切り取るべきだった。

4 おわりに

本研究では、物体検出の基本となる深層学習について学習し、YOLOv5を用いて実際に物体検出を行なった。今後は3.6節で示した内容を踏まえて実験を行い、信号機認識の精度向上を目指していく。

参考文献

- [1] Ding-Ruei Shen · Hong-Liang Chin · Chih-Hung Tu etc :『Navigation and Task Planning of a Mobile Robot under ROS Environment:A Case Study Using AutoRace Challenge』. the SICE Annual Conference 2021.
- [2] 日本ロボット学会 監修, 香月 理絵 編著, 荒井 幸代 etc 共著『自動運転技術入門』. オーム社, 東京, 2021.
- [3] 株式会社アルファシステムズ:『YOLOv5を使った物体検出』. アルファテックブログ, 2021-08-20, <https://www.alpha.co.jp/blog/20210802>.