

車載カメラ画像からカーブミラーを検出する深層学習モデルの性能比較

宮柱太一^{†1} 羽倉輝^{†2} 栗達^{†1} 河合由起子^{†2} 小野晋太郎^{†1}
^{†1} 福岡大学 ^{†2} 京都産業大学

1 はじめに

近年、自動車業界において自動運転に関する研究が盛んに進められている。日本においても2022年にはレベル4の無人運転に関する道路交通法の改正案が可決されるなど、自動運転に向けた積極的な傾向が見られている。これらの動向に伴い、自動車走行中の安全性向上は今まで以上に重要な課題と考えられる。特に見通しの悪い交差点は交通事故の主要な要因の一つであり、走行中の安全性向上には不可欠な要素である。

見通しの悪い交差点での事故を防止する手法として、Hinoら[1]は車載カメラとカーブミラーによって危険事象を認識するシステムの研究を行っており、Satoら[2]はカーブミラーの画像認識における課題である鏡面部分が低解像度である問題を解決するため、移動体の特徴を抽出する研究を行っている。しかしながら、これらの研究において、システムの起点となるカーブミラーの検出に適した物体検出モデルについての思案は行われていない。

本研究では、カーブミラーの検出に最適な深層学習モデルを探索することを目的とする。3種類の物体検出モデル及び2種類の最適化アルゴリズムを使用し、路上のカーブミラーを検出するための学習を行い、各モデルの推定結果を出力し、検出性能を比較する。

2 カーブミラーの検出方法

2.1 データセット



図1: カーブミラーを含んだ路上の画像データ

本研究は車載カメラ画像からカーブミラーを検出す

Performance Comparison of Deep Learning Models for Curve Mirror Detection from In-vehicle Camera Images.
^{†1} Taichi Miyabashira ^{†2} Hikaru Hagura ^{†1} Da Li ^{†2} Yukiko Kawai ^{†1} Shintaro Ono
^{†1} Fukuoka University ^{†2} Kyoto Sangyo University

ることを想定していたが、初期段階として、図1のように撮影した画像を用いる。福岡県内で撮影したカーブミラーを含んだ路上の画像とGoogle Street Viewの道路画像の合計771枚をデータセットとして使用した。これらのデータは事前処理としてリサイズを行い、各モデルの学習・検証に利用する。リサイズした画像の解像度は500×500を基準としているが、YOLOv4のみCNNの構造から480×480としている。この際、縦横比は考慮していない。

2.2 深層学習モデル

本稿では、Faster R-CNN [3], SSD [4], YOLOv4 [5]の3種類の物体検出モデルを比較する。Faster R-CNNのバックボーンにはResNet-50を使用し、学習は実行環境の制約によりオンライン学習(バッチサイズ1)で行う。SSDのバックボーンに関しても同様に、ResNet-50を使用し、学習はバッチサイズ32のミニバッチ学習で行う。YOLOv4のバックボーンにはCSP-DarkNet53を使用し、オンライン学習を行う。

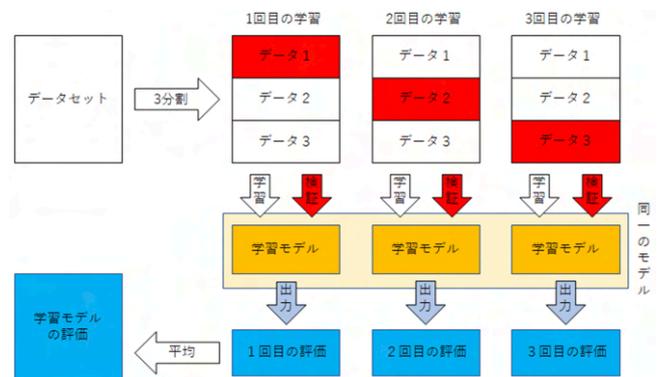


図2: 3分割交差検証の流れ

2.3 検証方法

深層学習ではニューラルネットワークの初期値や学習過程でのデータ入力によって同じモデルであっても性能に大きな差が生まれることがある。そのため、本稿ではデータセットを3分割し、2つを学習用データ、残りを検証用データに分ける3分割交差検証を行う。交差検証の流れは図2に示す。

検出結果は、正解の領域(グラウンドトゥルース)と検出した領域の重なる割合を示す値(IoU)が一定以上

検出モデル	最適化アルゴリズム	AP ₃₀	AP ₆₀	AP ₇₀	AP ₈₀	AP ₉₀
Faster R-CNN	SGDM	0.9454	0.8993	0.7821	0.5328	0.1233
	Adam	0.8046	0.7403	0.5540	0.2406	0.0122
SSD	SGDM	0.8325	0.8272	0.8043	0.6619	0.0949
	Adam	0.8408	0.8397	0.8128	0.6763	0.1781
YOLOv4	SGDM	0.9326	0.9277	0.8891	0.6461	0.0890
	Adam	0.9424	0.9238	0.8544	0.5931	0.0707

表 1: 各カーブミラー検出器の評価結果

であるときを True Positive とする。モデルの評価は IoU が 5 割から 9 割までの各段階で AP (Average Precision) を計算することで行う。AP は検出と分類が正確であるほど 1 に近づく評価指標である。

3 実験結果と考察

各物体検出モデルの性能評価を表 1 に示す。IoU ごとに AP を比較すると、IoU の閾値を 5 割以上とした場合は SGDM を使用した Faster R-CNN が最大 (0.9454) となっている。しかし、閾値を 6 割以上とした段階で SGDM を使用した YOLOv4 が上記の Faster R-CNN を 0.028 ほど上回る結果になった。同様に、閾値が 8 割以上の場合は Adam を使用した SSD が他のモデルを上回っている。

検出のみを考えた場合、AP が最も高い SGDM を使用した Faster R-CNN が最適であるが、鏡面部分の分析を行うことを考慮すれば IoU は 7 割以上であることが好ましい。また、閾値を 8 割以上とすると急激に AP が低下することから、現状では SGDM を使用した YOLOv4 がカーブミラーの検出に最適のモデルだと考えられる。図 3 に上記した YOLOv4 の検出例を示す。黄色の矩形が検出された領域であり、赤色の矩形と数値がグラウンドトゥールズと二つの矩形の IoU である。

最適化アルゴリズムに関して各モデルで比較したところ、オンライン学習を行った Faster R-CNN と YOLOv4 は SGDM を使用した方が性能が高くなり、ミニバッチ学習を行った SSD は Adam の方が性能が高くなる結果となった。従来研究 [6] により Adam は SGDM に比べて良い性能を出すと言われていたが、バッチサイズが非常に小さい場合は SGDM が勝ると考えられる。

4 おわりに

Faster R-CNN, SSD, YOLOv4 の 3 種類の物体検出モデルと SGDM, Adam の 2 種類の最適化アルゴリズムの組み合わせを比較・検討したことで、現段階における本研究の目的に適した YOLOv4 と SGDM の組み合わせを導き出した。図 4 に本研究の応用例として、カーブミラーマップの生成例を示す。しかしながら、本稿で



図 3: カーブミラーの検出例 (YOLOv4)



図 4: カーブミラーマップの生成例

の実験では実行環境の制約により十分に調整できていないミニバッチサイズ等のパラメータも存在している。今後は未検討の物体検出モデルを試行しつつ、データセットの増量や他の深層学習手法による精度向上を目指す。

謝辞: 本研究の一部は科研費 JP21K03962, JP21H01457, JP19K12240, JP22H03700 の助成を受けた。

参考文献

- [1] Yusuke Hino, Shintaro Ono, Noriaki Itagaki, and Yoshihiro Suda. Recognition of risky events reflected in road safety mirror considering ego vehicle's motion. *7th International Symposium on Future Active Safety Technology toward Zero Accidents (FAST-zero)*, 2021.9.
- [2] Rikiya Sato and Kazunori Onoguchi. Scene recognition for blind spot via road safety mirror and in-vehicle camera. In *2022 26th International Conference on Pattern Recognition (ICPR)*, pp. 3714–3720. IEEE, 2022.
- [3] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, Vol. 28, , 2015.
- [4] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. Ssd: Single shot multibox detector. In *European conference on computer vision*, pp. 21–37. Springer, 2016.
- [5] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*, 2020.
- [6] Dami Choi, Christopher J Shallue, Zachary Nado, Jaehoon Lee, Chris J Maddison, and George E Dahl. On empirical comparisons of optimizers for deep learning. *arXiv preprint arXiv:1910.05446*, 2019.