

画像認識のための教師データセット作成における 半自動化方式の評価

天野 隼斗† 大川 猛† 小野 智寛‡* 田向 権‡

東海大学情報通信学部組込みソフトウェア工学科†

九州工業大学 大学院生命体工学研究科‡

日本学術振興会特別研究員 DC*

1. はじめに

画像認識を用いて推論を行うには、画像を教師データとして使用するためのアノテーションという作業が必要である。ディープニューラルネットワークを用いる際の教師データ作成と画像認識における認識精度の担保が困難であった。この課題を解決し、高精度な機械学習モデルを作成するには、多くの時間や人手が必要になる。人を雇い大量のデータセットを作成することが例として挙げられるが、多くの人を雇ったとして、人的コスト・クオリティ・データ管理等についての課題がある。

これらの課題に対し、少ない人的コスト・大量のデータセットを短時間で生成する技術が要求される。これらの課題に対して、アノテーションの作業を半自動化することにより、クオリティが確保された大量のデータセットを生成し、短時間・小工数を実現する手法が提案された[1]。しかし半自動データセット生成方式について生成できる教師データ・学習結果の定量的な評価が少ない。

我々は、自動運転を模した小型ロボットカーの運転画像を題材とし、半自動データセット生成方式の定量的な評価とデータセット生成を目的とした。

2. 従来の手法アノテーションファイル作成方式

本研究の比較対象としている従来手法のアノテーションファイル作成方式とは、

- web カメラを用いて認識したいオブジェクトの動画を撮影する
- 撮影した映像に対し、動画処理ソフトウェアである FFmpeg を使用し、1 秒あたり 30 枚の画像へ変換
- 変換された画像を選別

- 画像内のオブジェクトに対し YOLO を対象とした学習用データセット作成ソフト labelImg を用いアノテーションファイルを作成

3. 半自動データセット生成方式

本研究の評価対象としている生成方式とは、chroma key を用いて撮影対象を回転させ 360° 撮影する。撮影した動画を画像に変換し、背景を透過し生成した画像をオブジェクト画像として用いる。背景として用意した画像に座標を指定する。配置したオブジェクト画像の座標に対応したアノテーションファイルを生成することで、手動で生成するデータセットとそん色ないクオリティを持った大量のデータセット生成を短時間で実現する方式である。教師あり学習の場合アノテーションファイルの作成はすべて手作業で行うことが一般的である。

高精度な機械学習モデルを用意するには、大量の教師データが必要になる。少人数で高い認識精度の機械学習モデルを用意することは困難である。そのため、アノテーションファイルを作成する工数の一部を半自動化し、工数の削減・大量のデータセット生成を実現する半自動生成方式が提案された。

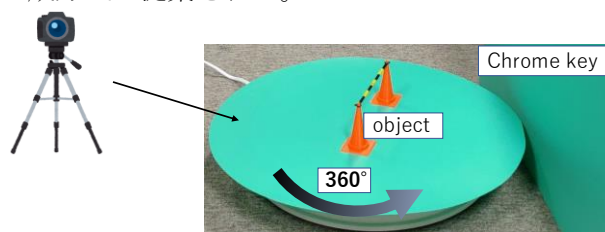


図1: chroma key を用いた撮影図

Evaluation of semi-automated method in creating teacher dataset for image recognition

Hayato Amano†, Takeshi Ohkawa†,

Tomohiro Ono‡*, Hakaru Tamukoh‡,

†Tokai University

‡Kyushu Institute of Technology

*JSPS

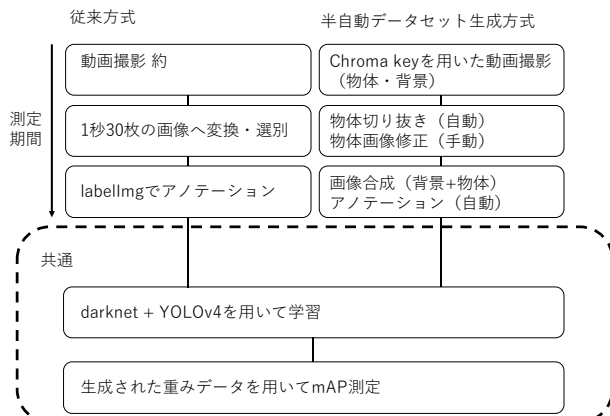


図 2：従来方式と半自動データセット生成方式のフロー比較

4. 評価

半自動データセット生成方式の定量的な評価を行うため、まずは比較対象として従来手法によるアノテーションファイル作成の実験を行った。自動運転を模した小型ロボットカーの運転画像を題材とし人手により、30分もしくは60分間で作成可能なアノテーションファイルの数を計測した。なお本評価では、動画撮影開始からアノテーションファイル作成完了までの時間を測定期間とする(図2)。結果を表1に示す。従来手法の場合、個人差はあるが、1分間に10枚程度作成できることが分かった。

次に半自動データセット生成方式を用いた際の、約60分間で作成できるアノテーションファイルの数・mAPの定量的な評価を行った。テストデータは学習に用いた教師データとは別に50枚ずつ用意した。従来手法については、通常の画像を用いた。一方半自動データセット生成手法については、背景画像+物体の合成画像を用いた。

評価のために用いた開発環境を表2に示す。

表 1:従来手法を用いたアノテーションファイル作成実験結果

	被験者A	被験者B	被験者C
時間(分)	60	30	30
ファイル数	609	401	260

表 2:開発環境

PC環境	OS:Ubuntu 16.04
	CPU:Intel Xeon W-2125
	GPU:NVIDIA RTX 2080
	Mem:DDR4-21300 256GB
使用ソフト	YOLOv4
	darknet

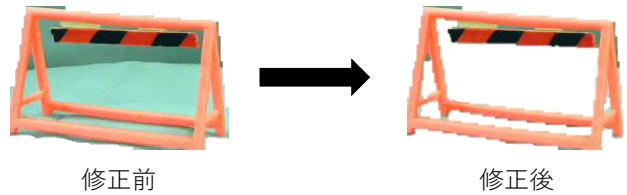


図 3：手作業による修正が必要な画像の例

表 3：測定結果

	従来手法	半自動生成手法
教師データ(クラス数3)	609	1088
作成時間(分)	60	10
画像修正時間(分)	0	46
mAP	73.75	64.71
学習時間(分)	187	253

図3のような対象のオブジェクトには空洞があったため、オブジェクト画像を修正する必要があり、背景部分を手動での透過を行った。

表3に生成時間及びmAPの評価結果を示す。教師データ作成時間に関して、従来手法は1分で約10枚程度作成可能であるのに対して、半自動生成手法は1分で約18枚程度作成可能であるという結果になった。修正時間を考慮しない場合さらに大量のデータセットを生成することが可能といえる。mAPは半自動生成手法よりも従来手法の方が約11ポイント高い。アノテーションファイル数は従来手法より、1.8倍の量アノテーションファイルを生成出来た。

5. おわりに

本稿は、高精度な機械学習モデルで使用するアノテーションファイル生成を半自動化する方式について定量的な評価を行った。

結果として、半自動生成手法は従来手法に対して約11ポイントmAPが低い。半自動生成手法は短時間で大量の教師データを生成できる可能性を示した。

参考文献

- [1] Ishida, Yutaro, and Hakaru Tamukoh. "Semi-Automatic Dataset Generation for Object Detection and Recognition and its Evaluation on Domestic Service Robots." *Journal of Robotics and Mechatronics* 32.1 (2020): 245-253.
- [2] Redmon, Joseph. "Darknet: Open source neural networks in c." (2013): 2018.
- [3] 田村爽, et al. "ROS ベースの自律移動ロボットにおける FPGA 統合開発プラットフォーム." *研究報告システム・アーキテクチャ (ARC)* 2019.8 (2019): 1-6.