

画像分類モデルを用いた物体検出用オートアノテーション処理手法の検討

A Study on Auto-Annotation in Object Detection Using Image Classification Models

○山田駿佑¹, 関弘翔², 細野裕行²

*Shunsuke Yamada¹, Hiroto Seki², Hiroyuki Hosono²

Abstract: This study examines an auto-annotation method using image classification models. We investigated a method to determine the location of objects by visualizing the basis for decisions using Grad-CAM, an Explainable AI for CNN image classification models.

1. 研究背景

近年、物体検出技術が様々な場面で使われている。物体検出で重要視されるものとしてアノテーションが挙げられる。物体検出におけるアノテーションとは、AIを学習するための教師データとして、画像内のオブジェクトに対する位置と種類のラベル付け作業のことを指す。この作業は物体検出AIを学習するうえで必須だが、人が手作業で行うと1つのオブジェクトあたり平均50.8秒かかるとされ^[1]、データセットの規模によってはアノテーションの時間と労力が膨大になる。

そこで本研究では、オートアノテーションの手法に着目する。現在オートアノテーションの手法として挙げられるものの多くは画像に属性を付与するような画像分類用のアノテーションであり、座標を決定するような物体検出用のオートアノテーションではない。

本稿では、物体検出よりも教師データの作成コストが低い画像分類を行うAIモデルを応用したオートアノテーションを検討する。

2. 報告内容

まず画像分類AIを学習し、説明可能なAIの一種であるGrad-CAM^[2]を用いて、画像分類を行う際に特徴となる領域をヒートマップとして可視化する。可視化したヒートマップに対して2値化を行い、連結する画素塊をラベリングにより抽出することでオブジェクトの座標決定し、オートアノテーションを実現する。本稿における2値化の閾値決定法にはトライアングルアルゴリズム^[3]を採用した。

結果は、オートアノテーション処理による輪郭データと手作業アノテーションでの輪郭データの重なり度合いをIoU(intersection over union)で確認し評価する。IoUは2つの領域の重なり度合いを示しており、完全一致の際に最大値の1をとる。



(a) heatmap

(b) auto-annotation result

Figure 1. heatmap and auto-annotation result

本稿では、画像分類モデルとして、ImageNet^[4]で学習済みの152層から成るCNNであるResnet152^[5]を用いた。オートアノテーションの対象物体としては、ImageNetの学習クラスに含まれる蟻を採用した。

Fig. 1(a)にGrad-CAMを用いて作成したヒートマップと元画像を重ねたものを示す。Fig.1(b)に、ヒートマップに対して2値化したのちにラベリングにより輪郭データを推定し、作成したオートアノテーション結果を示す。Fig. 1のIoUの値は72.9%であった。Fig.1では、概ね検出対象物体を囲むことができていることが分かる。

画像データ72枚に対して同様にIoUを確認したところ、平均値が33.8%、標準偏差が0.20という値になり、精度向上が必要である。

3. まとめ

本研究ではCNNの画像分類モデルを応用した、物体検出のためのオートアノテーション手法について検討した。

参考文献

- [1] H. Su, et al., In proc. of the AAAI Technical Report, 4th Human Computation Workshop (2012)
- [2] S. Ramprasaath R, et al., In proc. of IEEE CVPR 2017, pp.618-626 (2017)
- [3] G. W. Zack, et al., Journal of Histochemistry & Cytochemistry, 25(7):741, 1, 2.11 (1977)
- [4] J. Deng, et al., In proc. of IEEE CVPR 2009, pp.248-255 (2009)
- [5] K. He, et al., In proc. of IEEE CVPR 2016, pp.770-778 (2016)