

物体検出に用いられるニューラルネットワークモデル

最新モデルのサーベイと目的に応じたモデルの選択

Neural Network Models for Object Detection

A Survey of the Latest Models and Optimal Model Selections for Specific Tasks

金子 純也
Junya Kaneko

Morning Project Samurai 株式会社
Morning Project Samurai Inc.
junya@mpsamurai.com, <http://www.mpsamurai.com>

山田 貢己
Miki Yamada

(同上)
m.yamada@mpsamurai.com

keywords: survey, neural network, object detection, instance segmentation, deep learning

Summary

「ショートノート」は 200 ワード、それ以外は 200 ~ 500 ワード以内の英文で summary を記す (ここは、論文執筆後に書く。)

ここにチートシートを出力する。

1. ま え が き

作成中。

2. 物体検出 (Object detection)

2.1 Two-stage 検出器

§ 1 Faster R-CNN

§ 2 TFA Net

§ 3 Few-Shot Object Detection

2.2 One-stage 検出器

§ 1 YOLOv4

§ 2 EfficientDet

3. インスタンスセグメンテーション (Instance segmentation)

3.1 Mask Scoring R-CNN (MS R-CNN)

マスク品質 (インスタンスマスクと正解マスクとの IoU として定量化されるもの) を分類スコアと明示的に関連付けたモデルである [Huang 19]。MS R-CNN は、予測マスクの品質を学習するためのブロック (MaskIoU Head) を、Mask R-CNN[He 17] に導入したモデルになっている (図 1)。MaskIoU Head はインスタンスの特徴量と対応する予測マスクを一緒に取り込み、それを元に Mask IoU を回帰推定する。そして、推論時に予測 MaskIoU を分類スコアに掛け算して補正する。

§ 1 MS R-CNN の学習

学習サンプルとして RPN proposals を使う。proposal box と正解 box との IoU が 0.5 以上の学習サンプルが必要となる。これは Mask R-CNN の Mask head の学習サンプルの場合と同じである。各学習サンプルに対する回帰目標を生成するために、まず目標クラスの予測マスクを取得し、予測マスクを閾値=0.5 で 2 値化する。そして、2 値化マスクと正解との MaskIoU を使う。MaskIoU を回帰するのは L2 損失を使い、損失重みは 1 にする。ネットワーク全体は end-to-end で学習する。

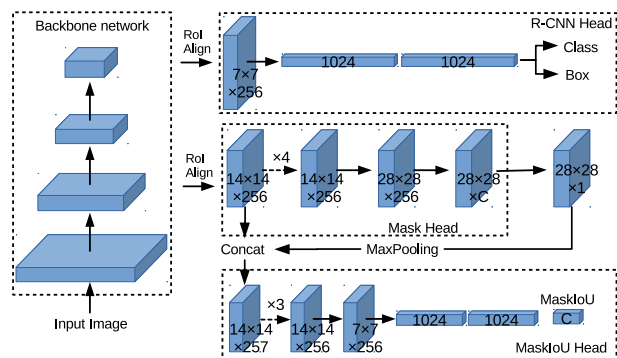


図 1 Mask Scoring R-CNN の構造。

§ 2 MS R-CNN の推論処理

MaskIoU Head は分類スコア (R-CNN head の出力) の調整に使う．推論の手順は次のようになる：

- (1) R-CNN head が N 個の bounding box を出力する．
- (2) N 個の bounding box のうち，SoftNMS[Bodla 17] で上位 k 個のボックスを選択する．
- (3) 上位 k 個のボックスを Mask Head に入力し， k 個のマルチクラスマスクを生成する（ここまでは標準的 Mask R-CNN の手順）．
- (4) これら k 個のマスクを目標として MaskIoU Head に入力し，予測 MaskIoU を出力する．
- (5) 予測 MaskIoU を，分類スコアに掛け算し，上位 k 個の修正された分類スコアを得る．

3.2 YOLACT++

4. パノプティックセグメンテーション (Panoptic(?) segmentation)

5. む す び

謝 辞

謝辞について

◇ 参 考 文 献 ◇

- [Bodla 17] Bodla, N., Singh, B., Chellappa, R., and Davis, L. S.: Soft-NMS Improving Object Detection with One Line of Code, in *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 5562–5570 (2017)
- [He 17] He, K., Gkioxari, G., Dollr, P., and Girshick, R.: Mask R-CNN, in *2017 IEEE International Conference on Computer Vision (ICCV), Venice*, pp. 2980–2988 (2017)
- [Huang 19] Huang, Z., Huang, L., Gong, Y., Huang, C., and Wang, X.: Mask Scoring R-CNN, in *Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6409–6418 (2019)

〔担当委員：××〕

19YY 年 MM 月 DD 日 受理

◇ 付 録 ◇

A. 付録のタイトル 1

付録の本文 1

著 者 紹 介

金子 純也(正会員)

著者 1 の略歴

山田 貢己(正会員)

1989 年東京大学大学院物理学専攻修了，理学博士．同年株式会社東芝入社．ニューラルネットワークの研究開発，セキュリティ技術，画像認識技術，テレビの高画質化技術，車載画像認識プロセッサ等の開発業務に従事．2020 年ジャパニクス株式会社に入社．現在，Morning Project Samurai 株式会社において AI 開発業務に従事．