

Hierarchical Head Design for Object Detectors

Shivang Agarwal and Frédéric Jurie

Normandie Univ, UNICAEN, ENSICAEN, CNRS — UMR 6072 GREYC shivang.agarwal@unicaen.fr



Introduction

- In object detection, anchors are the reference boxes having multiple pre-defined positions, scales and aspect ratios, which are used to classify as well as localize target object candidates of different sizes and shapes all over the image.
- The motivation of this paper comes from the observations that
 - 1. Each anchor learns to classify and regress candidate objects independently.
 - 2. Insufficient examples are available for each anchor in case of small-scale datasets.
- We propose a novel hierarchical head for the SSD detector. It allows anchors to assist each other. Each anchor can embed some specific information, not shared with other anchors,





jointly with some more general information shared with other anchors.

• State-of-the-art performance on FlickrLogos-47 and improved performance on PASCAL-VOC validate the method. The proposed design does not give additional performance gain over the original design in case of large scale datasets like MS-COCO.

Equations

General Head

 $\begin{cases} c_{a_i} = f^c (W_{a_i}^c x + b_{a_i}^c) \\ p_{a_i} = f^p (W_{a_i}^p x + b_{a_i}^p) \end{cases}$

Hierarchical Head

$$\begin{cases} c_{a_i} = f^c (W_{a_i}^c x + b_{a_i}^c + \sum_{j=1}^J (W_{R_j}^c x + b_{R_j}^c) \times \mathbf{M}_{ji}) \\ p_{a_i} = f^p (W_{a_i}^p x + b_{a_i}^p + \sum_{j=1}^J (W_{R_j}^p x + b_{R_j}^p) \times \mathbf{M}_{ji}) \end{cases}$$

Hierarchical Head





Figure 2: Graphical presentation of the proposed hierarchical head. In the configuration shown, R_2, R_3 and R_4 are responsible for learning common characteristics between anchors of different aspect ratios and same size. R_1 is responsible for learning common characteristics for all types of anchors.

Results on Pascal-VOC

Table 2: Mean Average Precision performance, reported on PASCAL-VOC07 test set (trained on VOC07+12 trainval set), for all the configurations.

Configuration	SSD300	SSD512
9An{9}	72.32	75.74
$9An\{9,1\}Both$	72.75	76.63
$9An\{9,3\}S$	72.68	76.27
$9An\{9,3\}As$	73.29	76.33
$9An\{9,3,1\}S$	72.71	76.60
$9An\{9,3,1\}As$	73.15	76.14

Conclusion

A novel hierarchical head is proposed:

• Anchors are able to gain from the training of other anchors when the data is scarce.

Acknowledgement This project was partly funded by the DGA through the RAPID-DRAAF project.

• The proposed method has no extra features at inference time.

• The proposed head can be used with different scales of datasets with performance gains to be expected in small-scale datasets.