Anchor free object detection method using mask optimization

CURRENT STATUS: UNDER REVIEW

EURASIP Journal on Image and Video Processing Springer

Beibei Fan
Shanghai University

fanbeibei@shu.edu.cn*Corresponding Author*

HE Yang
Shanghai University

Ling ling Guo
Shanghai University

DOI:
10.21203/rs.2.24760/v1

SUBJECT AREAS
Mechanical Engineering

KEYWORDS
object detection, anchor-free, convolutional neural network, attention mechanism

Abstract

The anchor-free method based on key point detection has made great progress. However, the anchor-free method is too dependent on using a convolutional network to generate a rough heat map. It is difficult for the network to detect objects whose shape changes greatly, and small objects are difficult to detect. In order to solve this problem, first of all, we use the most advanced mask attention mechanism algorithm in the network to increase the accuracy of the thermodynamic generation of network detection. Then, We also designed an optimized fire model to reduce the size of the model. The masking mechanism optimizes the feature map of the network to enhance the detection capability in network space and improve the accuracy of heatmaps generation. Our approach achieved an accuracy of 91.84% and a recall 89.83% in the Tencent-100K dataset. Our approach is also competitive with the most advanced approach.

Full Text

Due to technical limitations, full-text HTML conversion of this manuscript could not be completed. However, the manuscript can be downloaded and accessed as a PDF.

Figures

Figure 1

Illustrated some test results. Our system works very well in complex scene where objects are small and highly occluded.

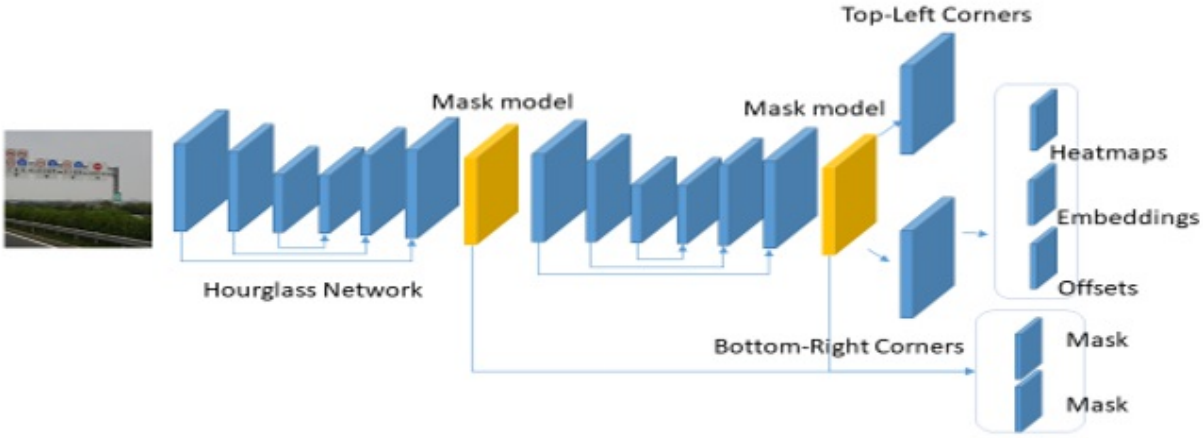

Figure 2

Describes the overall flow of our method. By stacking multiple Hourglass Network and mask modules to enhance

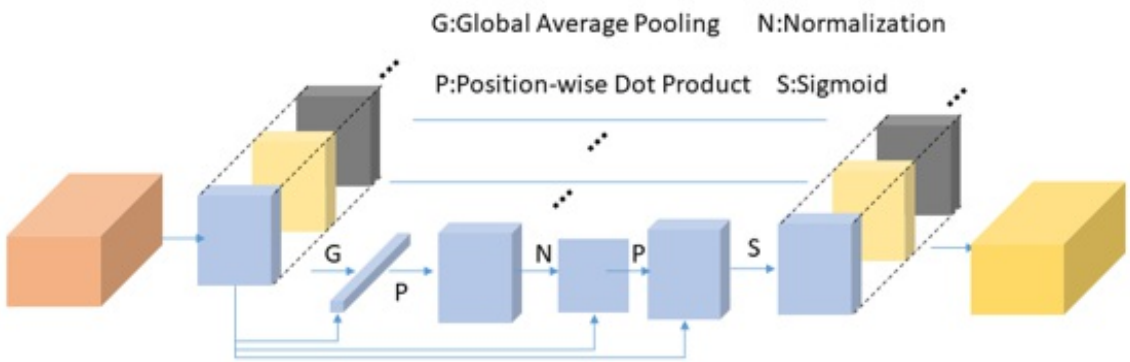

Figure 3

Diagram our improved fire module, which combines global and local features by grouping for incoming feature maps.

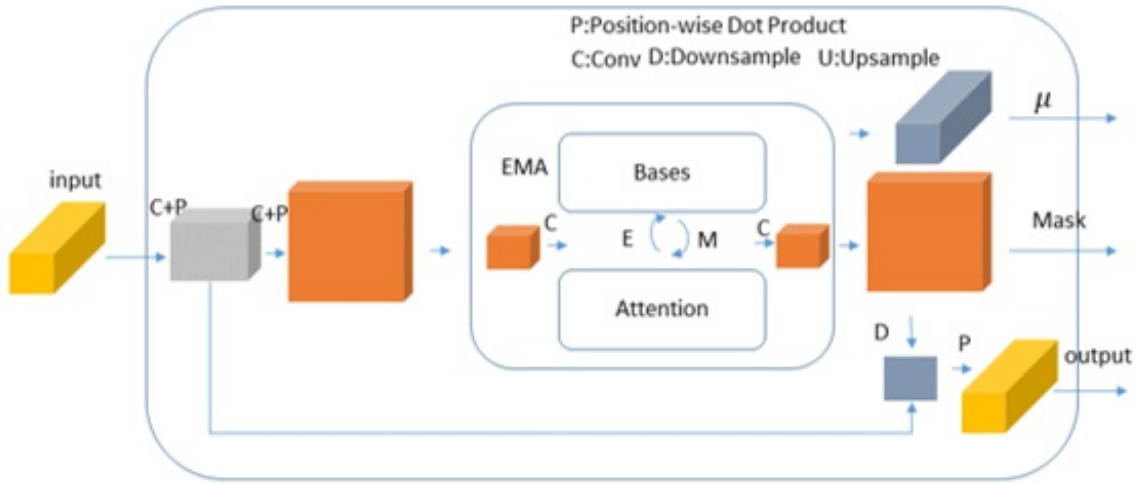

Figure 4

The mask module is shown in detail, and the feature is upsampled by multiple convolution kernels. Using the EMA module to estimate the mask, the mask module can get the segmentation result of the object. Aggregating the mask module can generate spatial constraints on the object and enhance local selection of features.

Figure 5

The data tag is displayed, the left side is the mark for coordinate positioning, and the right side is the label for image segmentation.

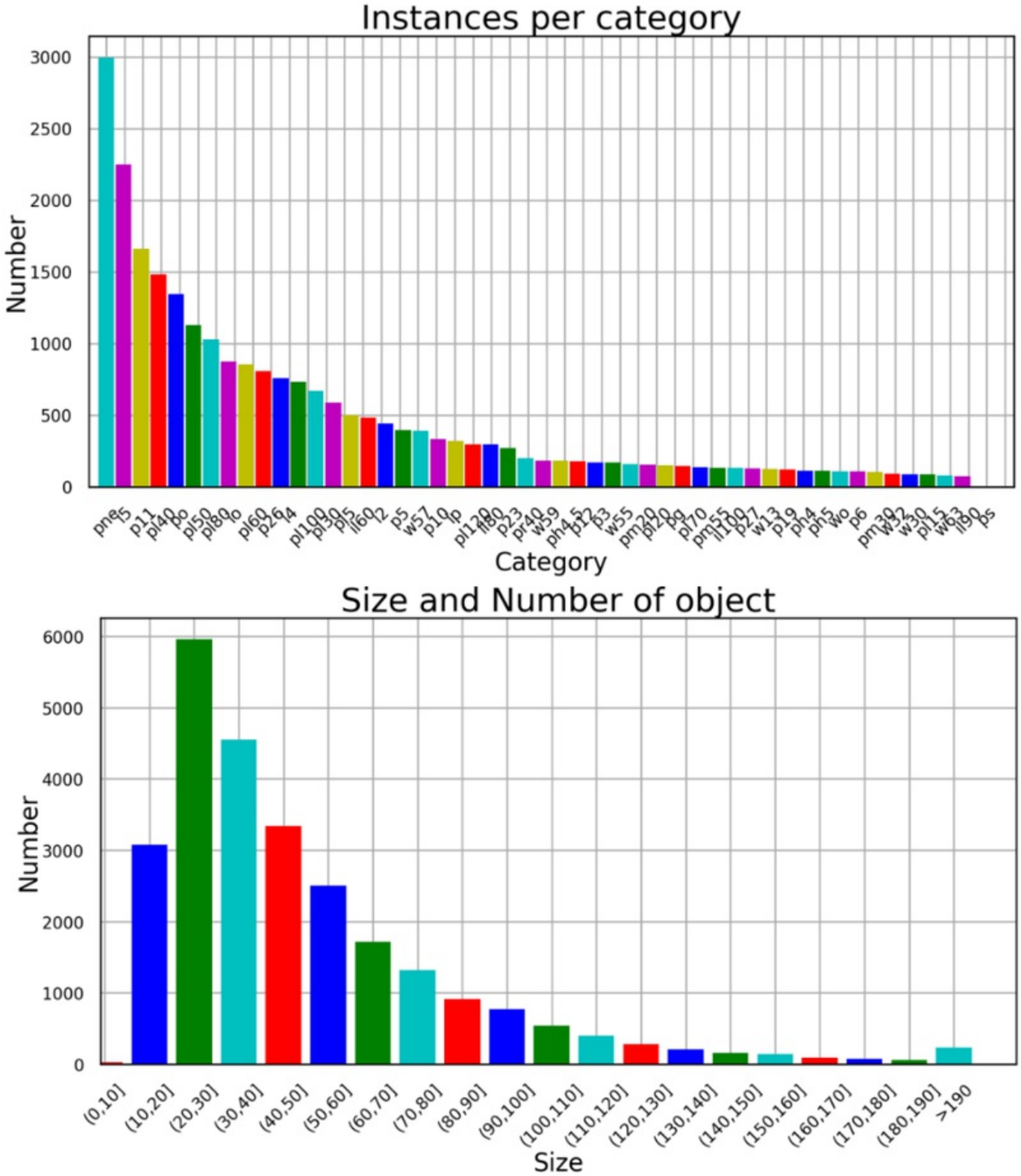

Figure 6

Object size and category statistics

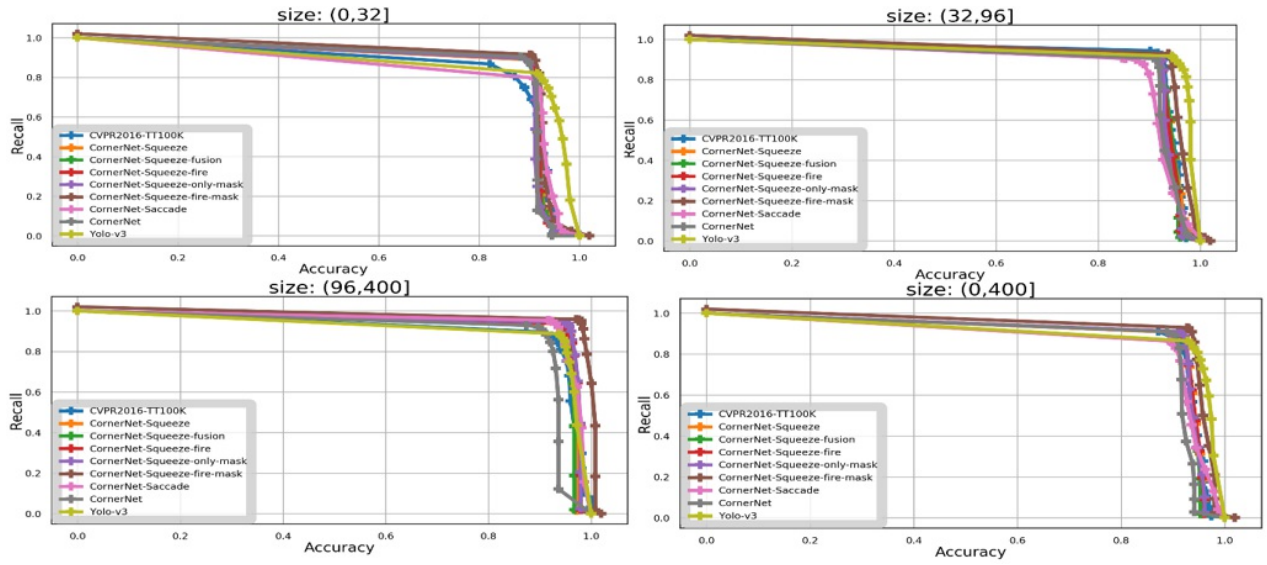

Figure 7

The performance of the proposed method is on the three scales of large, medium and small.

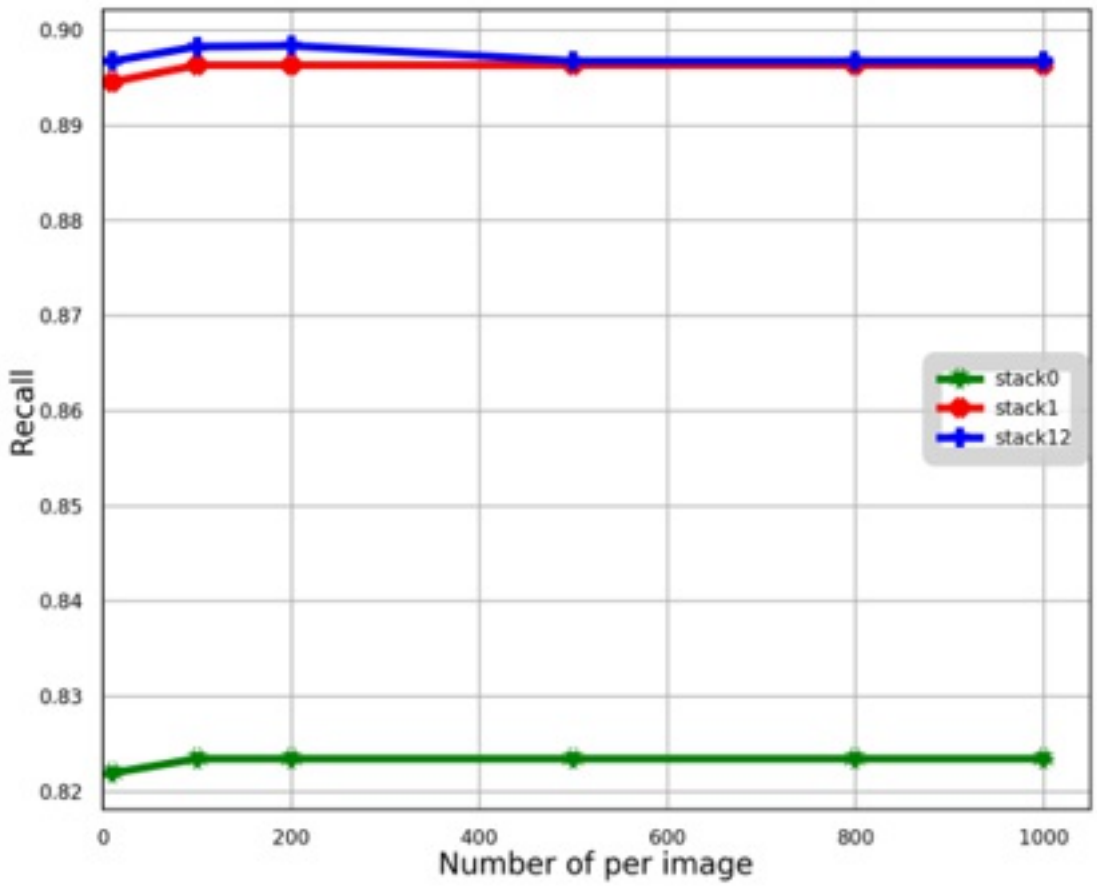

Figure 8

Diagram the effect of the recommended number

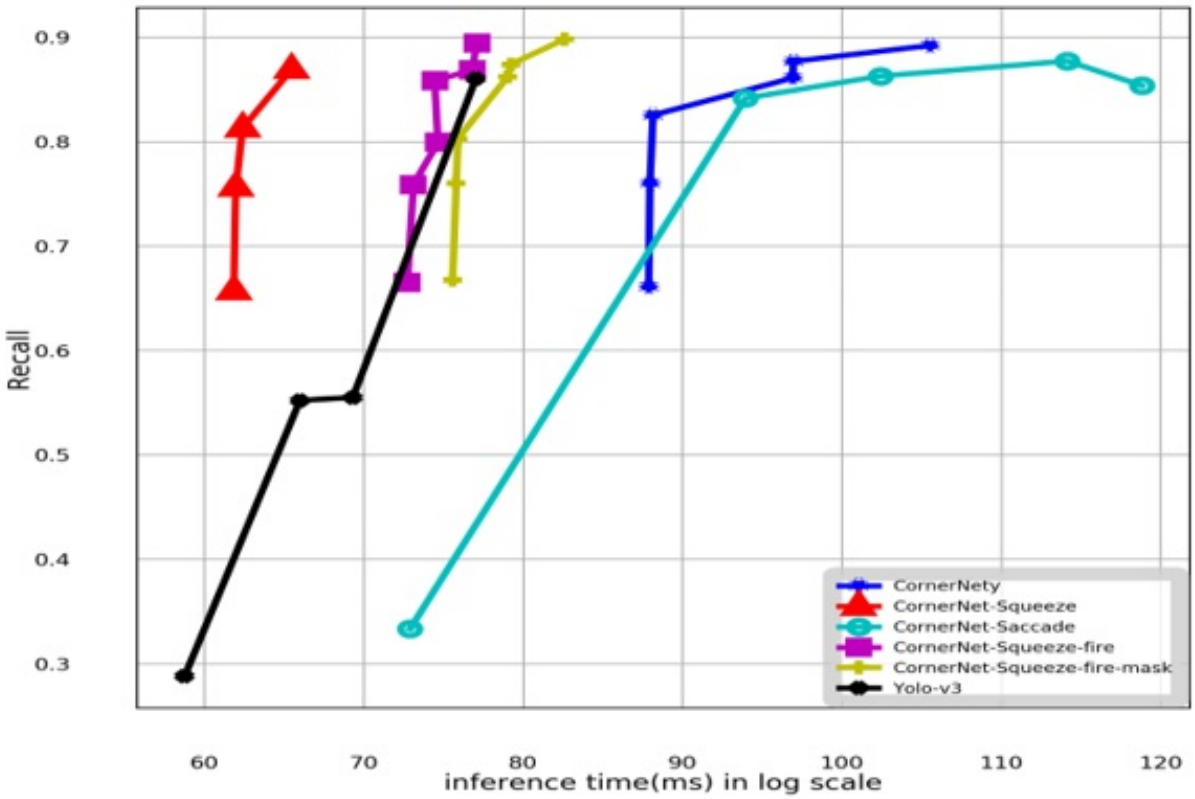

Figure 9

Diagram the effect of the recommended number on the recall.