Review for CenterMask : single shot instance segmentation with point representation

Can Koz

May 23, 2021

1 Summary

Instance segmentation is a task in computer vision which aims to locate classify and segment each instance in the image. Although one stage object detectors gained popularity over two stage detectors, there are few papers that focus on one stage instance segmentation. Authors aim to design a one stage and anchor box free instance segmentation model in this paper. This paper the authors introduce CenterMask which achieves a good performance on COCO dataset and they indicate that this can be easily integrated with existing object detectors.

2 Strengths

- A simple, fast and accurate design is proposed for instance segmentation. While being a single stage method, it shows good speed-accuracy trade off.

- Local shape representation allows more flexible representation for different instances so that different instances can be separated.

- The Global Saliency Map is a simple, fast and precise alternative for pixelwise feature alignment.

3 Weaknesses

- Although the method performs comparable among the one-stage approaches, the performance must be improved, since there exists better performing one-stage methods.

- The method is evaluated on a single dataset and it needed to be experimented on diverse datasets for analyzing generalization of the method.

4 Evaluation

The performance of the proposed CenterMask is evaluated on the MS COCO instance segmentation. Ablation studies are performed on shape size selection, backbone architecture, local shape branch, global saliency branch and combination of local shape and global saliency. Results are compared based on AP score. There are two models that achieves better than the proposed method, two-stage Mask R-CNN and one-stage TensorMask. Among one-stage algorithms, TensorMask performs the best but it is 5 times slower than the proposed method.

5   Final Comments and Future Work

Instead of multi stage approaches, one stage approaches promises simpler and more effective solutions to many computer vision problems while it requires well reasoned design. In the experiments, robustness and effectiveness of the method has been demonstrated. As mentioned in the experiments parts TensorMask achieves a better performance and the performance of CenterMask can be improved by sacrificing the runtime.