Mask R-CNN

Publisher: IEEE Cite This PDF

Kaiming He ; Georgia Gkioxari ; Piotr Dollár ; Ross Girshick **All Authors**

19518 Cites in Papers

40357 Full Text Views

Abstract

Document Sections

1. Introduction
2. Related Work
3. Mask R-CNN
4. Experiments: Instance Segmentation
5. Mask R-CNN for Human Pose Estimation

Authors

Figures

References

Citations

Keywords

Metrics

More Like This

Footnotes

Abstract:
We present a conceptually simple, flexible, and general framework for object instance segmentation. Our approach efficiently detects objects in an image while simultaneously generating a high-quality segmentation mask for each instance. The method, called Mask R-CNN, extends Faster R-CNN by adding a branch for predicting an object mask in parallel with the existing branch for bounding box recognition. Mask R-CNN is simple to train and adds only a small overhead to Faster R-CNN, running at 5 fps. Moreover, Mask R-CNN is easy to generalize to other tasks, e.g., allowing us to estimate human poses in the same framework. We show top results in all three tracks of the COCO suite of challenges, including instance segmentation, bounding-box object detection, and person keypoint detection. Without tricks, Mask R-CNN outperforms all existing, single-model entries on every task, including the COCO 2016 challenge winners. We hope our simple and effective approach will serve as a solid baseline and help ease future research in instance-level recognition. Code will be made available.

Published in: 2017 IEEE International Conference on Computer Vision (ICCV)

Date of Conference: 22-29 October 2017

DOI: 10.1109/ICCV.2017.322

Date Added to IEEE *Xplore*: 25 December 2017

Publisher: IEEE

ISBN Information:

Conference Location: Venice, Italy

Electronic ISSN: 2380-7504

1. Introduction

The vision community has rapidly improved object detection and semantic segmentation results over a short period of time. In large part, these advances have been driven by powerful baseline systems, such as the Fast/Faster R-CNN [9], [29] and Fully Convolutic] frameworks for object detection and semantic segmentation, respectively. These met ntuitive and offer flexibility and robustness, together with fast training and inference time. Our goal in this work is to develop a comparably enabling framework for instance segmentation.

Continue Reading

Authors

Figures

References

Citations

Keywords

Metrics

Footnotes

ng [MathJax]/extensions/MathZoom.js

Samsara EX1030
U.S. Patent No. 11,875,580

IEEE Personal Account

CHANGE USERNAME/PASSWORD

Purchase Details

PAYMENT OPTIONS

VIEW PURCHASED DOCUMENTS

Profile Information

COMMUNICATIONS PREFERENCES

PROFESSION AND EDUCATION

TECHNICAL INTERESTS

Need Help?

US & CANADA: +1 800 678 4333

WORLDWIDE: +1 732 981 0060

CONTACT & SUPPORT

Follow

About IEEE *Xplore* | Contact Us | Help | Accessibility | Terms of Use | Nondiscrimination Policy | IEEE Ethics Reporting | Sitemap | IEEE Privacy Policy

A public charity, IEEE is the world's largest technical professional organization dedicated to advancing technology for the benefit of humanity.

© Copyright 2025 IEEE - All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies.

Loading [MathJax]/extensions/MathZoom.js