Team Solaria COMP9517 Project

Yishun Jin
yishun.jin@unsw.edu.au
z5235653

Wei Wang

xxx
xxx@unsw.edu.au

I. Introduction

intro for dataset

II. Literature Review

Image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as image objects). [1] The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics.

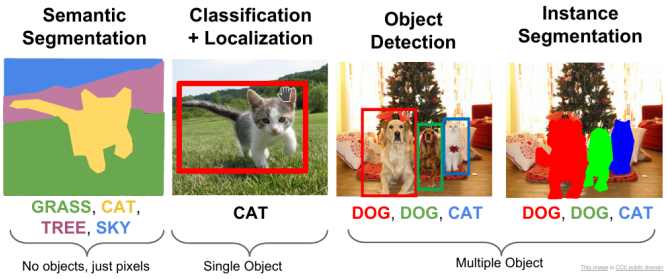


Fig. 1: Object Detection and Segmentation

Instance segmentation is an approach that identifies, for every pixel, a belonging instance of the object. It detects each distinct object of interest in the image. For example, in this project, each leaf in the picture is segmented as an individual object. [2]

III. Methods

A. Task 1

B. Task 2

C. Task 3 Instance Segmentation

It is a multi-instance segmentation problem. We use **Detectron 2** [3] which is FAIR's next-generation platform for object detection and segmentation. And we use **Mask-RCNN** [4] to approach our instance segmentation.

Mask R-CNN is a state-of-the-art model for instance segmentation. Mask-RCNN can be easily considered as ResNet-FPN + Fast-RCNN + mask.

To understand Mask R-CNN, it works mainly in two stages:

Stage1: The first stage consists of two networks, backbone (ResNet, VGG, Inception, etc..) and region proposal network.

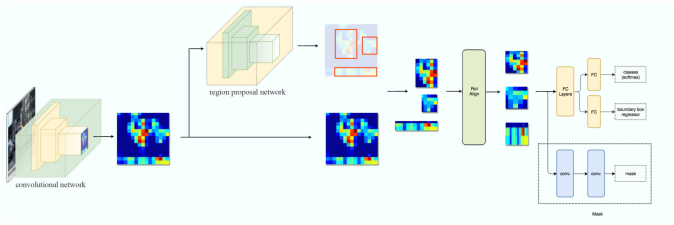


Fig. 2: Mask R-CNN

These networks run once per image to give a set of region proposals. Region proposals are regions in the feature map which contain the object.

Stage2: In the second stage, the network predicts bounding boxes and object class for each of the proposed region obtained in stage1. Each proposed region can be of different size whereas fully connected layers in the networks always require fixed size vector to make predictions. Size of these proposed regions is fixed by using either RoI pool (which is very similar to MaxPooling) or RoIAlign method.

For our implementation, we convert *_label.png into mask image and write a adapter for loading dataset.

Because our dataset is not large, we use transfer learning [5] for training process. Use the provided pre-trained weight for MS COCO dataset [6].

IV. Experimental Setup

IoU / Dice

V. Results and Discussion

A. Task 1

B. Task 2

C. Task 3 Instance Segmentation

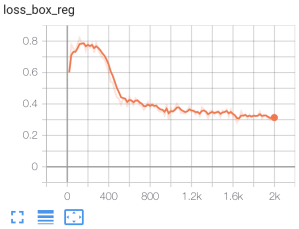
We use TensorBoard to visualize our training process data. Shown on Fig.3

VI. Conclusion

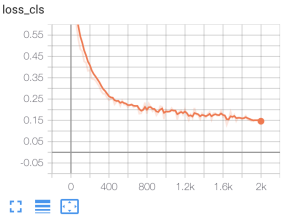
The conclusion goes here.

Acknowledgment

The authors would like to thank...



(a) loss bbox reg



(b) loss classification

(c) loss mask

Fig. 3: loss

VII. Contribution of Group Members

References

[1] R. M. Haralick and L. G. Shapiro, "Image segmentation techniques," *Computer vision, graphics, and image processing*, vol. 29, no. 1, pp. 100–132, 1985.

[2] A. M. Hafiz and G. M. Bhat, "A survey on instance segmentation: state of the art," *International Journal of Multimedia Information Retrieval*, vol. 9, no. 3, p. 171–189, Jul 2020. [Online]. Available: http://dx.doi.org/10.1007/s13735-020-00195-x

[3] Y. Wu, A. Kirillov, F. Massa, W.-Y. Lo, and R. Girshick, "Detectron2," https://github.com/facebookresearch/detectron2, 2019.

[4] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask r-cnn," *2017 IEEE International Conference on Computer Vision (ICCV)*, Oct 2017. [Online]. Available: http://dx.doi.org/10.1109/ICCV.2017.322

[5] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[6] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," *Lecture Notes in Computer Science*, p. 740–755, 2014. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-10602-1_48