

CS280 - Final Project Proposal

Xiangyi Zhang
Shanghaitech University
zhangxy9@shanghaitech.edu.cn

We propose to work on Amodal segmentation, which tackles the problem that predicts the complete shape of an object encompassing both its visible and occluded parts. To be more specific, the problem requires a model which can predict the missing information. This problem is very interesting, and it could be used to preceive the whole object when occluded and would be capable of providing valuable insights into human perception.

To get more context and background of this problem, we plan to read the related work [2],[3],[4]. And we will use the Cityscape amodal dataset that has both visible and amodal object masks annotated.

Amodal instance segmentation has gained increasing attention but it still remains problems of missing information, occlusion and expensive labeling. There is a realdata benchmark [4] and a synthetic benchmark[1], providing large-scale, accurate training data is still challenging.

Based on these challenges, we want to propose a novel semi-supervised model approach to infer multiple amodal masks from a single RGB image, which consists of three parts: A reference set constructed from existing amodal data, visual concept based network for similarity and a generative model to infer multiple solution.

References

- [1] K. Ehsani, R. Mottaghi, and A. Farhadi. Segan: Segmenting and generating the invisible. *arXiv preprint*, 2017.
- [2] K. He, G. Gkioxari, P. Dollár, and R. B. Girshick. Mask r-cnn. *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2980–2988, 2017.
- [3] Z. Zhang, C. Xie, J. Wang, L. Xie, and A. L. Yuille. Deepvoting: A robust and explainable deep network for semantic part detection under partial occlusion. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1372–1380, 2018.
- [4] Y. Zhu, Y. Tian, D. N. Metaxas, and P. Dollár. Semantic amodal segmentation. In *CVPR*, volume 2, page 7, 2017.