References

Here we list papers referred to in this part of the project. It mainly involves three types of papers, that is object detection, semantic segmentation and their backbone network.

Object detection

- Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." *Advances in neural information processing systems* 28 (2015): 91-99.
- Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." *arXiv preprint arXiv:1804.02767* (2018).
- Ge, Zheng, et al. "Yolox: Exceeding yolo series in 2021." *arXiv preprint arXiv:2107.08430* (2021).

Semantic segmentation

- Strudel, Robin, et al. "Segmenter: Transformer for Semantic Segmentation." *arXiv preprint arXiv:2105.05633* (2021).
- He, Kaiming, et al. "Mask r-cnn." *Proceedings of the IEEE international conference on computer vision*. 2017.

Backbone

- Liu, Ze, et al. "Swin transformer: Hierarchical vision transformer using shifted windows." *arXiv preprint arXiv:2103.14030* (2021).
- He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.