EE 599 Deep Learning – Initial Project Proposal

Multi Label Fully Convolutional Network for Segmentation

Yi Zheng, Ruyi Zhang, Yifan Wang

{zhen030, ruyizhan, wang608}@usc.edu

April 1, 2019

Project Summary: In this project we intend to improve image segmentation result based on fully convolutional network by adding segmentation boundary ground truth when computing loss function. In previous work, segmentation gives excellent performance on big objects. However, results on object boundary are not satisfactory. In that case, boundary label is used to enhance segmentation. Loss function is computed with combination of loss from both pixel wise segmentation labels and their boundaries. We are expected to see a higher segmentation accuracy. This method can be explained as emphasis more on segmentation boundaries.

Data Needs and Acquisition Plan: Our network would based on structure in [**?**, **?**] with code realization from [**?**]. Some modification is base on this code. Ground truth boundary can be derived from segmentation labels. Basically, our training and test would based on Pascal VOC 2012 dataset [**?**]. If time admits, more test would be carried out on NYUDv2 dataset [**?**]. Evaluation would use pixel accuracy, mean accuracy, mean IU and frequency weighted IU according to [**?**].

Primary References and Codebase:

[1] Shelhamer E , Long J , Darrell T . Fully Convolutional Networks for Semantic Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016:1-1.

[2] Long J , Shelhamer E , Darrell T . Fully Convolutional Networks for Semantic Segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 39(4):640-651.

[3] Pascal VOC 2011: `http://host.robots.ox.ac.uk/pascal/VOC/voc2012/`

[4] NYUDv2: `https://cs.nyu.edu/~silberman/projects/indoor_scene_seg_sup.html`

[5] GitHub: `https://github.com/aurora95/Keras-FCN`

Architecture Investigation Plan: Architecture is based on FCN-8s citec1 which is based on FCN VGG-16 with pre-trained weights.

Estimated Compute Needs: According to [**?**], fine tune these network would need about one day on a single GPU. Due to thr fact the this paper is published in 2014, we are confident that it would take less than one day according to the power of GPUs today.

Team Roles: The following is the rough breakdown of roles and responsibilities we plan for our team:

- Yifan Wang: Data collection and cleaning, video production. Code bi-directional GRU.

- Yi Zheng: Get existing codebase running on an AWS instance.

- Ruyi Zhang: Data Augmentation.

All team members will work on the final presentation, slides, and report.

Requested Mentor with Rationale: We request Jiali to be our team mentor because he has expertise in image processing. We have already had a talk with Jiali about our work and gotten some advice from him.