Notice of the Final Oral Examination
for the Degree of Master of Science

of

ABHISHAKE KUMAR BOJJA

BTech (Indian Institute of Technology, Indian School of Mines, Dhanbad, 2015)

“Deep Neural Networks for Semantic Segmentation”

Department of Computer Science

Wednesday, April 1, 2020
2:00 P.M.
Remote Defence

Supervisory Committee:
Dr. Kwang Moo Yi, Department of Computer Science, University of Victoria (Co-Supervisor)
Dr. Andrea Tagliasacchi, Department of Computer Science, UVic (Co-Supervisor)

External Examiner:
Dr. Madeleine McPherson, Department of Civil Engineering, UVic

Chair of Oral Examination:
Dr. Raad Nashmi, Department of Biology, UVic

Dr. David Capson, Dean, Faculty of Graduate Studies
Abstract

Segmenting image into multiple meaningful regions is an essential task in Computer Vision. Deep Learning has been highly successful for segmentation, benefiting from the availability of the annotated datasets and deep neural network architectures. However, depth-based hand segmentation, an important application area of semantic segmentation, has yet to benefit from rich and large datasets. In addition, while deep methods provide robust solutions, they are often not efficient enough for low-powered devices. In this thesis, we focus on these two problems. To tackle the problem of lack of rich data, we propose an automatic method for generating high-quality annotations and introduce a large-scale hand segmentation dataset. By exploiting the visual cues given by an RGBD sensor and a pair of colored gloves, we automatically generate dense annotations for two-hand segmentation. Our automatic annotation method lowers the cost/complexity of creating high-quality datasets and makes it easy to expand the dataset in the future. To reduce the computational requirement and allow real-time segmentation on low power devices, we propose a new representation and architecture for deep networks that predict segmentation maps based on Voronoi Diagrams. Voronoi Diagrams split space into discrete regions based on proximity to a set of points making them a powerful representation of regions, which we can then use to represent our segmentation outcomes. Specifically, we propose to estimate the location and class for these sets of points, which are then rasterized into an image. Notably, we use a differentiable definition of the Voronoi Diagram based on the softmax operator, enabling its use as a decoder layer in an end-to-end trainable network. As rasterization can take place at any given resolution, our method especially excels at rendering high-resolution segmentation maps, given a low-resolution image. We believe that our new HandSeg dataset will open new frontiers in Hand Segmentation research, and our cost-effective automatic annotation pipeline can benefit other relevant labeling tasks. Our newly proposed segmentation network enables high-quality segmentation representations that are not practically possible on low power devices using existing approaches.