

Speech Title: Semantic Scene Segmentation by Deep Machine Learning

Abstract: Scene segmentation is a challenging task as it need classify every pixel in the image. It is crucial to exploit discriminative context and aggregate multi-scale features to achieve better segmentation. Context is essential for semantic segmentation. Due to the diverse shapes of objects and their complex layout in various scene images, the spatial scales and shapes of contexts for different objects have very large variation. It is thus ineffective or inefficient to aggregate various context information from a predefined fixed region. In this talk, I will first present a novel context contrasted local feature that not only leverages the informative context but also spotlights the local information in contrast to the context. The proposed context contrasted local feature greatly improves the parsing performance, especially for inconspicuous objects and background stuff. Furthermore, I will present a scheme of gated sum to selectively aggregate multi-scale features for each spatial position. The gates in this scheme control the information flow of different scale features. Their values are generated from the testing image by the proposed network learnt from the training data so that they are adaptive not only to the training data, but also to the specific testing image. Finally, I will present a scale- and shape-variant semantic mask for each pixel to confine its contextual region. To this end, a novel paired convolution is proposed to infer the semantic correlation of the pair and based on that to generate a shape mask. Using the inferred spatial scope of the contextual region, a shape-variant convolution is controlled by the shape mask that varies with the appearance of input. In this way, the proposed network aggregates the context information of a pixel from its semantic-correlated region instead of a predefined fixed region. This talk is based on 3 papers: H. Ding, X. Jiang, et al, "[Context contrasted feature and gated multi-scale aggregation for scene segmentation](#)," *CVPR'2018 Oral*, H. Ding, X. Jiang, et al, "[Semantic Correlation Promoted Shape-Variant Context for Segmentation](#)," *CVPR'2019 Oral*. and H. Ding, X. Jiang, et al, "[Semantic Segmentation with Context Encoding and Multi-Path Decoding](#)," *IEEE Trans. Image Processing*, 2020.