THE UNIVERSITY OF CHICAGO | **DEPARTMENT OF STATISTICS**

PhD Dissertation Presentation

Angzhi Fan

Department of Statistics
The University of Chicago

"Object Detection and Panoptic Segmentation through Likelihood Optimizations"

Monday, February 26, 2024, at 3:00 PM
Jones 304, 5747 S. Ellis Avenue

Abstract

This thesis focuses on two pivotal subjects within the domain of Computer Vision: object detection and panoptic segmentation. Fueled by deep neural networks, substantial advancements have been witnessed in these fields in recent years. Many efforts in object detection and panoptic segmentation rely on feed-forward approaches, lacking a probabilistic interpretation. In response to this, this thesis puts forth three innovative algorithms: the Detection Selection Algorithm, the Detection Selection Algorithm with Mask, and the Maximizing the Posterior for Panoptic Segmentation Algorithm. The initial algorithm is tailored for object detection, while the latter two are specifically devised for panoptic segmentation. These three algorithms are rooted in three distinct probabilistic frameworks. Notwithstanding, they still depend on feed-forward models like Faster R-CNN and Mask R-CNN to generate raw object detections and instance segmentations. Given an image and a hypothesis regarding object configuration and latent codes, the probabilistic frameworks define their respective likelihoods. The primary objective of these algorithms is to identify a configuration hypothesis that maximizes these likelihoods. They employ greedy search procedures to mitigate computational complexity. The computation of likelihoods necessitates auxiliary tools, including Deep Generative Models that capture the distribution of object appearances. To conduct inference on the distribution of latent codes, Single Reconstruction Algorithms are designed. Additionally, Whole Reconstruction Algorithms are introduced to amalgamate the probability model of individual objects into a comprehensive probability model for the entire image. They necessitate occlusion relationship reasoning methods to identify the visible components of objects. Experimental results demonstrate that our algorithms yield improvements in tasks such as object counting and enhancement of Panoptic Quality scores. This thesis aims to showcase the potency of probabilistic modeling in the world of contemporary Computer Vision.