PhD Dissertation Proposal Presentation

Soumyabrata Kundu  
Department of Statistics  
The University of Chicago

“Steerable Architectures for Vision”

July 29, 2024, at 10:00 AM  
Zoom Meeting

Abstract

Recent advances in computer vision have been significantly driven by the availability of extensive labeled datasets and the development of deep neural networks capable of processing vast amounts of data. While these advancements have solved numerous practical problems, the reliance on large, labeled datasets remain a critical limitation. To address this, there is a growing interest in enhancing the statistical efficiency of machine learning methods through the learning of equivariant representations. In vision tasks, particularly, equivariance to rigid body transformations—embodied by the action of the special Euclidean group SE(d)—is highly desirable. Steerable neural networks are specifically designed for this purpose and are gaining popularity as they are more tailored and aligned with the intrinsic structure of most natural images. Standard vision techniques, such as Convolutional Neural Networks (CNNs) and vision transformers, commonly employed for tasks like classification and segmentation, can be adapted to exhibit equivariance. This discussion aims to explore various neural network architectures and demonstrate how they can be systematically modified to develop steerable versions of themselves, thus improving their performance in vision tasks by aligning them with the geometric properties of image data.