samsi
NSF·Duke·NCSU·UNC

Deep Learning Program
Triangle Machine Learning Day
September 20, 2019

SPEAKER TITLES/ABSTRACT

Sayan Mukherjee
Duke University

"Machine Learning for 3D Imaging"

It has been a longstanding challenge in geometric morphometrics and medical imaging to infer the physical locations (or regions) of 3D shapes that are most associated with a given response variable (e.g. class labels) without needing common predefined landmarks across the shapes, computing correspondence maps between the shapes, or requiring the shapes to be diffeomorphic to each other. In this talk, we introduce SINATRA: the first statistical pipeline for sub-image analysis which identifies physical shape features that explain most of the variation between two classes without the aforementioned requirements. We also illustrate how the problem of 3D sub-image analysis can be mapped onto the well-studied problem of variable selection in nonlinear regression models. Here, the key insight is that tools from integral geometry and differential topology, specifically the Euler characteristic, can be used to transform a 3D mesh representation of an image or shape into a collection of vectors with minimal loss of geometric information.

Crucially, this transform is invertible. The two central statistical, computational, and mathematical innovations of our method are: (1) how to perform robust variable selection in the transformed space of vectors, and (2) how to pullback the most informative features in the transformed space to physical locations or regions on the original shapes. We highlight the utility, power, and properties of our method through detailed simulation studies, which themselves are a novel contribution to 3D image analysis. Finally, we apply SINATRA to a dataset of mandibular molars from four different genera of primates and demonstrate the ability to identify unique morphological properties that summarize phylogeny.