Session H Abstracts

Evaluating Open-set Recognition Approaches for Re-identification
Avirath Sundaresan
Mentors: Serge Belongie, Nico Lang, and Pietro Perona

Re-identification, or re-ID, is a computer vision task that involves recognizing individuals across images. Re-ID techniques can be used for a variety of scenarios, from matching human fingerprints and faces for use in biometric recognition systems to identifying individual animals from natural body markings for biodiversity monitoring. In real-world re-ID scenarios, the set of possible identities is an open-set, where query images encountered during test time may not be represented in the database of known identities. Recent studies in visual recognition have shown that the ability for a classifier to recognize that a query image belongs to an unseen class is correlated with the performance of the classifier on known classes. In this research, we adapt these "familiarity-based" open-set recognition approaches for human and animal re-ID tasks, evaluating different feature extraction methods in their ability to generate meaningful ID representations that can be used for both closed-set re-identification and open-set ID recognition. Our work shows promising results for familiarity-based open-set ID recognition for both human and animal ID datasets.

Classifiers for Measurement Error Correction in Causal Inference
Sayuj Choudhari
Mentors: Zach Wood-Doughty and Chris Umans

Current methods for measurement error correction for classification of unobserved confounding variables in causal inference involve a manual calculation of error rates for different examples of data for a training set that includes the true unobserved confounding. This method becomes vulnerable to miscalculated error rates for cases with increased dimensionality of the observed data that influences the classification of unobserved confounding. A larger variation in examples of data over a set training size would result in less data per example and risk of miscalculation. We propose an alternative method of training a classifier to predict the error rate of a given example of data from which the predicted error rate on each example can be used to correct the distribution of the unobserved confounding variable. This classification method should perform better in higher dimensional cases where the model can learn the impact of each dimensional variable on the error rate and better learn the true error rate of examples with little training data. Our results show that in higher dimensional cases of classification on simple synthetic text generation, that the classification model performs significantly better than the original method.

Segmentation of Volumetric Cell Images With Segment Anything Model
Angel Rodrigo Avelar Menendez
Mentors: Yisong Yue, Uriah Israel, and Markus Marks

Qualitative measurements of volumetric biomedical images achieve a better understanding of biomedical systems such as diseases. Segmentation is a method to take these qualitative measurements. Existing models that perform segmentation are not generalizable, as they need to be trained for specific datasets and tasks, and are mostly designed for 2D images. This project aims to develop a novel 3D segmentation model by building upon Meta's Segment Anything Model (SAM). SAM was trained on millions of images to perform 2D segmentation. Through transfer learning, we extract relevant information on 2D slices of cell images and aggregate their encodings to perform 3D segmentation. We also compare how existing models perform on segmentation of various 3D cell images.

Conformal Prediction for Protein Design
Chu Xin Cheng
Mentors: Yisong Yue and Raul Astudillo

In the field of protein engineering, the discovery of sequences with high fitness is crucial. These sequences exhibit increased potency, affinity, or specificity for a particular antigen, offering significant scientific value. As such, understanding both the fitness value and the associated model uncertainty is essential for purposes such as bias correction or risk mitigation. This project focuses on the utilization of Graph Neural Networks (GNNs) for predicting protein fitness values, while also integrating Conformal Prediction (CP) methods to provide robust confidence sets for these predictions. We propose an approach to apply conformal prediction to the learned feature embeddings of protein sequences, with the aim of creating more precise confidence sets. In future work, we plan to incorporate conformal prediction into multi-fidelity optimization and experiment design problems. Our ultimate goal is to design a comprehensive, end-to-end differentiable pipeline for the efficient and reliable optimization of protein fitness.

Non-Invasive Neural Decoding
Ahamed Raffey Iqbal
Mentors: Yisong Yue and Sabera Talukder
Millions of people suffer from debilitating neuromuscular diseases and injuries that severely limit motor function and communication, such as ALS, Parkinson's, and Dystonia. Brain-machine interfaces (BMIs) work by extracting and subsequently decoding signals from the brain and muscles to enable people with such diseases to regain functionality and independence. Unfortunately, these systems typically require electrodes implanted via highly invasive neurosurgeries, which greatly reduces their accessibility.

Non-invasive electroencephalogram (EEG) recordings collected from a person’s scalp result in noisier signals, but enable wider adoption. Therefore we explored the feasibility of a high-accuracy, generalizable, non-invasive neural signal decoding pipeline to facilitate widespread BMI adoption. To this extent, we developed a multimodal data collection and processing pipeline and are developing and evaluating deep-learning neural decoders. Our deep learning models utilize an autoencoder backbone to generate low dimensional latent spaces and contrastive loss for performant latent space separation. Together these techniques extract useful neural-task-dependent information from the noisier EEG data. We aim to extend our data-collection platform and models to non-invasively control a robotic arm via shared autonomy.

**Kanerva Networks as Models of Syntactic Parameters**
Thanth thanh Nguyen
*Mentor: Matilde Marcolli*

In the field of linguistics, the “Principles and Parameters” model of syntax suggests that natural languages’ grammars are guided in terms of vectors of binary variables, known as “syntactic parameters”. These parameters are thought to be responsible for global language properties such as word order, negation, etc. Some criticisms of this model include its inability to produce a complete set of such parameters across languages, the lack of clear relations between parameters, and whether or not a clear set of independent variables exists among the known parameters.

Using a mathematical model developed by Pentti Kanerva in 1988, which is a network designed to represent human memory, this project aims to test the rigor of the “Principles and Parameters” model. The Kanerva network, also known as sparse distributed memory (SDM), is responsible for modeling how the human brain stores syntactic parameters and deriving dependence relationships between them. This project builds off of prior work done under Professor Marcoli’s direction, which investigated the reoccurrence of certain syntactic parameters in different languages and the level of “recoverability” of certain parameters as a function of others using associative memory.

**Diffusion Models for Brain-Machine Interfaces**
Christina Liu
*Mentors: Pietro Perona and Markus Marks*

Latent models are frequently used to model neuronal activity from neural spike trains. This form of modelling allows for better estimation of the underlying neuronal manifold, which in turn can be the input source of downstream tasks such as the decoding of motor activity based on recorded neural activity. In this study we will explore how diffusion models can be used to estimate the latent neural state space from noisy samples of neural activity spikes. We build upon the existing diffusion model to represent a time-dependent score-based generative model that contains an additional epsilon variable that marks the minimum diffusion time for perturbing data in the forward stochastic differential equation (SDE) process. From this model design, we evaluate the modified diffusion model's ability to denoise and learn the distribution of noisy data without knowing the ground truth data.

**Contrastive Language-Image Pre-training Meets Masked Video Distillation: A New Paradigm for Zero-shot Video Representation Learning**
Andrew Zabelo
*Mentors: Pietro Perona, Markus Marks, and Neehar Kondapaneni*

Knowledge Distillation and Representation Learning via Masked Autoencoders are powerful tools for Computer Vision, enabling state of the art performance on video tasks through pre-training frameworks such as Masked Video Distillation (MVD). Likewise, Contrastive Language-Image Pre-training (CLIP) sets the standard for zero-shot performance on image classification tasks by contrastively learning a joint embedding space via image and text encoders. However, the synergy between these two paradigms has not been fully captured, as previous efforts to finetune CLIP features have struggled to preserve the joint space and current extensions of CLIP to video introduce inductive biases that make it difficult for the model to generalize to new data, hurting performance. We propose to distill features from a frozen CLIP teacher model to a student encoder using the MVD architecture. We show that the student learns to extract rich spatio-temporal features while preserving the zero-shot capabilities provided by the CLIP joint embedding space. [Results, Applications, and Ablation Study TBD]. We release our code and pre-trained model weights at [github link TBD].
Unsupervised Re-Identification of Primates Through Self-Supervised Learning
Kevin Cai
Mentors: Pietro Perona and Markus Marks

Primate re-identification is a critical task in primate behavior research, which involves tracking and identifying individual primates from video footage. Primatologists, anthropologists, zoologists, ecologists, and wildlife biologists gather terabytes of video footage per camera each field season; thus, the current gold standard, manual annotation of this footage, is not scalable. Machine learning techniques have been applied to automate primate re-identification. However, the existing state-of-the-art methods are supervised, which require manually labeled training data, making them impractical in real-world scenarios where the data gathered is too large to actively label. Moreover, domain experts often do not know and therefore cannot label all the individuals. We propose an unsupervised primate re-identification framework that automatically identifies individual primates from video footage without requiring full training on a manually labeled dataset. Our framework will consist of a self-supervised student-teacher network that will train an identification model through which we will embed the video frames and cluster the individuals, so they can be tracked across multiple views. To evaluate the proposed framework, we will conduct experiments on various primate datasets to compare our model to existing benchmarks, and we will test our framework on a curated dataset with different environmental and lighting conditions to show that our framework can generalize across different environmental conditions, field sites, seasons, time of day, and primate groups.