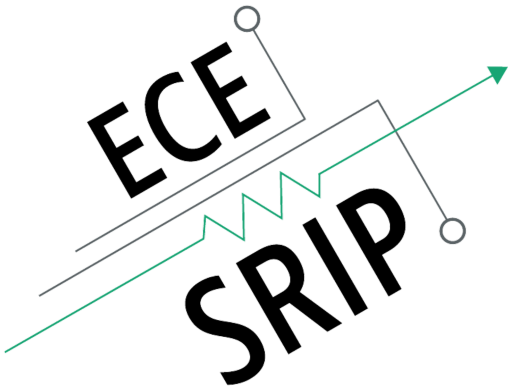


FACULTY MENTOR

Nuno Vasconcelos

PROJECT TITLE

Unsupervised Realistic Visual Question Answering (RVQA)

PROJECT DESCRIPTION

Visual question answering (VQA) is the problem of training a machine learning system to answer questions that refer to a picture, such as "what color is the person's jacket?" In realistic VQA (RVQA), the model has to both reject unanswerable questions (UQs) and answer answerable ones (AQs). RVQA is particularly important for applications such as indoor robots or assistants for the visually impaired. In this project, we aim to propose a more robust VQA model, less vulnerable to UQs, without sacrificing answer accuracy for AQs. More importantly, the robust VQA model should be trained without UQs (i.e., in an unsupervised or self-supervised manner). The model will then be evaluated on a dataset collected by SVCL, which contains 30K human-annotated UQs. Students will use language libraries like Hugging Face and work on Transformer architectures. The project aims for a top-tier conference publication.

Mentor: John Ho [Chih-Hui Ho <chh279@eng.ucsd.edu>]

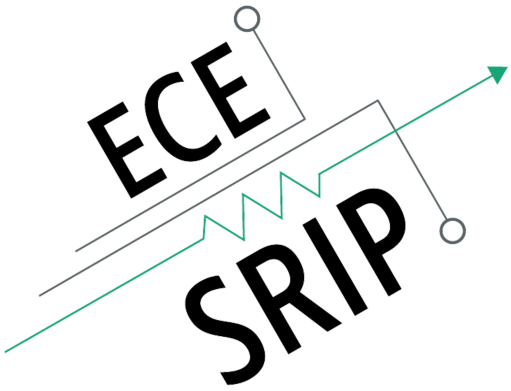
This project can accommodate both remote and in-person students.

INTERNS NEEDED

2+ Students

PREREQUISITES

- MS students. Experience with Python, Linux, and at least one popular deep learning framework such as PyTorch is an advantage. Stronger candidates will also have some knowledge of computer vision or natural language processing.



FACULTY MENTOR

Nuno Vasconcelos

PROJECT TITLE

Iterative Dataset Collection

PROJECT DESCRIPTION

While a wide spectrum of automated tools has been created for building deep learning models in the past decade, and dataset collection has remained a largely manual process with little systematic effort to account for bias in raw data or human annotations. The goal of this project is to build an iterative framework for dataset collection, annotator teaching, and model training. Under this unified framework, new examples are automatically selected for human annotation, cleaned for label bias, and added to the dataset progressively. Neural network models are trained on each iteration of data, and model explanation techniques are used to create teaching examples that reduce the bias of crowd-sourced annotators. The whole framework aims to produce datasets that are optimal for machine learning under multiple objectives, including classification accuracy and fairness. The research is connected to topics like active learning and human-in-the-loop AI systems. The project aims for a top-tier conference publication.

Mentor: Yi Li <yil898@eng.ucsd.edu>

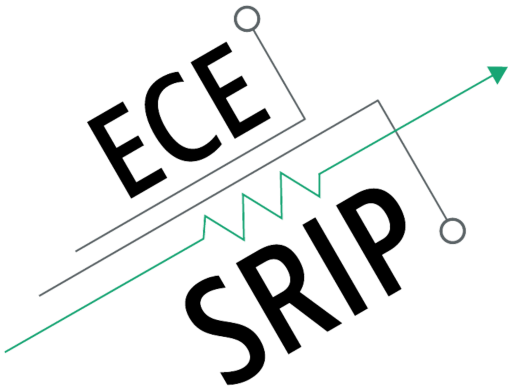
This project can accommodate both remote and in-person students.

INTERNS NEEDED

2+ Students

PREREQUISITES

- MS and undergraduate students. Experience with Python, Linux, and at least one popular deep learning framework such as PyTorch is an advantage. Stronger candidates will also have some knowledge of computer vision or natural language processing.



FACULTY MENTOR

Nuno Vasconcelos

PROJECT TITLE

Generalized Bird's-Eye View (BEV) Perception with Monocular Sequences

PROJECT DESCRIPTION

Many problems in computer vision, such as the navigation of autonomous agents (e.g., smart home robots, cars), or the detection of off-sides on soccer matches, require reasoning about the 3D world. However, 3D sensors like LiDAR are not always available due to high cost, lack of resolution, robustness, etc. Hence, there is interest in recovering 3D information from monocular video. An interesting representation is the bird's-eye view (BEV) of the scene. For example, for agents on the ground, such as cars or soccer players, the BEV is sufficient to identify the navigable area and helps with tracking, motion planning, etc. In this project, we are interested in training deep learning systems to estimate the BEV map from monocular sequences. To infer from 2D to 3D, in addition to geometric priors (e.g., door heights, object sizes, soccer pitch markers), we plan to leverage the geometric projection between consecutive frames in a monocular video to relax the need for depth supervision. The project aims for a top-tier conference publication.

Mentor: Gina Wu [Tz-Ying Wu <tzw001@eng.ucsd.edu>]

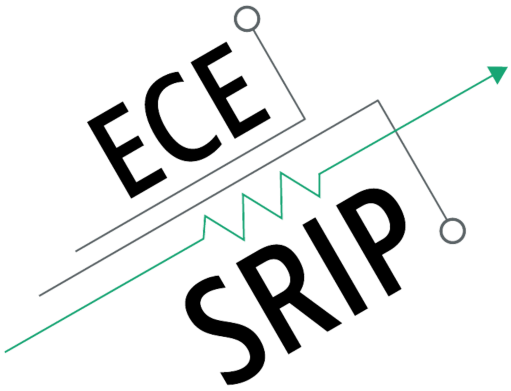
This project can accommodate both remote and in-person students.

INTERNS NEEDED

2+ Students

PREREQUISITES

- MS candidates are expected to be adept with Python and Linux. Stronger candidates will also have some knowledge of computer vision and experience with at least one popular deep-learning framework, such as PyTorch.



FACULTY MENTOR

Nuno Vasconcelos

PROJECT TITLE

Continual Learning with Deep Understanding of Objects

PROJECT DESCRIPTION

Humans live in a world that is continually changing. New objects and concepts emerge every day. Humans are able to learn new things during their lifetime without forgetting about old things. However, gradient descent-based Deep Neural Networks (DNNs) suffer from the catastrophic forgetting problem. Once they learn a new task, they tend to forget old ones. Continual Learning (CL) aims to find a solution to this. Existing CL methods mainly focus on consolidating the existing knowledge to prevent forgetting, but on the other hand, humans do not intentionally recall old objects and re-learn them in the memory while learning. By comparing new objects with old objects, more latent, deep connections are built to both consolidate the knowledge of old objects and enhance the learning of new objects. For example, when observing a new type of animal, humans naturally partition it into multiple parts, including the body, arms, legs, tail, etc., and compare those parts with existing animals to find their kinships. In this work, we will try to solve the CL problem by learning a latent space that represents our deep understanding of objects and as the bridge between old and new objects.

Mentor: Zhiyuan Hu <z8hu@ucsd.edu>

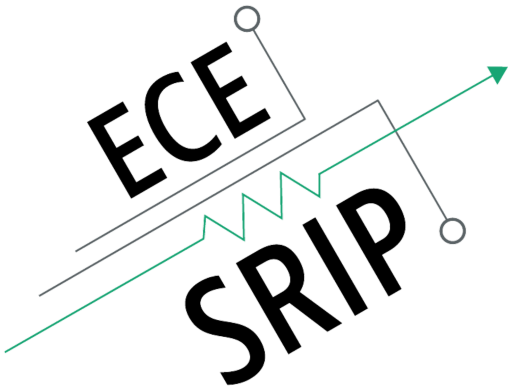
This project can accommodate both remote and in-person students.

INTERNS NEEDED

2+ Students

PREREQUISITES

- MS students. Experience with Python, Linux, and at least one popular deep learning framework such as PyTorch is an advantage. Stronger candidates will also have some knowledge of computer vision and continual learning.



FACULTY MENTOR

Nuno Vasconcelos

PROJECT TITLE

Customizing Radiation Cancer Treatment with Deep Learning

PROJECT DESCRIPTION

Brachytherapy is a treatment in which a radioactive source is used to deliver radiation internally to treat cancers such as cervical cancer. Currently, clinicians manually tune treatment parameters to customize the radiation to individual patients' anatomy. This process can take over an hour, which is problematic because patients are waiting in discomfort and often under sedation for this to occur. Deep learning can identify anatomical features that relate to ideal, customized radiation treatments by learning from past patient imaging and treatment data. In this project, we will generate new networks and inputs and/or modify existing networks to accurately predict radiation treatment parameters. The end goal is to automate the treatment customization process to ensure high-quality radiation treatments can be produced in a matter of minutes with a single button click. This project will involve working with a team of medical physicists (including Dr. Sandra Meyers), radiation oncologists, and electrical engineers and is a collaboration between the Vasconcelos and Meyers labs. The project aims for a top-tier conference or journal publication.

Mentors: Lance Moore <lcmoore@health.ucsd.edu>; Sandra Meyers <smmeyers@health.ucsd.edu>

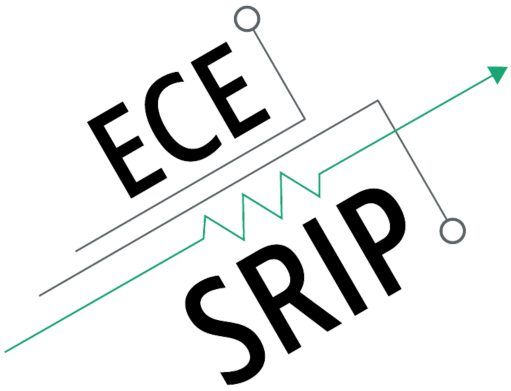
This project can accommodate both remote and in-person students.

INTERNS NEEDED

2+ Students

PREREQUISITES

- MS students. Experience with Python, Linux, and at least one popular deep learning framework such as PyTorch is an advantage.



FACULTY MENTOR

Nuno Vasconcelos

PROJECT TITLE

Accurate 3D-Hand Pose Estimation via Multi-Modal Fusion

PROJECT DESCRIPTION

3D hand pose estimation is a fundamental and challenging problem in vision. The goal is to recover the joint angles of the different finger and hand sections. This can be used, for example, in applications such as assistive remote surgery, training robots to grasp objects, or AR/VR problems that warrant the need for accurate 3D hand pose. However, accurate pose estimation using only RGB or depth cameras is impossible in occluded scenarios without the help of additional information. Such situations arise commonly since hands, fingers, and objects tend to occlude each other. To circumvent this, we propose to leverage the use of sensory information from a sensor attached to the human in conjunction with camera frames. In particular, we aim to explore different image/video and sensor fusion methods, including but not limited to early, mid and late fusion methods. Further, the sensor information should be complementary to visual input that can provide predictions independent of the sensor inputs. This project involves developing an end-to-end pipeline from data collection, pre-processing, and designing the algorithm with a focus on transformer attention architectures. The project aims for a top-tier conference publication.

Mentor: Deepak Sridhar <desridha@.ucsd.edu>

This project can accommodate both remote and in-person students.

INTERNS NEEDED

2+ Students

PREREQUISITES

- MS students. Experience with Python, Linux, and at least one popular deep learning framework such as PyTorch is an advantage. Experience with 3D vision is an asset.

FACULTY MENTOR
Nuno Vasconcelos

PROJECT TITLE
Generalizable Neural Radiance Fields (NeRF) with Few Images

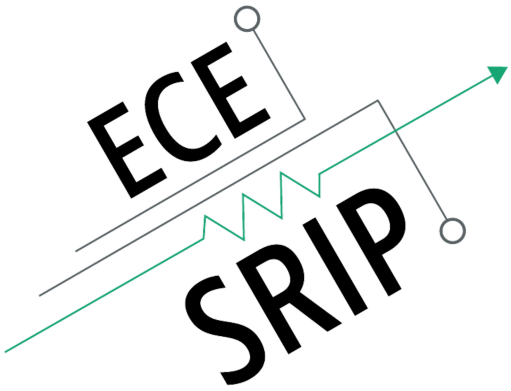
PROJECT DESCRIPTION
Neural Radiance Fields lean a continuous representation through multiview consistency, which can then be rendered from any viewpoint. Recently, NeRF-based representations have made significant progress in novel view synthesis and producing photo-realistic rendering results. However, NeRF optimization usually requires a large number of images to model accurate geometry and texture. It is observed that the rendering results decay fast as the number of image inputs decreases. In this project, we will investigate how to learn NeRF with fewer images. Training NeRF with fewer images would open the door for various applications in the real world. For example, one can possibly turn several images taken by your phone into an interactive 3D scene. The project aims for a top-tier conference publication.

Mentor: Jiteng Mu <jmu@ucsd.edu>

This project can accommodate both remote and in-person students.

INTERNS NEEDED
2+ Students

PREREQUISITES
- MS students. Candidates are expected to be adept with deep learning (DL) and Python/PyTorch. Experience with 3D vision is an asset.



FACULTY MENTOR

Nuno Vasconcelos

PROJECT TITLE

Crowd Localization with Diffusion Models

PROJECT DESCRIPTION

Crowd counting and localization aim to count the number of people and localize individuals in images. This can be applied to many applications such as public transportation, public safety, and business management. Although the counting performance has improved dramatically in recent years, the localization of individuals in dense crowds is still challenging. The localization is usually done by finding the local maximum and thresholding, which is sensitive to the threshold. In this project, we will explore diffusion models to localize individuals by gradually focusing on their locations from a noisy prediction. In particular, a smooth heat map will be predicted from an image. The smooth map is easy to use for counting but hard for localization. We will gradually predict a sharper map from a smooth one and generate the locations in the end. The project aims for a top-tier conference publication.

Mentor: Jia Wan <jiw162@ucsd.edu>

This project can accommodate both remote and in-person students.

INTERNS NEEDED

2+ Students

PREREQUISITES

- MS students. Candidates are expected to be adept with deep learning (DL) and Python/PyTorch.
- Experience with 3D vision is an asset.