Neehar Kondapaneni

Email: nkondapa@caltech.edu **Website:** https://nkondapa.github.io/

My research interests are in *machine teaching for humans* and *eXplainable Al*. I am developing algorithms to (1) identify the features used by models to achieve superhuman performance on visual tasks and (2) teach those features to human learners.

Education

California Institute of Technology

Computation & Neural Systems, PhD Candidate

Sept 2018 - Present

University of California - San Diego

B.S. Neuroscience and Physiology, Minor in Computer Science Magna Cum Laude

2014 - 2017 GPA: 3.91/4.00

Publications

Text-image Alignment for Diffusion-based Perception - CVPR 2024

Neehar Kondapaneni*, Markus Marks*, Manuel Knott*, Rogério Guimarães, Pietro Perona

Less is More: Discovering Concise Network Explanations - ICLR Re-Align Workshop 2024

Neehar Kondapaneni*, Markus Marks, Oisin Mac Aodha, Pietro Perona

Visual Knowledge Tracing - ECCV 2022

Neehar Kondapaneni, Oisin Mac Aodha, Pietro Perona

A Number Sense as an Emergent Property of the Manipulating Brain - Scientific Reports 2024 Neehar Kondapaneni, Pietro Perona

Transformation of Cortex-wide Emergent Properties during Motor Learning - Neuron 2017 Hiroshi Makino, Chi Ren, Haixin Liu, An Na Kim, Neehar Kondapaneni, Xin Liu, Duygu Kuzum, Takaki Komiyama

Preprints

A Closer Look at Benchmarking Self-Supervised Pre-training with Image Classification

Markus Marks, Manuel Knott, Neehar Kondapaneni, Elijah Cole, Thijs Defraeye, Fernando Perez-Cruz, and Pietro Perona

Research Projects

Caltech Vision Lab

Advisor: Dr. Pietro Perona

Text-image Alignment for Diffusion-based Perception (2023)

- Proposed a novel technique for prompting the diffusion backbone. Applied a captioning model on the dataset to generate prompts for the diffusion backbone. Found that this enforced text and image alignment and improved the intermediate feature maps.
- Extended this idea to the cross-domain setting and found that aligning the prompt to target domains improved performance on several datasets.
- Achieved SOTA results in NYUv2 Depth, Pascal to Watercolor2K, and Cityscapes to Nighttime Driving.

Less is More: Discovering Concise Network Explanations (2023)

- Developed a dataset to measure the conceptual alignment between model concepts and human expertdefined concepts.
- Developed a method to compare how different XAI methods aligned to human expert-defined concepts.
- Proposed an XAI method to reduce explanation complexity while preserving explanation alignment to human expert concepts.

Visual Knowledge Tracing (2022)

- Introduced visual knowledge tracing as a novel problem for the computer vision community, where the goal is to predict a student's response to a visual stimulus given their history of responses.
- · Collected three new datasets for evaluating models for visual knowledge tracing.
- Proposed and evaluated prototype and exemplar cognitive models, and LSTM and transformer neural network model. All models predicted student responses significantly above chance, with the LSTM as the top performer.

A Number Sense as an Emergent Property of the Manipulating Brain (2020)

- Developed a *cross-modal unsupervised* learning framework where self-directed actions act as a supervisory signal for perception.
- Our framework organizes numerosity sensitive neurons in an abstract, generalizable representation that reproduces classic results in psycophysics.

Automated mouse behavior environment for Two-Alternative Forced Choice (2-AFC) Tasks (2019) Advisor: Dr. Pietro Perona & Dr. Markus Meister (2018 - 2019)

- Developed a framework to determine whether an automated behavior environment with curriculum learning could improve the speed of mice learning two-alternative forced choice tasks.
- Tested a curriculum learning algorithm that scaled task difficulty with prior performance.

Komiyama Lab - Transformation of Cortex-wide Emergent Properties during Motor Learning Advisor: Dr. Takaki Komiyama (2014 - 2016)

- Designed a motor learning task for mice, built hardware and software to collect data, trained mice, and analyze the neural recordings.
- Developed and implemented algorithms to (1) match image annotations across image channels of immunohistochemistry and (2) partial automation of parsing and annotating videos of mouse behavior.
- Found motor learning modulates whole brain activity by creating a sequential activity pattern, driven by the pre-motor cortex, that is temporally compressed over the course of learning.

Teaching

Selected Topics in Computational Vision (EE/CNS/CS 148)

- (2023, Head TA) Oversaw the development of new course material to focus on recent advances in AI, including transformers, diffusion models, and large language models.
- (2022) Developed and delivered a lecture introducing the mathematical details and intuition behind transformers and vision transformers' attention, including high-level differences between CNNs and ViTs and differences in their latent representations. (Slides)
- (2021) Provided tutoring and guidance to students on class material. Worked closely with Amazon SageMaker EDU team to provide a managed environment for students to train their models. Developed tutorials to teach students how to use PyTorch, Docker, and AWS.

Activities

Reviewer - ICCV 2023, Neurips 2023 (Workshop) **SURF Mentor**

<u>Andrew Zabelo (2023 - current)</u> is an undergraduate at Caltech. We are studying how to use distillation from vision-language models to improve zero-shot performance for action recognition in video. <u>Daniel Israel (2021 - 2022)</u> is a current graduate student UCLA. Daniel assisted in creating ablation experiments to study the limits of our number sense model.

STEM Stall - brought STEM to a rural community through demos and question and answer sessions. **Freshman Summer Research Institute** - Co-mentored an incoming freshman from a minority background. Taught them the basics of computer vision research through a project analyzing images collected from camera traps in the wild.

Rise Tutoring (2018-2020) - tutored underprivileged Pasadena high-school students in mathematics.