

Ce Zhang

✉ cezhang@cs.unc.edu | 🏠 [cezh.github.io](https://github.com/cezh) | 🐦 twitter.com/cezhhh | 🎓 [Google Scholar](#)

EDUCATION

UNC-Chapel Hill

Ph.D. student in Computer Science, GPA: 4.0 / 4.0

- Advisor: [Prof. Gedas Bertasius](#)

Chapel Hill, NC, US

Aug. 2023 – Present

Brown University

M.S. in Computer Science, GPA: 3.75 / 4.0

- Advisor: [Prof. Chen Sun](#)

Providence, RI, US

Aug. 2021 – May 2023

Southeast University

B.S. in Computer Science, GPA: 3.75 / 4.0, Rank: Top 10%

Nanjing, Jiangsu, China

Aug. 2016 – Jun. 2020

RESEARCH INTEREST

I'm broadly interested in Computer Vision, Multimodal learning and Robotics. Currently, I'm mainly working on **video understanding**, with a focus on leveraging **foundation models** (LLMs, VLMs, etc.) to solve multiple video understanding tasks. I'm also interested in **offline decision making**, especially learning from videos. I believe that videos can provide rich sources of demonstrations for robot learning, and that the commonsense knowledge encoded in foundation models can help solve robotic tasks more efficiently and robustly.

EXPERIENCE

Research Assistant

UNC-Chapel Hill, Advisor: [Prof. Gedas Bertasius](#)

- Proposed a framework for long-range videoQA by decomposing videoQA into short-term captioning and long-range language modeling. Achieved SOTA on EgoSchema, NeXT-QA, IntentQA and NeXT-GQA. (**Preprint**) [link](#)

Aug. 2023 – Present

Chapel Hill, NC

Research Assistant

Brown University, Advisor: [Prof. Chen Sun](#)

- Represent the video with discretized action labels and utilized LLMs for reasoning and long-term action anticipation. Achieved SOTA on Ego4D, EPIC-Kitchens-55 and EGTEA Gaze+. (**ICLR 2024**) [link](#)
- Extracted task-specific object-centric representations from pretrained models (e.g. GLIP). Utilized object-centric representations for long-term action anticipation. Achieved competitive results on Ego4D, 50Salads and EGTEA Gaze+. (**WACV 2024**) [link](#)
- Condensed expert trajectory demonstrations into useful representations for policy learning. Achieved competitive performance on performance on AntMaze, FrankaKitchen and Locomotion. (**NeurIPS 2023**) [link](#)

Apr. 2022 – May 2023

Providence, RI

Machine Learning Engineer Intern

QCraft, Inc.

- Worked on 3D Multi-Object Tracking for vehicles and pedestrians by fusing 2D and 3D appearance feature (from ResNet and MLP) and 2D motion feature (from Kalman Filter).

Mar. 2021 – July 2021

Beijing, China

Machine Learning Engineer Intern

Momenta, Inc.

- Generated pseudo labels for 2D facial key points by fusing 2D detection from multi-view cameras. Bootstrapped the model by training on pseudo labels.

Feb. 2020 – July 2020

Suzhou, China

PUBLICATION

- *Ce Zhang**, *Taixi Lu**, *Md Mohaiminul Islam*, *Ziyang Wang*, *Shoubin Yu*, *Mohit Bansal*, *Gedas Bertasius*. A Simple LLM Framework for Long-Range Video Question-Answering (**Preprint**) [link](#)
- *Qi Zhao**, *Shijie Wang**, *Ce Zhang*, *Changcheng Fu*, *Minh Quan Do*, *Nakul Agarwal*, *Kwonjoon Lee*, *Chen Sun*. AntGPT: Can Large Language Models Help Long-term Action Anticipation from Videos? (**ICLR, 2024**) [link](#)
- *Ce Zhang**, *Changcheng Fu**, *Shijie Wang*, *Nakul Agarwal*, *Kwonjoon Lee*, *Chiho Choi*, *Chen Sun*. Object-centric Video Representation for Long-term Action Anticipation (**WACV, 2024**) [link](#)
- *Zilai Zeng*, *Ce Zhang*, *Shijie Wang*, *Chen Sun*. Goal-Conditioned Predictive Coding as an Implicit Planner for Offline Reinforcement Learning (**NeurIPS, 2023**) [link](#)

SERVICES

Reviewer: ECCV 2024, T4V @ CVPR 2024

Organizer: T4V @ CVPR 2024

TECHNICAL SKILLS

Languages: Python, C/C++, Go

Frameworks: PyTorch, TensorFlow

Other: Git, Slurm, Docker