

# Hand and Object Pose Tracking in Multi-Camera System

Team Members: Rebecca Bo, Isaac Chen, May Espinola, Reanna Jin, Nicole Song, Lyndon Yang, Ida Zou

Mentors: Jikai Wang, Qifan Zhang | Professor Yu Xiang

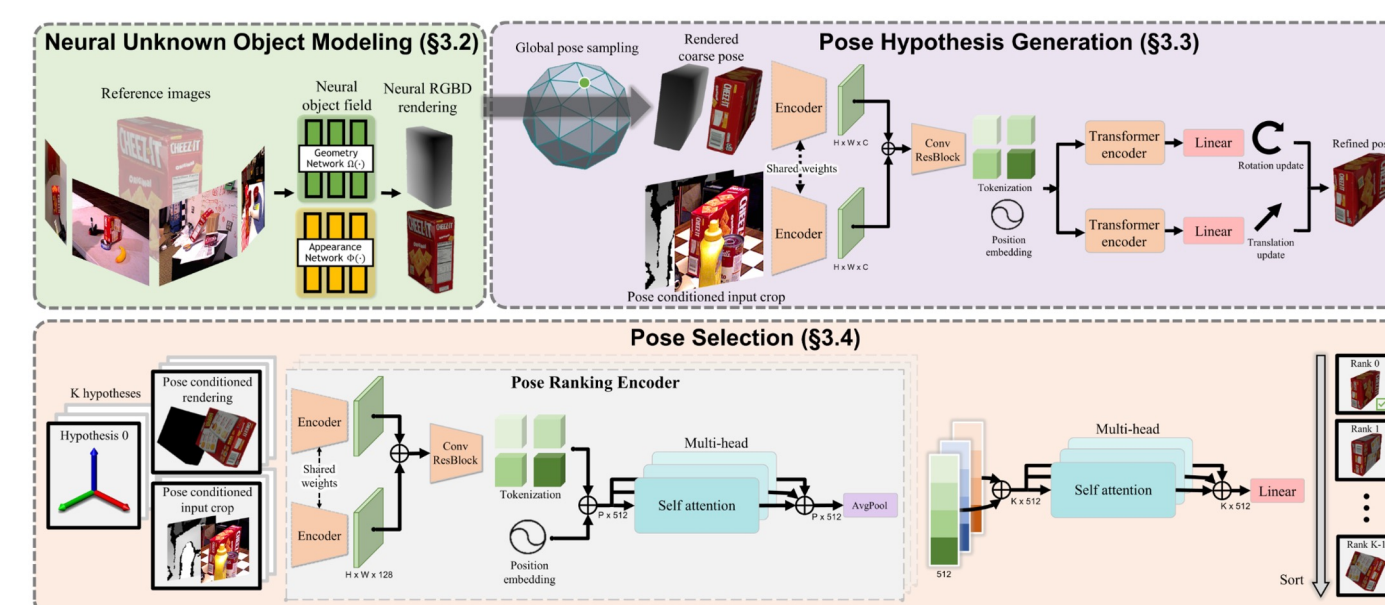
## Motivation

Understanding hand-object interactions from images or videos is crucial in computer vision, the key challenges include pose estimation and shape reconstruction.

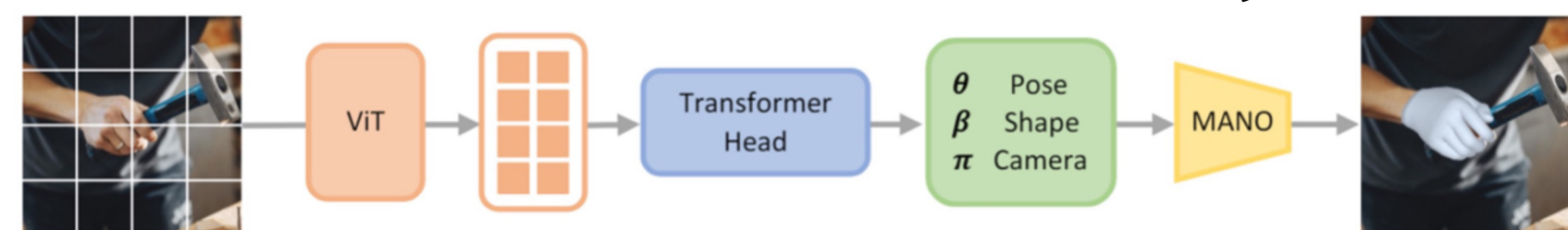
This project aims to develop a pipeline for recognizing hand and object poses from data collected by a multi-camera capture system.

## State-Of-The-Art

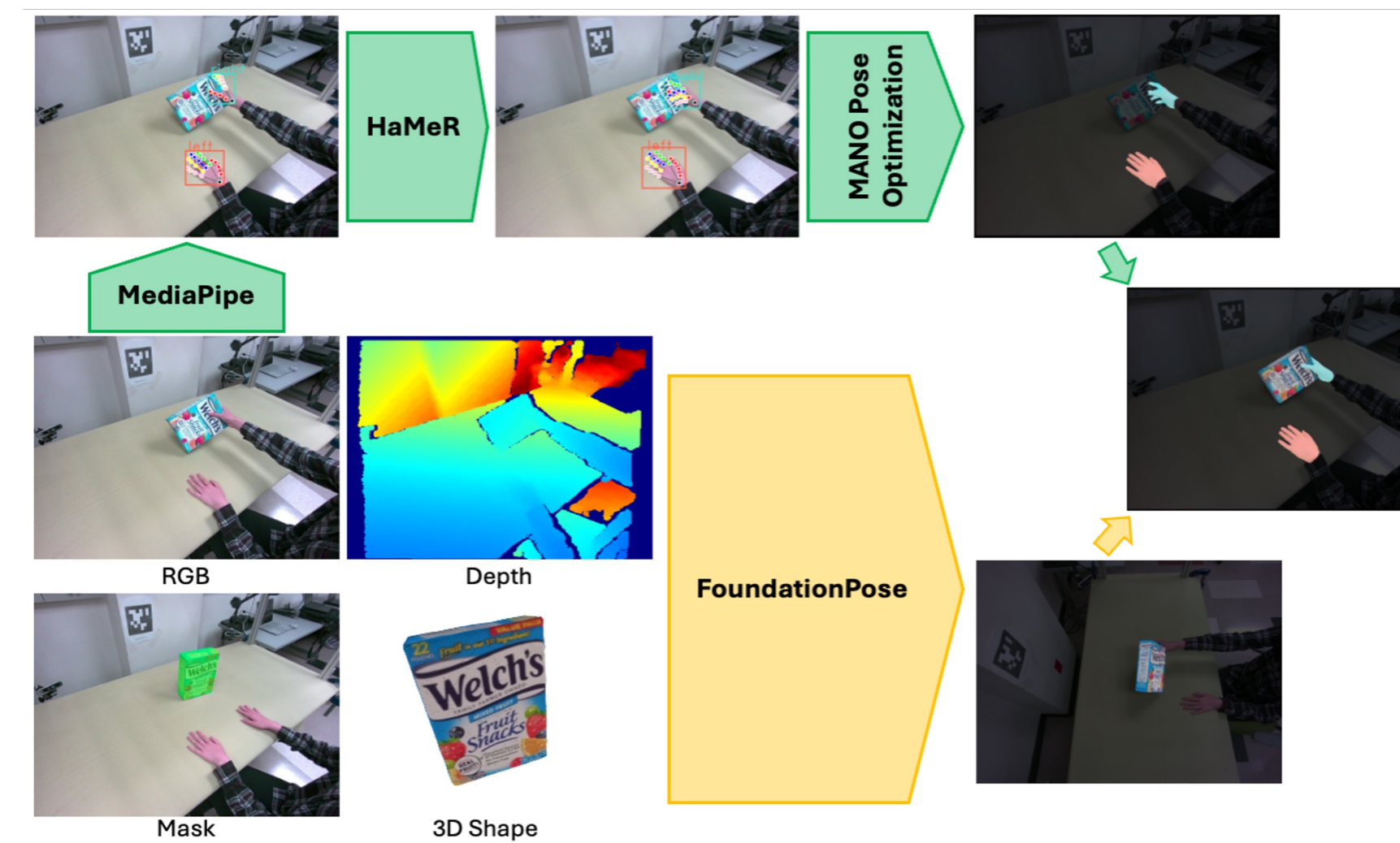
FoundationPose: 6D Pose Estimation and Tracking of Novel Objects



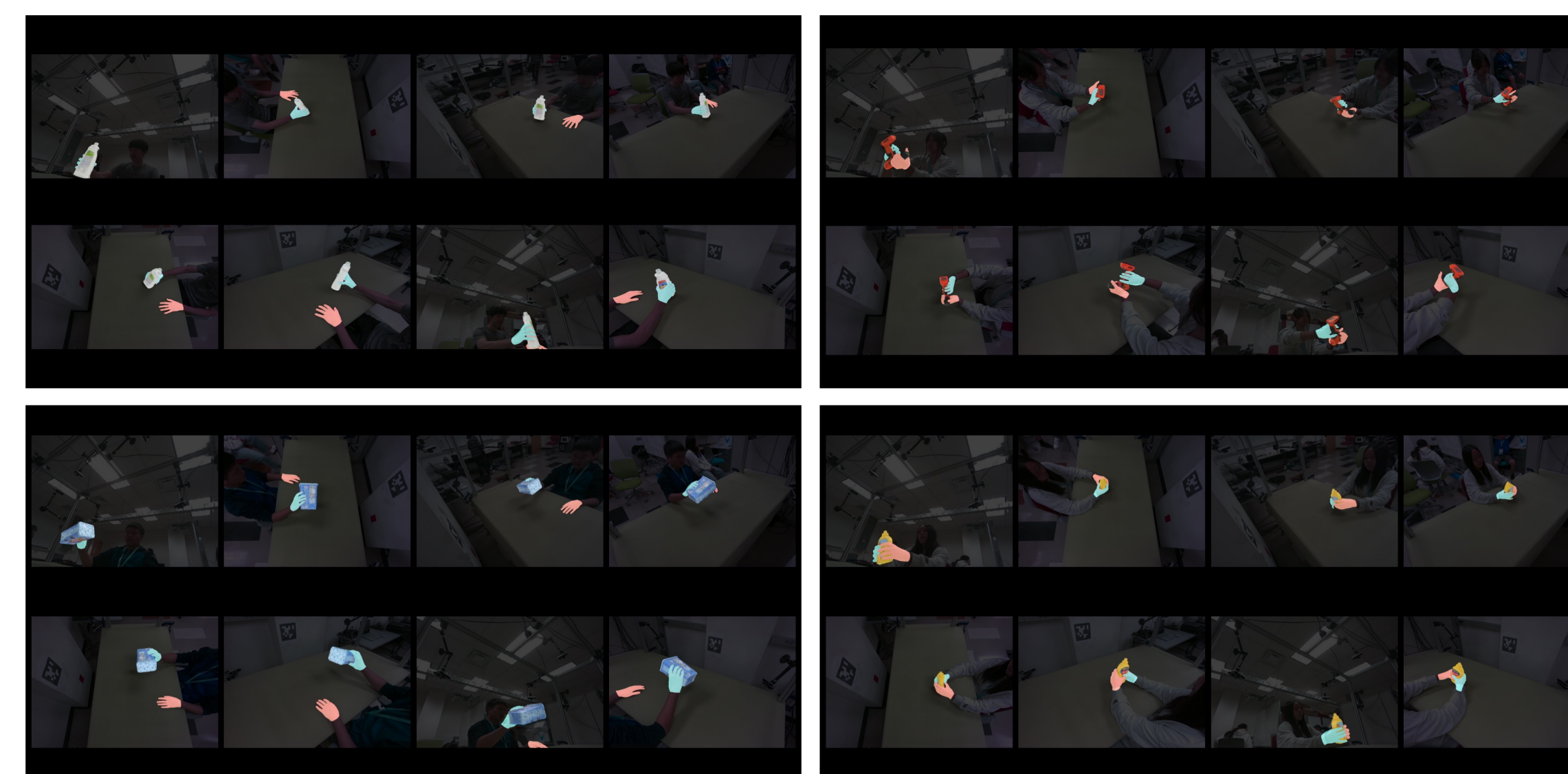
HaMeR: Hand Mesh Recovery



## Pipeline



## Results

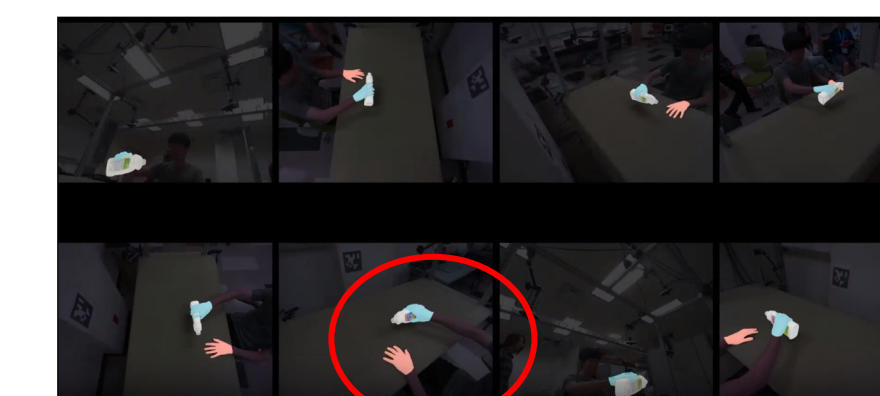


## Conclusions and Outlook

We successfully collected data using the multi-camera capture system. The developed pipeline efficiently and accurately estimates both hand and object poses jointly.



It is important to note that while the results are highly effective, the system may fail to accurately detect poses under challenging conditions, such as a vibrating hand.



Vibrating hand

## Acknowledgments

CAST STEM Bridge 016 would like to thank Yu Xiang, Jikai Wang, and Qifan Zhang for their guidance and support throughout this project. Thank you to UTD for hosting and allowing the usage of their equipment and facilities which were necessary for this project.

## Resources

Bowen Wen, Wei Yang, Jan Kautz, Stan Birchfield, "FoundationPose: Unified 6D Pose Estimation and Tracking of Novel Objects," in CVPR, 2024.

Georgios Pavlakos, Dandan Shan, Ilija Radosavovic, Angjoo Kanazawa, David Fouhey, Jitendra Malik, "Reconstructing Hands in 3D with Transformers," in CVPR, 2024.

## Data Collection

Multi-Camera Capturing System



Objects



3D Shapes and Textures

