

# Yuheng Wang

Seattle, WA | arrokothwhi@gmail.com | 206-390-3402 | github.com/Arrokothwh

## Education

---

**University of Washington**, BS in Mathematics, Computer Science Sept 2023 – May 2027

GPA: 3.9/4.0

**Coursework:** Linear Algebra, Probability, Optimization, Deep Learning, Advanced Machine Learning, NLP

## RESEARCH EXPERIENCE

---

**Student Researcher**, Ai2, Seattle, WA Apr. 2026 – Present

- Work on video-based vision encoder training infrastructure for multimodal models.

**Research Assistant**, Raivn Lab, University of Washington, Seattle, WA Sep. 2025 – Present

- Focus on generative modeling, vision understanding, and multimodal learning for visual AI systems.

## SELECTED RESEARCH PROJECTS

---

**RefDecoder: Conditional Video Decoding** Sep. 2025 – Mar. 2026

Raivn Lab, University of Washington

[refdecoder.github.io/](https://refdecoder.github.io/)

- Developed a conditional VAE decoder for reference-aware image-to-video generation, improving reconstruction and generation quality.
- Designed decoder modules that inject reference-frame information during video VAE decoding and improved reconstruction quality over baseline approaches.
- Trained and evaluated models on large-scale video datasets using reconstruction quality and decoding-efficiency benchmarks across model variants.

**Molmo Encoder** Apr. 2026 – Present

Ai2

- Develop video training pipelines for a Molmo2 vision encoder with video inputs for vision-language models.
- Build the video data pipeline, including clip loading, frame sampling, preprocessing, batching, and distributed training support.
- Improve training reliability by debugging pipeline issues, validating batch behavior, and monitoring large-scale multimodal runs.

**Virtual VLA** Apr. 2026 – Present

Raivn Lab, University of Washington

- Contribute to a vision-language-action project on multimodal sequential learning from visual observations, language, and action trajectories.
- Develop video-based data and training pipelines for embodied learning with aligned visual, language, and action sequences.

## Technologies

---

**Languages:** Python, Java, C, JavaScript, TypeScript, MATLAB

**Technologies:** PyTorch, TensorFlow, Distributed Training, Computer Vision, Generative Models, Transformer Architectures, Data Pipelines