

WENBO (GORDON) HU

Los Angeles, CA whu@cs.ucla.edu [Personal Website](#) [Linkedin](#) [Github](#)

SUMMARY

I'm a machine learning engineer and language generation researcher advised by [Prof. Nanyun Peng](#) at UCLA. My primary research interest lies in enabling large multimodal/language models to have perception and comprehension following human values. I graduated from UCSD and used to work on unsupervised image classification under language supervision, 2D and 3D object detection, and model-free reinforcement learning.

EDUCATION

University of California, Los Angeles *Sept.2023 - Present*
Master of Science in Computer Science
University of California, San Diego *Sept.2019 - March.2023*
Bachelor of Science in Data Science; Minor in Business *Major GPA: 3.93, Total GPA: 3.88*

RESEARCH

Student Research Assistant at PLUS Lab mentored by [Nanyun Peng](#) at UCLA *Sept.2023 - Present*
- Research of a comprehensive benchmark for evaluating hallucinations of objects, attributes, and relations information in large vision language models' answers. Employed GPT3.5 in extracting and formatting the generated outputs.

Research Intern at mlPC mentored by [Zhuowen Tu](#) at UCSD *May.2023 - Present*
- Led research of a simple multimodal LLM for better handling of text-rich visual questions. Under Review AAI 2024.
- Our model, BLIVA, significantly enhances performance in processing **4** text-rich VQA benchmarks (up to **17.76%** in OCR-VQA task) and in undertaking **8** typical VQA benchmarks (up to **7.9%** in Visual Spatial Reasoning task).

Undergraduate Researcher Supervised by [Tsui-Wei Weng](#) at UCSD *Sept.2022 - March 2023*
- Research focuses on improving Vision-Language models such as CLIP (Contrastive image representation learning under natural language supervision) by both visual and text prompt learning.
- Independently researched and designed a new model and outperformed the SOTA models by **6.33%** classification accuracy on average of **10** image recognition datasets for 1 shot training and **4.43%** on the 16 shot training setting.

Computer Vision Research Intern at [Hao Su Lab](#) at UCSD *Feb.2022 - December 2022*
- Conducted follow-up research of [ManiSkill 2022 Challenge](#), which is minimizing simulation to real robot transferability gap with Xarm7 joint robot arm and implemented hand-eye calibration algorithm with Intel Realsense camera.
- Employed Unet and implemented Iterative Closest Point algorithm to solve the 6D Pose estimation problem.
- Experienced in solving Forward and Inverse Kinematics problems, Motion Planning tasks(RRT, PRM), PID control and Reinforcement Learning algorithms such as PPO, SAC using the [SAPIEN environment](#) (CVPR 2020).

WORK EXPERIENCE

Algorithm Engineer Intern at [Synthesis Electronic Technology](#) *June.2021 - Aug.2021*
- Researched lightweight object detection models such as Yolo Series; Compressed and accelerated deep learning models to run on small devices (CPU chip & mobile end).
- Improved standard Yolov5s model with **5% mAP@.5 increase** by utilizing in-depth data augmentation, TTA, and applying Transformer as the backbone after reading 10+ papers in Transformer domain.
- Converted models from different frameworks to NCNN, ONNX, and TensorRT that can run on mobile devices and deployed int 8 quantized yolov5s on intel i5 CPU with speed 180ms in 1920*1080 video.

Software Engineer Intern at [Inspur Groups](#) *July.2020 - Sept.2020*
- Assisted back end development and support for the customer company using MyBatis and SpringMVC framework under Maven and Tomcat.
- Developed a comprehensive data monitoring and visualization system. Achieved through SQL database querying, Java-based development of the controller, DAO, data, and service layers, and effectively addressing front-end requests.

SKILLS

Proficient in Python, Java, R, SQL, Shell, Javascript, and MATLAB; Utilize Pytorch, Numpy, Sklearn, Scipy, Pandas, AWS, Docker, Kubernetes, PySpark, Dask, OpeanAI Gym and PostgreSQL.