



Brandon Yifan Yang

 [brandonyifanyang.com](https://github.com/brandonyifanyang)  yang52@seas.upenn.edu  [branyang02](https://github.com/branyang02)

EDUCATION

-  **University of Pennsylvania** Philadelphia, PA
MSE in Robotics Aug 2025 - May 2027
-  **University of Virginia** Charlottesville, VA
BS in Computer Science with Highest Distinction August 2021 - May 2025
- **Relevant Coursework:** Robotics*, CUDA Programming & Architecture*, Machine Learning, Reinforcement Learning*, Natural Language Processing*, Probabilistic ML*, Optimization*
*Graduate-level courses.

EXPERIENCE

- **Embodied Intelligence Research Intern** Beijing, China
Spirit AI May 2025 - Aug 2025
 - Trained and deployed Vision-Language-Action (VLA) model variants in PyTorch on a dual-arm robot, experimenting with contrastive objectives, observation-noise curricula, reduced guidance inputs, and mixture-of-experts (MoE).
 - Designed and implemented DAGger data-collection pipeline for VLA models, including operator-takeover logic and post-training workflows for scaling corrective demonstrations.
 - Integrated robotics simulation environments into a unified, reproducible setup, enabling efficient evaluation of VLA models; containerized training and evaluation workflows with Docker and Slurm. Open-sourced codebase and datasets on [GitHub](#).

RESEARCH EXPERIENCE

- **General Robotics, Automation, Sensing, and Perception Lab, UPenn** Philadelphia, PA
Advisor: Dinesh Jayaraman, Junyao Shi Aug 2025 - Present
 - Designing retrieval-augmented VLA models for robot manipulation (ongoing, target: RSS 2026).
 - Building scalable data pipeline, training and evaluation infrastructure with Docker, Slurm, JAX, Ray, and Hugging Face on GPU clusters.
- **Learning and Interactive Robotics, University of Virginia** Charlottesville, VA
Advisor: Yen-Ling Kuo Aug 2024 - May 2025
 - **Interpretable Vision-Language-Action Models via Skill Conditioning**
 - * Led *SkillVLA*, a skill-conditioned VLA model for language-conditioned manipulation with improved action interpretability via subgoal instructions and learned skill library.
 - * Presented as an oral talk at the 2024 UVA LLM Workshop ([slides](#)), earning the Audience Choice Award (top 3 of 28 presentations).
 - **Contrastive Learning for Robot Manipulation**
 - * Performed contrastive learning over action sequences to learn behavior-grounded visual embeddings, improving visuomotor policies under heterogeneous camera poses and object appearances. ([CoRL 2025](#))
 - * Modified simulation environments and trained PyTorch-based visuomotor policies to evaluate learned embeddings.
- **Research Assistant, Collaborative Robotics Lab, University of Virginia** Charlottesville, VA
Advisor: Tariq Iqbal May 2022 - May 2024
 - **Grounded Location for Object Manipulation (GLOMA)** [[code](#)]
 - * Led team of 3 to develop zero-shot image-editing model grounded by language instructions for object relocation and manipulation tasks, designed for downstream robotic applications using goal-conditioned RL and Behavioral Cloning (BC).

- * Integrated language grounding with visual perception by using bounding box guidance from pre-trained language models, enabling precise object relocation without external supervision and improving baseline performance by 65%.
- * Collected and annotated custom dataset for fine-tuning pre-trained language and vision models.
- * Presented poster at 3 conferences and symposiums.
- **Centralized multi-agent RL for Collaborative Tasks** [code]
 - * Developed long-horizon on/offline centralized MARL for robotic bolt screwing tasks.
 - * Designed and optimized custom reward functions in multi-agent framework for task completion and agent collaboration, improving task success rate by 20%.
 - * Deployed and tested custom simulated environments in IsaacGym for training and evaluation.

PUBLICATIONS

Lee, Sung-Wook, Xuhui Kang, **Brandon Y. Yang**, and Yen-Ling Kuo. “Class: Contrastive learning via action sequence supervision for robot manipulation.”

In *Conference on Robot Learning*, pp. 4743-4766. PMLR, 2025.

[Website][arXiv][PDF]

Sethi, Amish*, **Brandon Y. Yang***, Yuchen Zheng, Jiani Huang, Jianing Qian, Chris Watson, Junyao Shi, Mayur Naik, and Dinesh Jayaraman. “Retrieval-Augmented Vision-Language-Action Model”.

Ongoing project.

SOFTWARE PROJECTS

notie-markdown [Website]: Markdown rendering web app with support for equation previewing, graphing, code running, and more. Built with React, TypeScript, Python, Flask. Deployed on Vercel and Heroku.

Blogs and Notes [blog, notes]: Detailed blog posts and course notes with visualizations, graphs, and code on CS and AI topics, written with notie-markdown.

SmartOH [GitHub]: Queue management system designed for office hours. Features real-time queue updates, notifications, and analytics. Built with Next.js, TypeScript, Python, FastAPI. Deployed with CI/CD pipeline on Vercel and AWS EC2.

Voy: Collaborated with 7 non-profits to develop Voy, a volunteer and driver management platform; competed and received \$1000 in funding from UVA’s Entrepreneurship Cup. Built with React, Node.js, TypeScript, JavaScript.

SKILLS

Programming Languages: Python, C/C++, CUDA, Java, JavaScript, TypeScript, Coq

ML + Robotics: PyTorch, Jax, ROS, HuggingFace, IsaacGym, Habitat, RLBench, Maniskill

Tools & Frameworks: Docker, Slurm, Ray, AWS, Git, CI/CD, React, Next.js