Research Interests: We are now in a new renaissance era of robotics and artificial intelligence. Consider the Large Language Model (LLM) as an example; once an academic dream swiftly transforms into a practical reality. Rather than online AI assistance like ChatGPT in a virtual environment, using robots in the real world would face significant challenges such as zero-shot generalization ability, task-level intelligence, safety constraints, and robustness against physical uncertainty. Captivated by Isaac Asimov's robot novels from a young age, I have long pondered whether we can develop robotic systems capable of moving and interacting with humans and environments in a clever, stable, and efficient manner.

Research Experiences: My research begins with integrating visual perception and manipulation. I first engaged in a vision-based FDM-printing project: ***Vision-based FDM Printing for Fabricating Airtight Soft Actuators*** (RoboSoft 2024 Oral Presentation)[1]. This project enhances the fabrication quality of desktop FDM printers by correcting defects during printing with a webcam. My contributions to this project were: first, propose a layer-wise image segmentation method to map the G-code contours with each layer's captured images, and second, provide a computation efficiency solution for real-time defects segmentation using the Laplacian of Gaussian method. Ultimately, airtightness was achieved for printed soft pneumatic actuators.

Furthermore, considering human-level vision, we can capture 3D geometrical information in the real world easily. I raised my interest in 3D point cloud tasks. A crucial prerequisite for vision-based robotics manipulation is the availability of comprehensive data input, and I finished a paper as **co-first author** focused on loss function optimization for deep learning-based point cloud completion: ***Loss Distillation via Gradient Matching for Point Cloud Completion with Weighted Chamfer Distance*** (IROS 2024 Oral Presentation)[2]. In this project, I proposed a family of CD-based losses (weighted CD) using a gradient weighting scheme to mimic the teaching neural network learning behavior and developed a bilevel optimization formula to train the backbone network based on the weighted CD loss, which needs no hyperparameter tuning. In the end, I conducted comprehensive experiments with novel networks in real (KITTI) and synthesis (ShapeNet) datasets to examine the findings.

Beyond focusing solely on point cloud completion. I have realized the importance of high-level distributional tools to guide neural networks in learning robust representations, which are essential for solving a wide range of set-to-set matching problems and our understanding of such vision tasks. Inspired by previous work, I explored statistical learning methods and extreme value theory to identify shared distributional representations that can address distinct tasks, such as point cloud completion, zero-shot image classification, and sparse point cloud registration, which can all be identified as set-to-set matching. To this end, I finished a paper as **co-first author**: ***GPS: A Probabilistic Distributional Similarity with Gumbel Priors for Set-to-Set Matching*** (ICLR 2025). In this project, I proposed a similarity learning framework that measures the similarity between the underlying distributions generating the sets. It is achieved by learning a Gumbel prior with minimum distances between the set items to maximize the log-likelihood. This was motivated by the observation that the distributions of minimum distances metrics, as encountered in the aforementioned tasks. Next, I formulated a reparametrizing technique based on distribution shift. Finally, I demonstrated multiple analyses; training efficiency dramatically improved compared with Optimal Transport methods in zero-shot image classification, completion results showed both visual and metrics improvements, and registration recall was boosted in sparse point cloud registration.

The aforementioned projects developed my knowledge of deep learning, optimization, probabilistic methods, and research mindset. However, to reach the ultimate intelligence, the only study on vision tasks cannot satisfy my research ambitions. Robot learning, a complex system of interaction and reasoning with the real world like humans, came to mind. To this end, I joined Professor Robert Platt's The Helping Hands Lab as a research intern student. My first engaged project is "***Imagination Policy*** (CoRL 2024)[3]" guided by Haojie Huang. In this project, my contributions are the following. First, I investigated various baseline models in robot manipulation learning. Then, I was responsible for the real robot experiments. At last, I introduced an articulate object task (open microwave) and a multi-step task (stack chairs) to show the generalization ability of our method.

Based on the Imagination Policy, we finished a follow-up project: "***Match Policy*** (ICRA 2025)[4]," a plug-and-play policy designed for leveraging point cloud registration methods, which needs no training after demo collection. The "Match" is to capture vital local geometry features and calculate the correspondence between objects. I am responsible for implementing point cloud registration methods onto real robot demos. Here are the details of this approach: First, our data inputs of manipulation objects and test environments are all represented by point clouds. Second, the problem can be described as, given point clouds P_a, P_b and P_{ab}, where P_{ab} shares a similar part with P_a and P_b but there is no overlapping between P_a and P_b. Our goal is to infer two rigid transformations T_a and T_b to match the configuration illustrated by P_{ab}. To this end, such point cloud registration methods can effectively capture both local and global geometrical features and then calculate the transformations for robot manipulation, which satisfies our requirements.

These two projects astonished me with their incredible geometrical learning abilities and promising performance in robot learning by applied equivariance property. And I will continue to study related knowledge to develop my original contribution.

Research Vision and Role of NEU: My goal is to work in academia, contributing to foundational research and interdisciplinary problem-solving. At NEU, I aspire to continue working on geometrical learning to empower the generalization of robot learning for manipulation, i.e., skill learning. Furthermore, task-level reasoning and planning, i.e., task and motion planning (TAMP), are underexplored in robot learning methods. My PhD goal is to leverage an elegant bridge between the TAMP and robot skill learning. The robot can execute stably under a complex task domain by understanding geometry to develop new skills and doing dynamic replanning to comprise uncertainty. NEU's collaborative environment and research strengths, particularly research work from Professor **Robert Platt**, Professor **Robin Walters**, and Professor **Lawson Wong**, align perfectly with my aspirations.

Community and Leadership Contributions: I am committed to promoting inclusivity and access to education for underprivileged communities through workshops, teaching assistant roles, and leadership in Robot Learning groups. I believe my experiences as an international student and advocate for diversity will contribute positively to NEU's graduate community.

References

[1] Y. Wu, Z. Dai, H. Liu, L. Wang, and M. P. Nemitz, "Vision-based fdm printing for fabricating airtight soft actuators," in *2024 IEEE 7th International Conference on Soft Robotics (RoboSoft)*. IEEE, 2024, pp. 249–254.

[2] F. Lin, H. Liu, H. Zhou, S. Hou, K. D. Yamada, G. S. Fischer, Y. Li, H. K. Zhang, and Z. Zhang, "Loss distillation via gradient matching for point cloud completion with weighted chamfer distance," *arXiv preprint arXiv:2409.06171*, 2024.

[3] H. Huang, K. Schmeckpeper, D. Wang, O. Biza, Y. Qian, H. Liu, M. Jia, R. Platt, and R. Walters, "Imagination policy: Using generative point cloud models for learning manipulation policies," *arXiv preprint arXiv:2406.11740*, 2024.

[4] H. Huang, H. Liu, D. Wang, R. Walters, and R. Platt, "Match policy: A simple pipeline from point cloud registration to manipulation policies," *arXiv preprint arXiv:2409.15517*, 2024.