Portfolio

"Shaping the Future of Robotics with Vision, Passion, and Innovation."

Two-Part Portfolio: Projects, Research, and Personal Statement

Author: Tianyi Xia

Date: March 2, 2025

The two-part Portfolio comprises my projects, research, and Personal Statement.

Contents

Projects and Researches Demonstrations

I'll illustrate some projects and research here (instead of GitHub) due to privacy reasons.

1.1 Design of a social balloon robot in domestic scenes $'24 Summer$

The research was funded by the Access Computing Summer Program of the Global Innovation Exchange of Tsinghua University and the University of Washington. I worked as a student researcher at the Pervasive HCI Laboratory, proposing and developing the first social balloon robot from the concept. I completed a conference paper for CHI '25, which is currently under the second round review (the passing rate of the first round is $< 50\%$). Due to the anonymous principles, I can not share the full paper until it is accepted.

In general, the paper aims to draw a clearer picture of using a balloon robot as an interactive and friendly social agent. By developing a prototype and its web user interface, we first proposed a series of use cases that are grounded on the application topics in the relevant literature and characteristics. We then conducted an open-ended lab study (n=33), which comprised a **stimuli session** using a pre-recorded video showcasing these use cases and a **hands-on session** where participants were allowed to operate, check, and touch the balloon robot. In particular, the video was crafted using a Wizard-of-Oz (WoZ) design. To the best of our knowledge, this paper is the first to explore users' feedback on a social balloon robot with experiments using a robot prototype and WoZ scenarios. Finally, each participant was invited to complete a questionnaire and participate in an on-site semi-structured interview. Looking into the results, we provided insights on how the balloon robot is perceived by users from different dimensions, their diverse expectations about the functionality it may carry, and concerns it should address in the future.

After conducting an in-depth review of the literature on social robots, including drones, ground robots, and desktop robots, and studies specifically related to balloon robots, I envisioned the first-ever balloon robot with social attributes. Figure 1 is a simple depiction of "A Day with My Social Balloon Robot."

To better leverage the balloon's unique physical characteristics for social functions, I designed and complemented a hardware platform, shown in Figure 2.

In addition to designing the hardware platform and the control system, the most challenging part is designing downstream use cases while considering common needs. We consider two key characteristics of the balloon robot, namely **spatial maneuverability** and **touchability**. In addition, we deem the balloon robot a perfect platform for intelligent multimodal interactions that have been implemented in other robot types, e.g., natural communication, target recognition, and pose estimation, thereby considering **multimodal intelligence** as another characteristic in its social functioning. Specifically, we propose the following applications: Fitness Coach, Housekeeper, Emotional Companion, and Indoor Navigator, shown in Figure 3.

We fulfilled each function and applied LLM(for conversation) and object detection and segmentation algorithms for autonomous interactions, shown in Figure 4 and Figure 5.

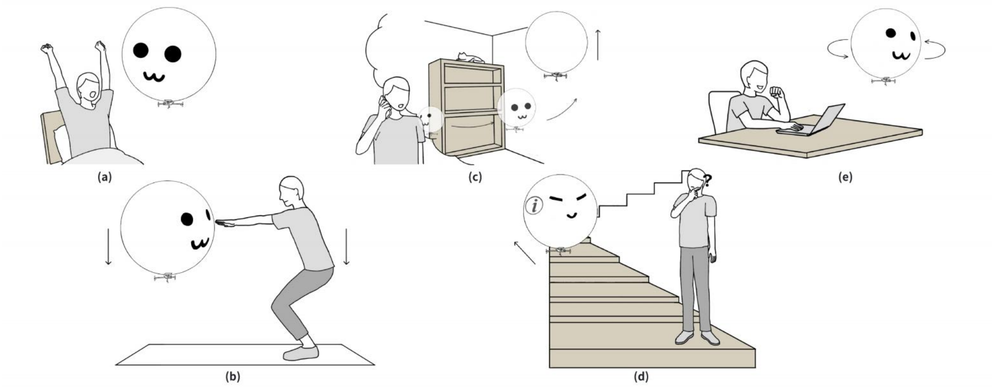


Figure 1: Daily use examples of the balloon robot in domestic scenes. (a) At 8:00 a.m., the balloon robot flows over to wake up the user and provides a cheerful greeting. (b) One hour later, it becomes a fitness coach, guiding the user with touches. (c) At 1:00 a.m., the balloon robot is the housekeeper, helping locate a roaming cat at home. (d) By 3:00 a.m., the balloon robot in the library assists the user in book searching across different building levels. (e) Finally, the balloon robot provides relaxing and supportive companionship while the user works on their laptop.

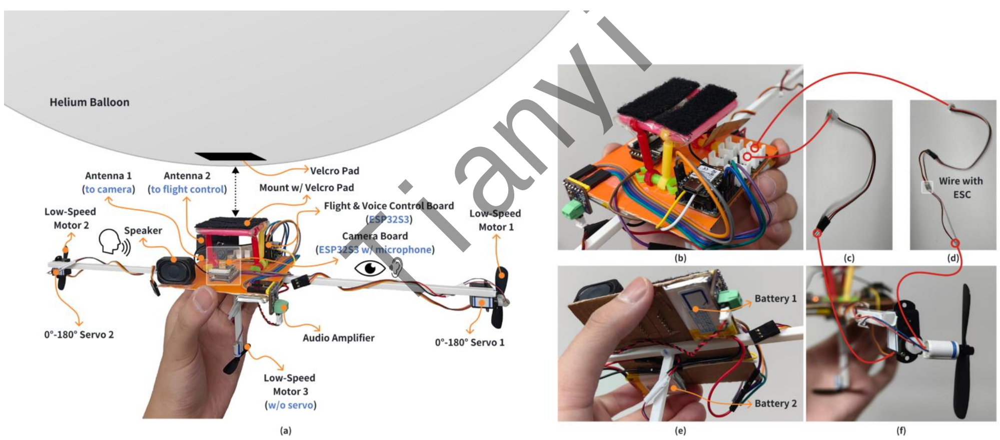


Figure 2: (a) An overview of the balloon robot. (b) The upside panel from the backside view, where the flight & voice control board is connected via the wires shown in (c) and (d) with the servo and the motor on each end, respectively. It is also connected via another ESC-equipped wire with the motor at the bottom. (e) Two batteries are attached to the downside panel. (f) The lateral view of a servo and a motor.

After implementing all the functions, we invited 33 participants for an open-lab study. This study employed a 5-point Likert scale questionnaire to evaluate participants' familiarity with robots, attitudes (using NARS), interaction perceptions, and functionality expectations of the balloon robot. Key dimensions assessed included safety, noisiness, empathic alignment, ease, and privacy. Additionally, specific use cases like fitness coaching and housekeeping were examined in detail. Follow-up semi-structured interviews provided qualitative insights, with content analysis used to identify patterns in participant responses, complemented by

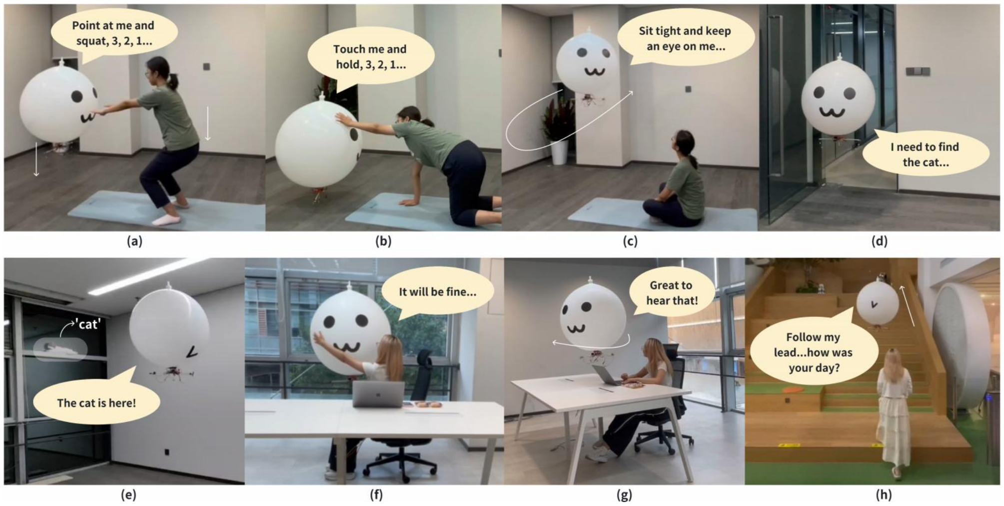


Figure 3: We prepared a stimuli video covering the four use cases, namely (a-c) fitness coach, (d-e) housekeeper, (f-g) emotional companion, and (h) indoor navigator. These demo contents were created with a Wizard-of-Oz design.

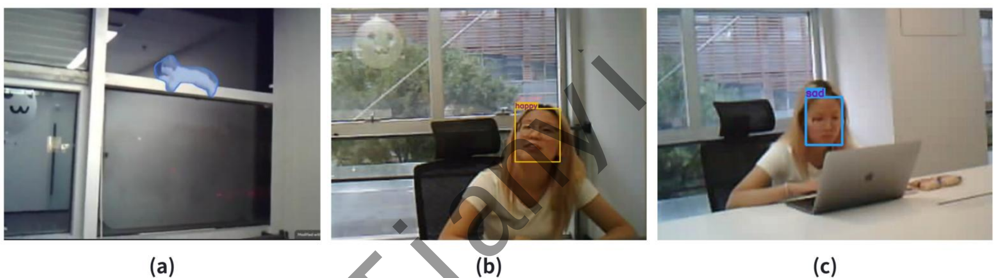


Figure 4: Results of image segmentation and emotion detection. (a) shows the image segmentation result when the balloon robot is finding a cat. (b) shows the detected happy emotion, while (c) shows the sad emotion.

quantitative analysis for correlations between familiarity, attitudes, and perceptions. Non-parametric tests (e.g., Spearman's correlation) were used due to data non-normality, yielding correlation coefficients (r_s) and error rates (p).

We prepared a web user interface for the wizard's real-time control of all the interactive channels of the balloon robot; the video and sound captured by the robot are also streamed back to this interface. Participants could use keyboards to control. All the communications are built based on the network connection, while the control boards serve as the server and the local computer is the client, as shown in Figure 6.

1.2 Experiments design and analysis on patients with cerebellum degradation 02/2024–07/2024

Please refer to the link Connect4 game design and Behavior Modeling.

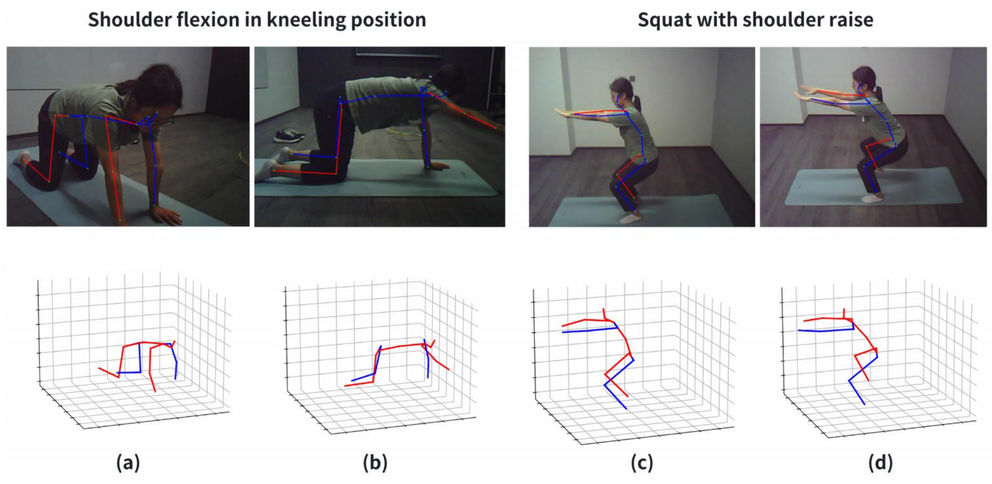


Figure 5: Visualization of the 2D and 3D pose estimation results of a user performing two different exercise movements (a, b and c, d). Additionally, (c) and (d) demonstrated the corrections made to the user's form of squat.

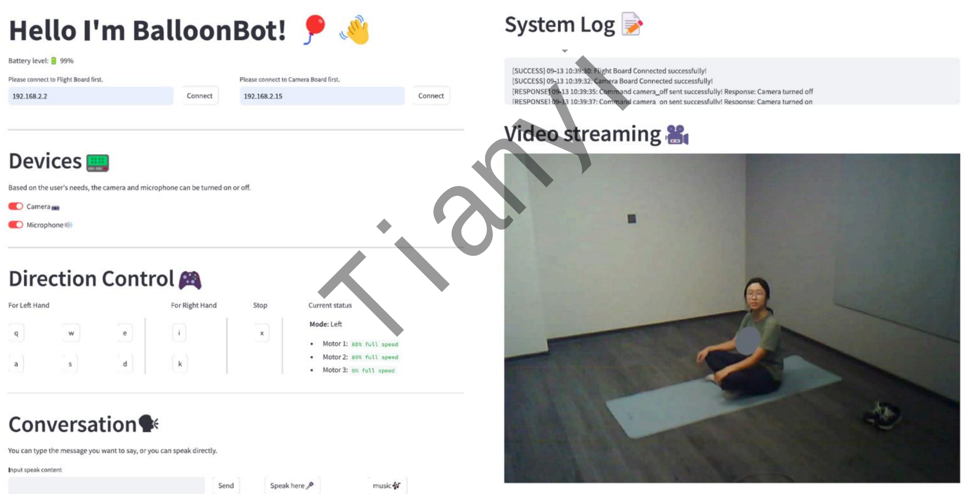


Figure 6: WebUI: built on Streamlit

1.3 Target Tracking and Aiming System in Complex Motion Environment 09/2022 – 07/2023

In Robomaster, the robot's auto-aiming system assists the operator in targeting the opponent robots' armors, significantly improving projectile accuracy by quickly locking onto and striking the armors.

Figure 7: My balloon robot and Me.

1.3.1 Target Detection and Recognization

This section primarily introduces the implementation of the auto-aiming recognition process using traditional and advanced algorithms.

Traditional Algorithms

The recognition of armor with traditional algorithms can be divided into two main parts. The first part uses conventional image processing techniques to identify all possible contours of armor plate light strips. These contours are then paired in all possible combinations, resulting in groups of light strips that meet the armor's characteristics. The bounding rectangle of each group is considered a candidate armor plate. If operating in Track mode, the search area for the armor plate is narrowed from the full image to the ROI (Region of Interest) around the armor plate detected in the previous frame, thus speeding up the filtering process.

The second part involves using a neural network to build a simple classifier. All candidate armors from the first step are fed into the classifier, which determines their respective categories while filtering out any misidentified plates from the first step. This process ultimately provides the corner point information for all genuine armors in the image, which can be showcased in Figure 8. The algorithm results are showcased in Figure 9 and Figure 10.

Armor Plate Recognition Based on NanoDet Four-Point Model

The NanoDet-based four-point model for armor plate recognition provides an end-to-end solution for locating and classifying armors in competition scenarios. Unlike traditional methods that require preprocessing and classification in separate steps, NanoDet directly predicts classification and four-corner coordinates of armors with high accuracy and confidence.

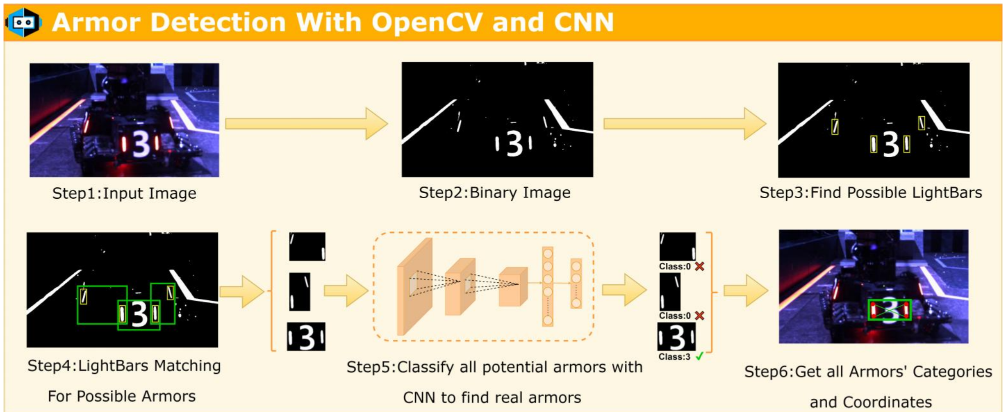


Figure 8: Process diagram of traditional auto-aiming algorithm

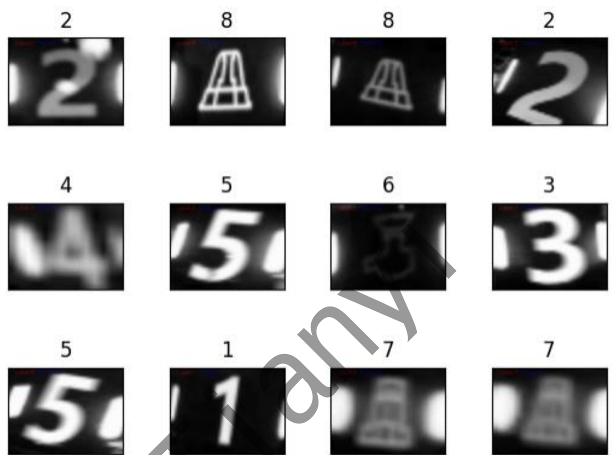


Figure 9: Confusion matrices based on the accuracy of each category.

This significantly simplifies the process and enhances robustness in varying environments. Moreover, the lightweight nature of NanoDet makes it suitable for deployment on mobile platforms, addressing the limitations of larger models like YOLO and SSD.

To meet the requirements of precise pose estimation, the model was modified to detect the four corner points of armors instead of bounding boxes. This involved preparing a custom dataset with annotated corner points, modifying the network structure, and training it using the AutoDL cloud platform to leverage advanced GPUs and save training time. The trained model was deployed using OpenVINO, optimized for Intel hardware, ensuring fast inference speeds and efficient resource utilization.

The model incorporates dynamic label assignment strategies, replacing ATSS with Matching Cost-based methods to improve the accuracy of positive and negative sample selection during training. Additionally, the Generalized Focal Loss was adapted for the four-point detection task by integrating keypoint-based MSE loss, enhancing the precision of corner predictions. Non-Maximum Suppression (NMS) was further refined with cross-category operations to minimize duplicate outputs for the same target.

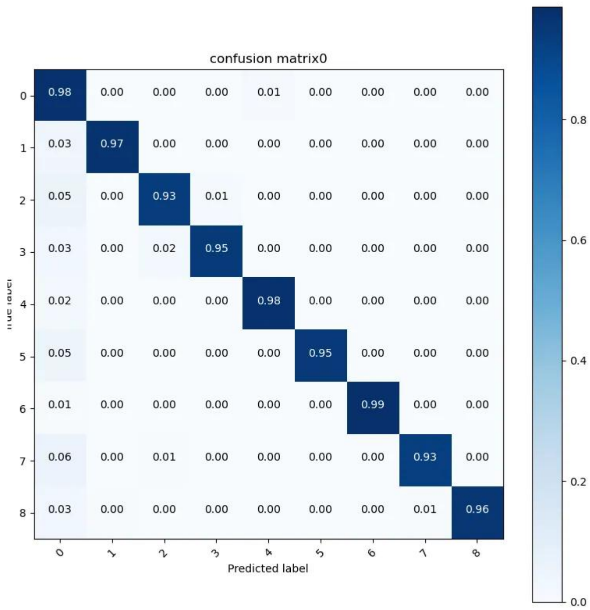


Figure 10: Classification result display, armor 1-5 correspond to categories 1-5 respectively, category 6 is sentinel, category 7 is an outpost, and category 8 is the base.

While the model demonstrates strong performance metrics (classification accuracy: 97%, precision: 95.2%, recall: 94.7%) and deploys effectively on lightweight hardware, challenges remain. The computational complexity of the four-point loss and the inherent trade-offs of lightweight models limit potential accuracy improvements. To address these, further optimizations in handling key points, backbone structure experimentation, and hyperparameter tuning are planned. These advancements aim to refine the model's adaptability to complex scenarios and further enhance its performance. The recognition result is in Figure 11.

Figure 11: NanoDet recognition example

1.3.2 Motion Modeling and Prediction

The auto-aiming algorithm in competitions processes the motion information of enemy robots and predicts the potential position of the armor plates in the next moment. This prediction serves as a crucial basis for calculating the gimbal angles, which is then provided to the control system. The core information used by the algorithm is derived from the three-dimensional (3D) data of the target armor plate in the world coordinate system. The algorithm filters the 3D data and estimates the optimal position, velocity, and acceleration of the target armor plate, which can then be used for subsequent gimbal angle adjustments and projectile trajectory calculations.

For coordinate conversion, we developed a detailed strategy using the Perspective-n-Point algorithm(PnP) with Bundle Adjustment(BA). To determine the armor plate's center in the absolute coordinate system, transformations are performed between three coordinate systems:

- **World Coordinate System:** Represents the armor plate's position.
- **Camera Coordinate System:** Varies with gimbal movement.
- **Absolute Coordinate System:** Fixed at the intersection of the yaw and pitch axes.

World Coordinate System to Camera Coordinate System The first transformation, from world to camera coordinates, uses the *solvePNP algorithm*, which provides the rotation matrix R and translation vector T. This transformation expresses the relationship as:

$$P_c = R \cdot P_w + T$$

where P_w is the world coordinate point. By defining the armor plate center as the origin in the world coordinate system, T directly represents the armor plate's position in the camera coordinate system. This approach simplifies the computation and aligns both systems.

Camera Coordinate System to Absolute Coordinate System The transformation from the camera to the absolute coordinate system involves three steps:

1. **Translation:** Aligns the camera origin with the absolute coordinate system origin. This uses pre-measured offsets for x, y, and z, handled programmatically.
2. **Axis Adjustment:** Converts the camera's coordinate system (right-hand system with x-right, y-down, z-forward) to the conventional right-hand system.
3. **Rotation:** Aligns the rotated camera axes with the absolute coordinate axes. The rotation uses the yaw and pitch angles, which are obtained from the control system.

The yaw is defined as a counterclockwise rotation around the z-axis, and pitch as a counterclockwise rotation around the x-axis.

Accounting for Misalignment Between Camera and Gimbal In practice, the camera and gimbal are misaligned by a 5° angle. While the provided yaw and pitch data from the control system already account for this offset, adjustments are needed:

- The transformation proceeds as usual with translation and rotation.

- Adjustments are made to the translation vectors to account for the camera's displacement from the gimbal's axis. These adjustments are calculated using geometric relationships and mechanical blueprints.

Camera Parameters and Modularization To enhance flexibility, camera intrinsic parameters are loaded dynamically from an external XML file. This avoids hardcoding and allows for parameter updates without recompiling the program. The parameters include Camera Intrinsic Matrix and Distortion Coefficients. The parameters are loaded via a `loadParam` function, which uses OpenCV's `FileStorage` to read the XML file. This modular approach improves maintainability and adaptability.

This season, the algorithm utilizes **yaw, distance, pitch, and their corresponding velocities and accelerations** as state variables. Building on the constant velocity model, the algorithm employs Kalman filtering and Extended Kalman filtering (EKF) (UKF was tested). To enhance noise adaptability and responsiveness, adaptive filtering techniques are integrated, significantly improving accuracy and performance in complex and dynamic environments. The algorithm demonstrates excellent performance under laboratory conditions. For stationary targets, the hit rate is consistently 100%, regardless of the distance. For targets moving in a sinusoidal pattern, the hit rate reaches 90% at 3.5 meters and 80% at 5.5 meters. For dynamic targets with random motion, the hit rates are 80% at close range and 60% at long range.

The algorithm has several advantages. Its Kalman filter-based framework, combined with EKF and adaptive algorithms, effectively models the system. It operates with minimal memory requirements by relying on current and previous data, enabling fast iteration speeds. On the control platform, the frame rate remains stable above 100 FPS. Furthermore, the algorithm achieves high accuracy, with an 80% hit rate for small armor plates at 5.5 meters in the lab, while maintaining stable gimbal performance with minimal shaking. However, the algorithm also has drawbacks. The use of nine-dimensional data increases the complexity of parameter tuning, making the adjustment process challenging and time-consuming.

To further enhance the algorithm's performance for predicting moving targets, several optimization measures were applied. Visualization tools like Matplotlib and OpenCV were used to display filtering, prediction, and hit results, aiding parameter tuning. Parameter initialization methods would be refined to improve convergence speed and accuracy. Adaptive algorithms would be further optimized to reduce noise effects and increase robustness in complex environments. Additionally, testing and compensating for discrepancies in the control system's positional data exchange would help improve overall hit accuracy.

You can refer to the link Video demonstrations for more exciting results.

Moreover, I designed decision-making algorithms based on the different priorities of the target and developed the functions of localization, navigation, obstacle avoidance, and others for the unmanned control robot (a special kind of robot in the competition)(in Figure 13). Apart from vision algorithm design, I am responsible for communication maintenance (e.g., serial or CAN communication between the main host NUC and the main control board), code repository maintenance, testing of new algorithms and stability, and script development.

Figure 12: Armor recognition and coordination solution result

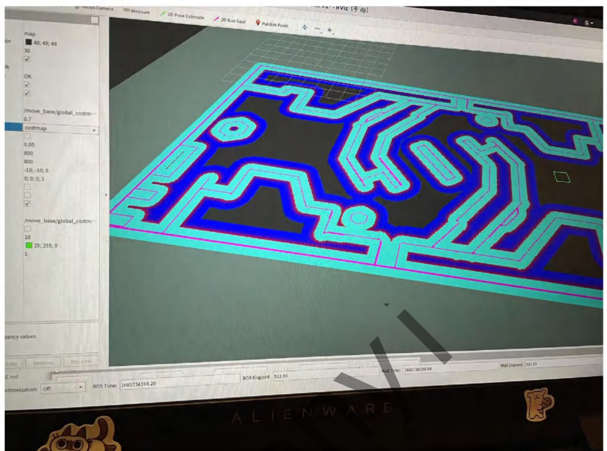


Figure 13: The map of the real competition environment in Rviz, which was built by simulations in Gazebo.

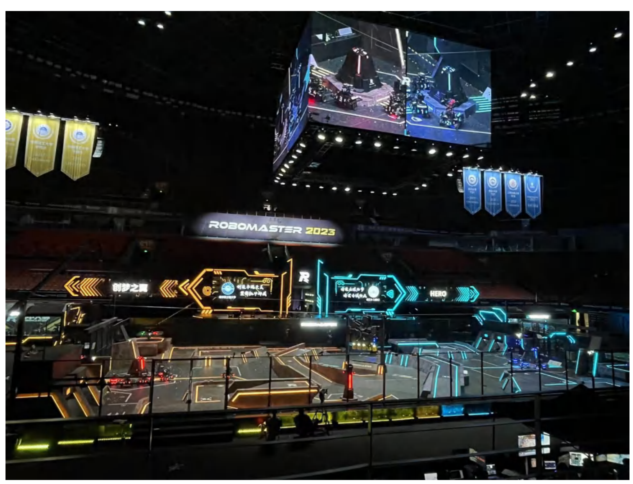


Figure 14: The competition environment in the final.

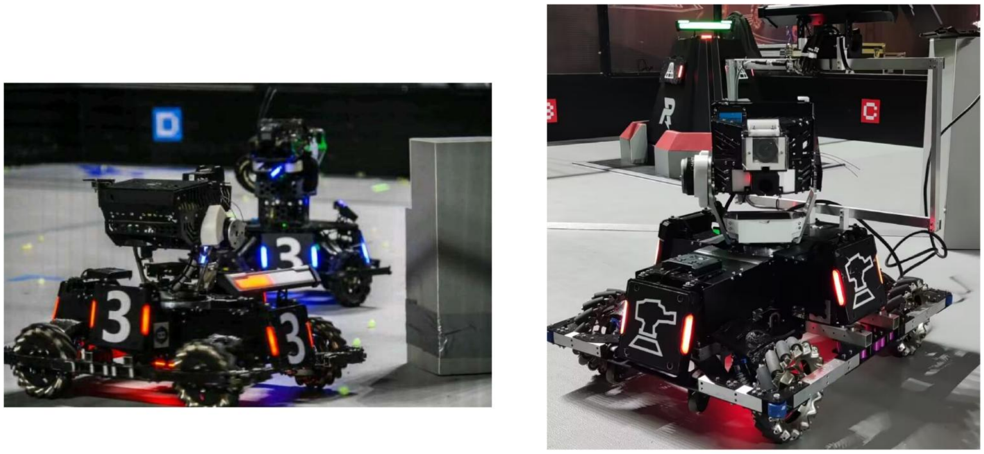
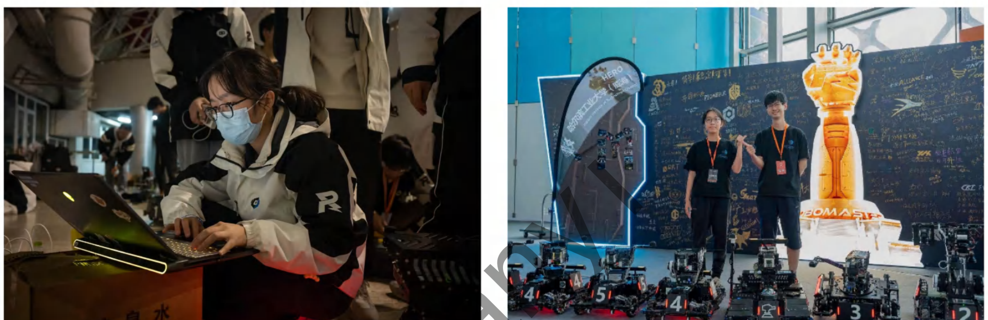


Figure 15: Two robots of our team.



I was testing and debugging between matches.

My friend and me after the final of the National Competition at Shenzhen.

Figure 16: Some photos of me as the team member.

1.4 Distributed Algorithm for Matrix-variable Optimization Problems with General Constraints 09/2022–09/2023

This is my Mathematic undergraduate thesis (still working on improvement), and please refer to the link Math thesis.

1.5 The Automated Guided Vehicle Based On Jetson Nano For Delivery 01/2022–08/2022

Please refer to the link Online Selection Tasks and Regional and National Competitions.

1.6 Automatic tracking and block grabbing vehicle based on STM32 11/2021

Please refer to the link Final Report.

1.7 Human-Robot Interaction and Policy Learning for Surgical Robotics *Ongoing*

I am currently collaborating with professors from the University of Macau and the Chinese University of Hong Kong, focusing on surgical robotics and targeting IROS 2025.

1.8 Design of a Vehicular Network Data Authentication System *Ongoing*

This is my undergraduate computer science thesis, and it is now ongoing. Please refer to the updating file CS thesis (ongoing).

1.9 Computer Science Course Projects

Please refer to the link CS Course Projects.

1.10 Mathematical Modeling Competitions

Please refer to the link documents.

Personal Statement

Besides academics, I am eager to demonstrate more about my background and personal experiences, sharing how they have profoundly shaped my character, values, and aspirations.

2.1 A Journey of Growth and Passion

I have always believed that my real freedom comes from living beyond labels and embracing the abundant opportunities to find my real self. My journey has been influenced by challenges, cultural insights, and a passion for technology, always breaking the boundaries of traditional expectations and finding my own path. I played Go and the piano in childhood, which nurtured my perseverance, creativity, and problem-solving skills. Winning a provincial first prize in the Chinese Physics Olympiad deepened my fascination with physics. Besides that, I am a sports enthusiast, particularly sprinting and soccer, where the training taught me about concentration, resilience, and that we need to go beyond the limit. As a first-generation college student, I had limited access to STEM resources in my hometown, which made entering university exciting but challenging for me. However, I quickly realized that pursuing excellence meant stepping outside my comfort zone. Pursuing a Mathematics major with a double major in Cyberspace Security, I actively embraced challenges and explored interdisciplinary applications, particularly in robotics and engineering, to combine theory with real-world impact.

2.2 Building Dreams Through Robotics

My 2-year experience with the university's robotics team was a milestone that emphasized the value of teamwork, persistence, and the passing down of technical knowledge. Motivated by our team's motto, 'Chasing boundless craziness, building earnest dreams together,' we dedicated countless nights to coding, debugging, and testing, passionate about our shared goals. Teamwork was at the core of everything we achieved. Collaborating with talented peers taught me to trust in others' strengths, communicate effectively under pressure, and adapt to challenges as a unified group. Persistence became a necessity during intense competitions and demanding preparations, reminding me that breakthroughs often come from embracing challenges and pushing through setbacks. One of the most meaningful aspects was the opportunity to pass down knowledge and skills to younger team members. Each year, we held selection tests and prepared lecture sessions to teach essential concepts to beginners. Whether it was explaining algorithm design or demonstrating hardware integration, I found immense fulfillment in sharing what I had learned and contributing to the team's technical environment and spirit. Additionally, maintaining detailed notes for every task is crucial to our team's success, a practice from which I greatly benefited and was deeply committed. These experiences honed my technical and leadership skills and cultivated my good habits of organization, mentorship, and continuous learning.

2.3 Embracing Diversity and Discovery

My semester as a visiting student at UC Berkeley further broadened my horizons and strengthened my commitment to diversity and inclusion. Initially, I was worried about being in a totally new environment, considering the potential language barriers, course pressure, and meeting new people. However, I was fortunate enough to be surrounded by talented and warm-hearted people. Immersing myself in AI and Machine Learning courses, I expanded my technical foundation and built lasting relationships with peers from a wide range of cultural and academic backgrounds. We shared our common knowledge but with different perspectives, which is a super exciting way to explore. It was a bit challenging for me to attend the Stochastic Process course due to the fast pace. However, the professor was so nice and patient, always taking the time to explain every detail of my questions after class. And I can't forget his words in our farewell email: "You should take confidence in your performance under such drastically new circumstances. I am sure that you will continue to improve if you apply yourself with the same determination that you showed this past semester." which touched and motivated me a lot. When I was seeking an RA position, I was warmly welcomed by another professor, acknowledging my status as a newly arrived visiting student. Thankfully, I was fortunate to work as a research assistant in his lab at the Helen Wills Neuroscience Institute, I explored the intersection of neuroscience and AI, designing cognitive experiments and analyzing data to better understand human behavior. This experience deepened my appreciation for interdisciplinary collaboration and cultural exchange as I learned to navigate complex challenges with the guidance of a warm, helpful mentor and professors whose expertise and support left a profound impact on me.

Beyond academics, I made lasting friendships, traveled widely, and embraced new experiences that broadened my horizons and personal growth. I was taken good care of by people, school, and the community. Those vivid memories stay with me, like noticing the abundance of disability-friendly infrastructure, which offered me a new lens through which to view inclusivity in public spaces. Watching a baseball game in Los Angeles gave me a first-hand sense of community and passion that transcends cultural boundaries. At the Berkeley Marina, my sailing class was a testament to the collaborative spirit of the campus—although we were strangers at the start, everyone helped and supported one another as we learned the ropes together. I'll never forget the nights spent with new friends playing board games, sharing cultural experiences like travel stories, festivals, and traditions, and even debating a variety of topics while gathered around the sofa.

This experience made me braver in facing new challenges. I traveled alone in San Francisco and Taiwan, engaged warmly with locals, and embraced every moment with an open heart. Academically, I became more confident in tackling challenges, motivated by the resilience and determination I gained during my time at Berkeley.