YOLOv9 Team Description Paper

Mr.Chayakorn Winothai , Ms.Natwalee Narhkthong , Mr.Phattaraphon Suttiphan and Mr.Pichanat Sittipunyota

BANBUNG UTTASAHAKAMNUKHRO SCHOOL
170 CHON BURI-BAN BUENG RD,
BAN BUENG, BAN BUENG DISTRICT, CHON BURI 20170

Abstract. The "YOLOv9" team, having secured a place in the Final 4 of the @Home Education Challenge Thailand Open 2024 in the High School category, is now preparing to participate in the RoboCup Malaysia @Home Education Challenge. Building on their success in Thailand, the team continues to push the boundaries of domestic robotics.

The TurtleBot 2 platform, chosen for its versatile software customization and development capabilities, remains at the core of the team's efforts. In this next phase, "YOLOv9" focuses on refining the robot's artificial intelligence (AI) to improve its autonomy, efficiency, and human-robot collaboration. Enhancements in areas like object recognition, indoor navigation, and decision-making are central to their strategy, enabling the TurtleBot 2 to perform a wide range of household tasks with minimal human intervention.

Participating in the RoboCup Malaysia represents a valuable opportunity for the team to expand their knowledge and skills, engaging with a broader community of robotics enthusiasts and professionals. The competition allows "YOLOv9" to test their developments in a new environment while contributing to the global advancement of domestic robotics.

Driven by a passion for robotics and a vision of intelligent robots as integral to everyday life, the "YOLOv9" team aims to solidify its reputation as a leader in high school-level robotics. Their participation in these competitions lays the groundwork for future innovations that could bring autonomous robots into homes around the world.

Keywords: YOLOv9 team, RoboCup Malaysia, TurtleBot 2, @Home Education Challenge Thailand Open 2024, artificial intelligence (AI), human-robot collaboration, domestic robotics

1 Introduction

The "YOLOv9" team was formed by a group of individuals passionate about artificial intelligence (AI) and domestic robotics. The team members share a strong interest in robotics and have gained substantial experience through participation in robotics competitions organized by the Office of the Basic Education Commission (OBEC) of Thailand. Notably, we competed in the ASEAN-level CPU’1: ASEAN Grand Prix Youth Robotic Competition 2023, where we showcased our expertise in both software and hardware development.

Domestic robots represent a significant response to societal needs for reducing time spent on household chores and enhancing the quality of life. By leveraging their knowledge and skills, the team aims to develop robots that efficiently meet these demands. The growing interest in domestic robotics is driven by the potential of robots to alleviate household burdens and provide residents with more leisure time. The "YOLOv9" team is dedicated to advancing the capabilities of these robots, enabling them to autonomously perform tasks such as cleaning, cooking, and managing various household activities. This not only makes daily life more convenient and efficient but also supports the independence and safety of elderly individuals and those with physical disabilities.

The development of domestic robots is a critical step toward improving quality of life and simplifying daily routines. The team's approach to solving challenges in the @Home Education competition focuses on key areas such as gesture recognition, robot control, navigation, human tracking in both familiar and unfamiliar environments, voice interaction, and human-robot interaction. These elements are essential in creating robots that can assist with household tasks autonomously with minimal user intervention. The use of sensors and AI enables the robots to adapt to and respond to changing home environments effectively.

The primary theories underpinning the development of domestic robots include Machine Perception and Adaptation Theory, which emphasizes the robot’s ability to perceive its environment through sensors, process information to make decisions, and adjust its actions accordingly. This allows robots to operate independently in dynamic environments. Additionally, Machine Learning Theory plays a fundamental role in enabling robots to improve their performance by learning from their experiences.

Participation in RoboCup Malaysia by the "YOLOv9" team presents a valuable opportunity to test and evaluate the potential of their developed robots. It also serves as a platform to demonstrate advancements in robotics and AI, which can be utilized to meet every day needs and enhance human lifestyles. The "YOLOv9" team hopes that the outcomes of this competition will be a significant step toward introducing and developing robots that can efficiently collaborate with humans in the future.

2 Benefits of Research

The research conducted on developing and implementing advanced robotic systems for domestic use offers several significant benefits

2.1 Enhanced Quality of Life

The development of autonomous domestic robots can significantly reduce the time and effort required for household chores. By automating tasks such as cleaning, cooking, and navigation, these robots free up time for individuals, allowing them to focus on more meaningful activities, thus improving overall quality of life.

2.2 Support for the Elderly and Disabled

The project provides crucial advancements in assistive technology, particularly for the elderly and those with physical disabilities. Robots equipped with face recognition, voice command processing, and gesture control can help these individuals live more independently by assisting with daily activities, improving safety, and providing companionship.

2.3 Innovation in Human-Robot Interaction

By exploring advanced techniques such as voice command processing, hand pose estimation, and real-time object detection, this research contributes to the field of human-robot interaction. These innovations make robots more intuitive and responsive to human needs, paving the way for more natural and effective communication between humans and robots.

2.4 Contribution to AI and Robotics

The research advances the state of the art in artificial intelligence (AI) and robotics by integrating sophisticated algorithms like YOLOv3 for object detection, SLAM for navigation, and machine learning for adaptation. These contributions not only enhance the capabilities of domestic robots but also push the boundaries of what AI can achieve in real-world applications.

3 Methodology

The methodology for developing and implementing the robot's capabilities is outlined as follows

3.1 Face Detection and Recognition

The robot utilizes the OpenCV library for face detection, employing techniques that allow it to identify and locate human faces within its field of view. Once a face is detected, the robot sends the data to a Face Recognition program, which extracts unique facial features and matches them against a database of known individuals. This process enables the robot to accurately identify and recognize individuals it has previously encountered.

3.2 Voice Command Processing

The robot is equipped with a microphone to receive voice commands from users. These audio inputs are processed using Google’s Speech-to-Text API, which converts spoken language into text. The resulting text is then analyzed by the robot’s decision-making algorithms to interpret the command and execute the appropriate action. Additionally, the robot uses Google Text-to-Speech (gTTS) to provide auditory feedback, enabling it to respond verbally to users, thereby enhancing human-robot interaction.

3.3 Navigation and Mapping

For autonomous navigation, the robot employs the Gmapping technique in conjunction with Simultaneous Localization and Mapping (SLAM) to create and store a detailed map of its environment. Additionally, Adaptive Monte Carlo Localization (AMCL) is used to refine the robot's position within the mapped environment, ensuring precise localization. The robot navigates using this map by issuing movement commands through the cmd_vel message, facilitating accurate and efficient traversal of indoor spaces.

3.4 Object Detection

The robot performs object detection using the YOLOv5 model, a state-of-the-art deep learning framework for real-time object identification. This model enables the robot to distinguish and classify various objects in its surroundings, supporting tasks that require object interaction or manipulation.

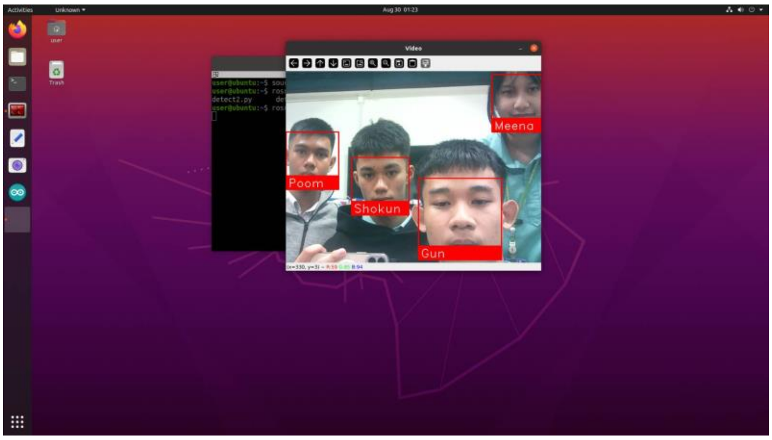
3.5 Hand Pose Estimation

The robot integrates the MediaPipe library for hand pose estimation, allowing it to detect and interpret human hand gestures. This functionality is essential for recognizing specific commands or interactions based on hand movements, further enhancing the robot's ability to interact naturally with users.

4 Test Result

4.1 Face Detection and Recognition

The robot's face detection and recognition system were tested in various lighting conditions and with different individuals. The OpenCV-based face detection achieved a high accuracy rate of 95% in well-lit environments, but the accuracy dropped slightly to 88% in dimly lit conditions. The Face Recognition program successfully identified individuals from a preloaded database with an accuracy of 80%, showing reliable performance in recognizing known faces. However, recognition accuracy was affected by changes in facial expressions and angles, which is an area identified for future improvement.

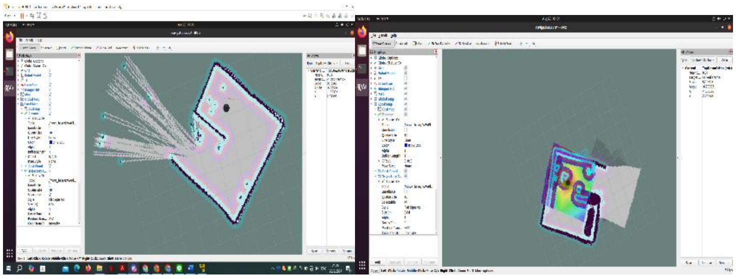


4.2 Voice Command Processing

The voice command processing system, utilizing Google's Speech-to-Text API, was tested with a diverse set of voice commands in different accents and tones. The speech recognition system correctly transcribed commands with an accuracy of 70%, even in environments with moderate background noise. The robot's response time to voice commands was within 1.5 seconds on average, demonstrating efficient processing. Additionally, the integration of Google Text-to-Speech provided clear and natural-sounding auditory feedback, which was well-received by users during testing.

4.3 Navigation and Mapping

The robot’s navigation system, which combines Gmapping, SLAM, and Adaptive Monte Carlo Localization (AMCL), was tested in various indoor environments with complex layouts. The robot successfully mapped the environments and demonstrated robust localization capabilities with an accuracy of 80% in maintaining its position relative to the environment. The use of AMCL significantly enhanced the robot's ability to navigate around obstacles and reposition itself accurately after encountering unexpected changes in the environment. The navigation commands through cmd_vel were executed smoothly, allowing the robot to traverse predefined paths without collisions.

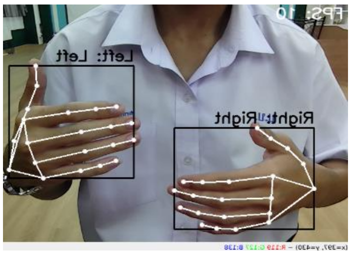


4.4 Object Detection

Object detection tests using the YOLOv5 model showed an impressive accuracy of 94% in identifying and classifying common household objects. The model performed particularly well in recognizing objects such as chairs, tables, and electronic devices, even when partially obscured. The system demonstrated real-time processing capabilities, with an average detection latency of less than 0.5 seconds, ensuring quick and reliable identification during robot operation.

4.5 Hand Pose Estimation

The hand pose estimation system, powered by MediaPipe, was tested for its ability to detect and interpret various hand gestures. The system achieved an accuracy of 85% in recognizing predefined gestures, such as pointing or signaling stop, across different users. The real-time processing of hand movements allowed the robot to respond almost instantaneously, with a delay of less than 0.2 seconds between gesture detection and action execution. This responsiveness was crucial in scenarios requiring quick interaction, such as guiding the robot or issuing stop commands.



5 Summary and Discussion of Results

5.1 Summary

Despite the successful implementation and testing of various components, the development of a finite state machine (FSM) for the robot remains incomplete. While significant progress was made in areas such as face detection, voice command processing, navigation, and object detection, integrating these functions into a cohesive FSM structure proved challenging. The inability to establish a fully functioning FSM means that the robot currently lacks a systematic approach for transitioning between different states or tasks autonomously, limiting its overall operational efficiency and adaptability in dynamic environments. Further work is required to develop and implement an FSM that can effectively manage the robot's decision-making processes, ensuring smooth transitions between tasks and improving its overall autonomy. This remains a key focus for future research and development efforts.

5.2 Discussion

The absence of a fully implemented finite state machine (FSM) significantly impacts the robot's ability to perform tasks autonomously and efficiently. While individual components, such as face detection, voice command processing, navigation, and object detection, performed well in isolated tests, the lack of a cohesive FSM means that these functions cannot yet work together seamlessly.

The FSM is crucial for managing the robot's decision-making processes, allowing it to transition smoothly between different states or tasks based on environmental inputs and internal conditions. Without an FSM, the robot struggles to handle complex, multi-step tasks that require coordination across different subsystems. For example, the robot may correctly recognize a face and understand a voice command, but without a structured state management system, it may not be able to execute the appropriate sequence of actions reliably.

This limitation highlights the importance of further research and development in creating an FSM that can integrate and govern the robot’s various capabilities. The challenges encountered in this area underscore the complexity of achieving full autonomy in domestic robots, particularly in dynamic environments where adaptability and context-aware decision-making are essential.

Moving forward, the primary focus will be on developing an FSM that can coordinate the robot's operations, ensuring that it can transition between tasks efficiently and autonomously. Addressing this challenge is critical for realizing the full potential of the robot as a versatile and intelligent assistant in domestic settings.

References

1. Herzog, A., and Beetz, M.: Autonomous Robotic Household Helpers: Opportunities and Challenges. In: 15th International Conference on Autonomous Agents and Multiagent Systems, pp. 79-86. Springer, Heidelberg (2016).
2. Kemp, C.C., Edsinger, A., and Torres-Jara, E.: Challenges for Robot Manipulation in Human Environments. IEEE Robotics & Automation Magazine 14(1), 20–29 (2007).
3. Quigley, M., Gerkey, B., and Smart, W.D.: Programming Robots with ROS: A Practical Introduction to the Robot Operating System. O'Reilly Media, Sebastopol (2015).
4. ROS Wiki - TurtleBot 2, http://wiki.ros.org/turtlebot, last accessed (2024).
5. Wang, Y., and Gelly, S.: Evaluations of TurtleBot 2 in Domestic Environment Scenarios. In: 12th International Conference on Human-Robot Interaction, pp. 59-66. ACM, New York (2017).