Enhancing Safety and Security in Construction Sites through Mobile Robot Monitoring

Group 2, Construction Robotics, WS 2023/24

Manav Thakkar (561745), Rohan Majigoudar (562193), Junaidali Surti (561774)

***Abstract*—Safety and security are significant concerns in industrial activities, particularly within the construction sector. However, existing practices for monitoring construction sites predominantly rely on manual observation, leading to labor-intensive and time-consuming processes. Inefficiency in monitoring procedures may compromise the overall safety and security of the construction site, highlighting the need for more streamlined and technologically advanced monitoring solutions. Recognizing the need for a more efficient approach, this paper explores the integration of small mobile robots as a flexible and cost-effective solution for monitoring construction sites. Mobile robots have gained prominence in the construction domain due to their stability on rough terrains, ability to navigate obstacles, climb stairs, and carry various sensors or arms for diverse tasks. The innovation not only saves time but also improves accessibility to challenging or unsafe spaces, ultimately reducing costs associated with construction tasks. This study proposes a comprehensive system that makes use of mobile robots and computer vision techniques to control the motion of the robot autonomously. Additionally, it features a web-based graphical user interface (GUI) enabling remote teleoperation of the robot, allowing users to obtain live status updates of the construction site through camera feedback. The proposed methodology demonstrates the potential of leveraging new technologies to enhance safety and security measures in the construction industry. The paper concludes with an analysis of the system, delving into the results and highlighting its viability as a practical solution for addressing safety and security concerns in construction sites.**

***Index Terms*—Landmark Detection, Teleoperation, Surveillance**

I. Introduction

In the rapidly evolving landscape of robotics and automation, the integration of advanced hardware with state-of-the-art algorithms has given rise to robotic systems, offering solutions across various industries. Notably, the construction sector, facing challenges of weak productivity growth and limited adoption of new technologies, is now set for advancement through the integration of robotics. The use of robotic technology in the construction industry is a promising trend. It has the potential to improve productivity, efficiency, and safety in various construction tasks.

Unlike sectors such as manufacturing, the construction industry lags in terms of efficiency and productivity. Recognizing the potential for enhancement, the focus has shifted towards leveraging robotic teleoperation. Teleoperation, a term deeply entrenched in established engineering, refers to the remote control of machines and systems. A key application within the construction industry involves the deployment of robotic teleoperation, with a focus on tasks such as payload handling, logistics, and the crucial function of monitoring construction sites.

In the context of telerobotics, the machine operates under the control of a human, with tasks such as data sensing, interpretation, and cognitive activities being orchestrated by the human operator. This paper delves into the exploration of the potential impact of robotic technology on safety, security, and productivity within construction sites. The primary focus is on autonomous operations, utilizing landmark detection to follow human operators, and concurrently examining teleoperation and monitoring functionalities. By delving into such aspects, the aim is to shed light on the potential of robotics in addressing concerns in the construction industry.

II. Background

The human-following aspect is achieved through the integration of the Mediapipe library, a tool for human skeletal posture detection. Leveraging deep learning and computer vision techniques, Mediapipe employs a combination of a detector and a tracker within its framework. The detector identifies the area of interest in the image, while the tracker focuses on recognizing posture landmarks. The Mediapipe library, developed by Google, utilizes Convolutional Neural Networks (CNNs) and deep learning to detect 33 distinct points on the human body, as depicted in Figure 1. Trained on extensive datasets with annotated examples of human body poses, the model generalizes well to diverse real-world scenarios. The neural networks enable accurate and robust landmark detection, providing detailed representations of body positioning in both image coordinates and 3-dimensional world coordinates.

Moreover, the ability of the robot to carry a payload introduces the potential for further enhancing its utility on construction sites. The comprehensive understanding of spatial aspects of human movement is made possible by the underlying architecture of the algorithm which establishes a solid foundation for the extraction of meaningful information about body posture. Additionally, it facilitates the extraction of information about movement, enabling diverse applications in construction site scenarios.

III. Research Methodology

The foundation of this study lies in the utilization of the Bittle Quadruped Robot (mIDOG) as the robotic platform,

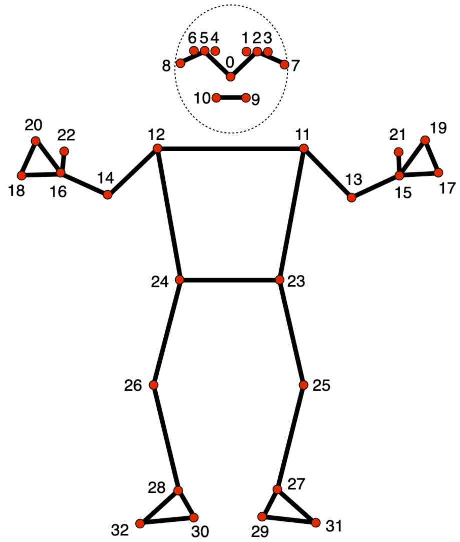

Fig. 1. Body Landmark Model

providing the groundwork for implementing a robust human-following mechanism. The mechanism leverages state-of-the-art landmark detection algorithms based on computer vision. The integration of Mediapipe body landmark detection model facilitates the human-following capability of the robot, with the hardware configuration of the robot illustrated in Figure 2. It presents the opportunity to deploy the robot as a mobile and hands-free assistant for a construction supervisor, ensuring efficient communication of instructions as the robot autonomously navigates the dynamic construction environment. The choice of a quadruped robot aligns with the objectives of the project, emphasizing the necessity for heightened mobility and adaptability in dynamic environments, particularly in construction sites. Inspired by nature, quadruped robots offer inherent stability and versatility, making them well-suited for navigating unstructured and challenging terrains.

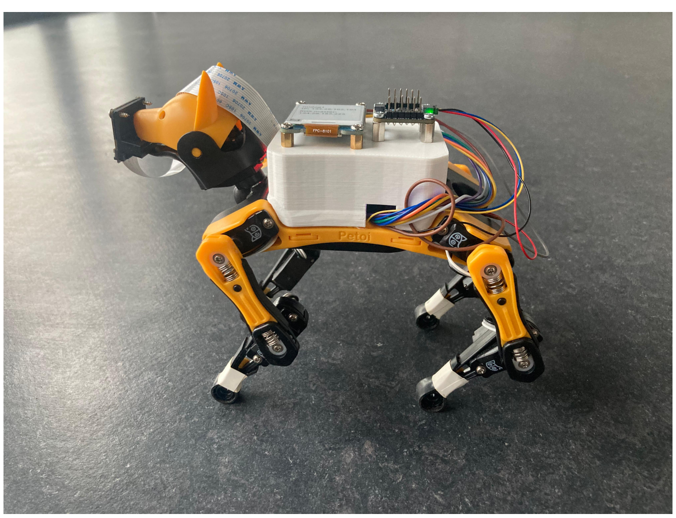

Fig. 2. Bittle Quadruped Robot (mIDOG)

The research employed in this study is designed to showcase the efficient integration of the Bittle Quadruped Robot (mIDOG) with the Robot Operating System Framework (ROS) for human-following and teleoperation functionalities. The utilization of ROS serves as the backbone for seamless communication among various components of the system. Due to the constraints of the robot's processing power, a decentralized communication model is adopted, with the ROS master running on a separate PC within the same network as the robot. The design ensures that resource-intensive computations such as landmark detection are offloaded to the PC, while the robot processor focuses on executing control actions, optimizing overall system efficiency.

Control decisions governing the movement of the robot are based on the tracking of a specific landmark on the human body. In the experiment, a point on the abdomen is selected for its consistent visibility in the robot's camera frame. The movement of the selected point dictates the corresponding movement of the robot. To achieve that, a virtual rectangular frame is superimposed on the video stream obtained from the camera of the robot. The dimensions of the frame determine the sensitivity of the movement of the robot. A smaller rectangle increases sensitivity, allowing the robot to respond more rapidly to the movement of the tracking point. The directional decisions are straightforward: turning left when the point moves left of the rectangle, turning right for rightward shifts, and applying similar logic for forward and backward directions, as shown in Figure 3.

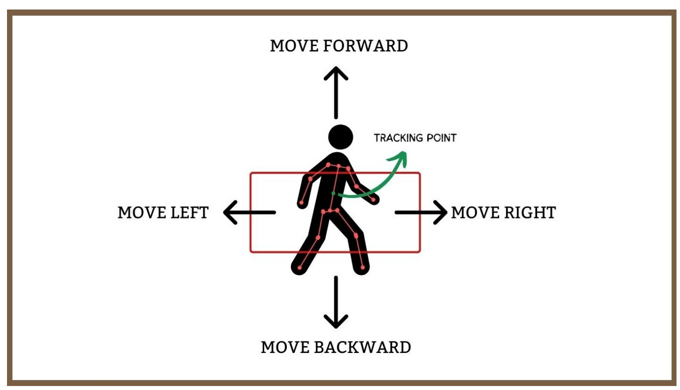

Fig. 3. Robot Movement Mechanism

The web-based GUI as shown in Figure 4, provides a user-friendly interface, enabling toggling between human-following mode and teleoperation mode using a switch. Additionally, the GUI offers a constant live video feed from the robot's perspective, enhancing situational awareness for the operator. The comprehensive methodology ensures an integrated and effective approach to human-following and teleoperation capabilities for construction site applications.

IV. Implementation

The implementation of the robotic system aims at human assistance, with a specific focus on potential applications within the construction industry, revolving around the utilization of the mIDOG platform as a prototype. The hardware platform serves as a stepping stone towards the development of a more robust robotic system tailored for the complexities of the construction environment.

Mediapipe is chosen for its lightweight architecture, real-time capabilities, and the precision with which it extracts relevant landmarks using a robust pre-trained model.

A key aspect of the implementation is the creation of a system capable of autonomously following a human user, coupled with a user-friendly interface for manual control. To enhance user accessibility and mitigate the potential complexity of the ROS system, a web-based GUI is developed. The GUI allows users to control the robot and access its camera feed through a web browser, featuring buttons for predefined robot motions, a joystick for teleoperation, and a toggle switch for switching between teleoperation and human-following modes.

The web-based GUI, hosted on a local network, serves as an intuitive control dashboard. While initially designed for local use, the potential exists to host it on a server, enabling remote monitoring and control from a location miles away from the construction site, contingent upon a fast and stable internet connection. Such implementation showcases the practicality and adaptability of the developed robotic system for human assistance in the construction industry.

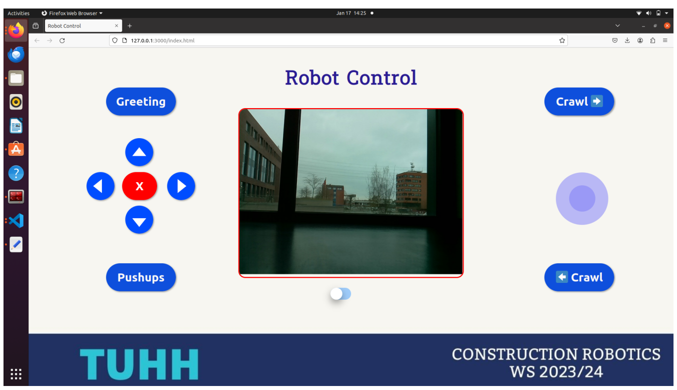

Fig. 4. Web-Based GUI

Several essential nodes, including the landmark detection node (for detecting body landmarks), the web video server node (displaying the camera feed to the GUI), the rosbridge websocket node (bridging ROS and the GUI), an optional teleoperation node (for keyboard control), and the robot control node (sending motion commands to the robot), contribute to the decentralized architecture and run on the PC. Conversely, nodes responsible for providing camera data and accepting motion commands run on the robot processor. The distributed setup ensures that the computational load on the robot remains minimal. Figure 5 provides an overview of the operational architecture and the data flow occurring within the system.

The hardware configuration involves the use of a Raspberry Pi camera mounted on the head of the robot, with each leg equipped with two servo motors providing two degrees of freedom per leg. The main processing unit, a Raspberry Pi 3A+, runs ROS and acts as a slave. Image data from the robot's camera is published on a topic and subscribed to by a node running on the PC. The image data, in the form of compressed images, is converted into OpenCV format for further processing. The resulting image stream, akin to a video, is fed into the landmark detection model implemented in Mediapipe.

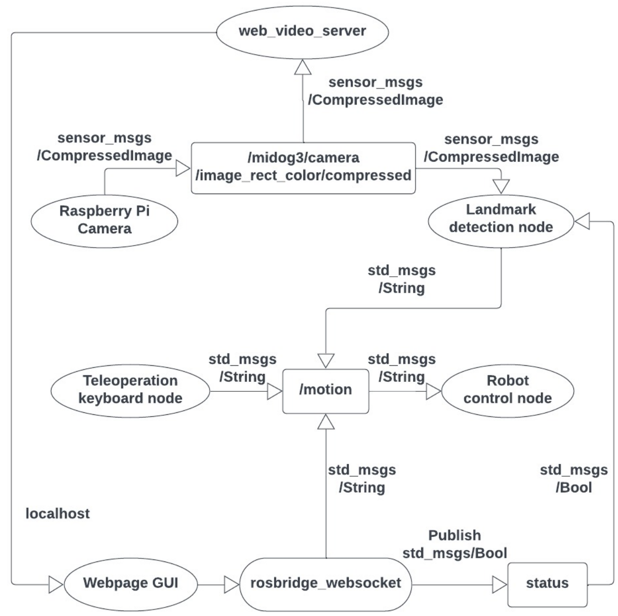

Fig. 5. RQT Graph of the system

The Mediapipe landmark detection model does not directly provide a landmark on the abdomen region. Instead, the locations of landmarks 23 and 24, representing left and right hips, are obtained. The mid-point between these landmarks is then calculated, with a vertical offset applied to align it correctly at the abdomen level.

The movement of the robot is dictated by the location of the tracking point. If the person moves out of the frame, the last command is executed until the tracking point reappears within the predefined rectangular region. For example, if the tracking point moves to the left and then disappears, the robot will continue moving left until the person reappears within the frame.

V. Validation and Results

The experiments aim to rigorously evaluate the performance and robustness of the developed robotic system under diverse conditions. Various tests have been conducted to assess the capabilities of the system and identify potential limitations. Tests are carried out in different conditions where a person stands a few meters away in front of the robot and moves on different surfaces and under various lighting conditions.

A. Surface Variation:

The system is evaluated on various surfaces, including a smooth table and a carpet. On a smooth table surface, the robot exhibits smoother movements, highlighting its adaptability. Challenges arise on carpets due to higher friction, resulting in difficulties dragging its legs, stability issues, and occasional falls.

B. Distance Variation:

The distance between the human subject and the robot was systematically varied to assess the system's performance. Optimal performance was observed when the distance between the camera and the person was within a specific range. Detection slightly deteriorated when the person was too close (less than 0.5 meters) or too far (more than 6 meters), highlighting distance-related dependencies in the system's operation.

C. Lighting Conditions:

The system's performance was evaluated under different lighting conditions to assess the robustness of the landmark detection algorithm provided by Mediapipe. The system exhibited consistent performance across varied lighting conditions, showcasing its adaptability to different environments.

D. Wi-Fi Transmission Latency:

During testing, issues related to Wi-Fi latency have been noticed, affecting the real-time performance of the system. To address the issue, tests have been conducted using a stationary PC camera, eliminating the need to transmit images over Wi-Fi. The modification has significantly improved the real-time performance and accuracy of the system, highlighting the importance of addressing connectivity issues for seamless operation.

The overall results of the project reflect a promising foundation for the developed robotic system, demonstrating its capability for monitoring and assistance in various conditions. However, noteworthy issues emerged during the experimental phase, particularly in relation to Wi-Fi latency, shedding light on critical considerations for the system's performance.

Issues related to delays in image transmission over Wi-Fi were noticeable and underscored the importance of network latency in the overall system functionality. The latency challenges manifested as delays in real-time responses, potentially affecting the robot's ability to follow a human subject seamlessly. The finding prompts a reevaluation of the communication infrastructure and potential optimizations to ensure seamless data transmission between the robot and the processing unit.

The results and insights gained from this experimentation phase contribute valuable information for future iterations, guiding the development of robotic systems with enhanced practicality and reliability in dynamic environments such as construction sites.

VI. Conclusion

In conclusion, the presented project introduces a robotic teleoperation and autonomous human-following system designed specifically for applications within the construction industry. The experimental results provide valuable insights into the potential of the system and underscore the challenges encountered in real-world conditions.

The experimental outcomes have highlighted the impact of network latency on the real-time performance of the robot. There is clear potential for improvement by upgrading the processing unit on the robot and optimizing the network architecture. Enhancements in such aspects could contribute to a substantial boost in real-time responsiveness and accuracy, crucial factors for the effective deployment of such robotic systems in dynamic environments like construction sites. The feasibility of deploying a more robust robotic platform, such as the Spot robot or similar models, using the same algorithms opens up possibilities for scalability and adaptability. The core algorithms, particularly ones related to human-following and teleoperation, remain consistent, offering a seamless transition to more advanced robotic systems. Overall, the project serves as a stepping stone towards the realization of sophisticated robotic assistance in the construction sector.

VII. Future Work

Future research should focus on refining adaptability, especially on diverse surfaces, and enhancing landmark detection accuracy at varying distances. Integration of advanced technologies like 5G can mitigate latency issues, while exploring sensors like depth sensors or LiDAR can improve environmental perception in challenging construction sites. Incorporating machine learning for dynamic adaptation and continuous real-world testing, guided by user feedback, are crucial for seamless integration into the construction industry.

The presented robotic teleoperation and autonomous human-following system lay the groundwork for potential future enhancements and expansions, particularly in the realm of safety and security on construction sites. While the current study focuses on teleoperation and human following, there are avenues for extending the capabilities of the system in future research endeavors.

One promising direction for future work involves leveraging the camera data to implement custom object detection algorithms. Specifically, the algorithms could be designed to assess whether workers at the construction site have properly donned their safety gear. Although not within the scope of the current study, the extension holds considerable significance for enhancing safety protocols within the construction industry. By identifying whether workers are adhering to safety regulations, the robotic system could serve as an additional layer of oversight, promoting a safer working environment. The potential extension aligns with the broader goals of advancing technology to address critical issues in construction site safety.

The web-based GUI, being a central interface for users, can be enriched with additional features for enhanced control and monitoring, including interactive map-based mission planning tools, real-time status updates, and the ability to define and customize patrol paths. Improvements to the GUI's user experience would contribute to the overall accessibility and effectiveness of the system.

Future research and development efforts can focus on integrating these enhancements, creating a more sophisticated and capable robotic system for addressing the diverse challenges in the construction industry.

Acknowledgment

The authors of this paper thank the IDAC at TUHH for providing the necessary resources and hardware used in the scope of this project work.

References

[1] Amrutha, K., P. Prabu, and Joy Paulose. "Human Body Pose Estimation and Applications." 2021 Innovations in Power and Advanced Computing Technologies (i-PACT). IEEE, 2021.

[2] Sánchez-Brizuela, Guillermo, et al. "Lightweight real-time hand segmentation leveraging MediaPipe landmark detection." Virtual Reality 27.4 (2023): 3125-3132.

[3] Singh, Amritanshu Kumar, Vedant Arvind Kumbhare, and K. Arthi. "Real-time human pose detection and recognition using mediapipe." International Conference on Soft Computing and Signal Processing. Singapore: Springer Nature Singapore, 2021.