2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

Design and Development of a Work Cell with a One-Handed Soldering Tool for Enhanced Human-Robot Collaboration

Natchanon Suppaadirek[1], Maximilien Sonnic[2], Raul Ariel Duran Jimenez[1], and Tomohiro Shibata[1]

Abstract— **The challenge of human-robot collaboration, particularly in the context of enhancing the productivity of work processes, has been a pivotal area of research and development for many years. Despite significant advancements, there remains a substantial gap in the design of these systems to cater specifically to individuals with disabilities. This paper presents an innovative approach in assistive robotics, focusing on the development of a work cell designed to facilitate individuals with single-arm functionality. Through the integration of depth camera technology, machine learning algorithms, and Mediapipe human tracking, our system is capable of interpreting human intentions, thereby making interactions with robots more intuitive and effective. Central to our research is the design of a specialized workspace that assists in object handling and incorporates a fully functional One-Handed Soldering Tool, integrated within a robotic arm setup. This work cell is tailored for users with limited arm functionality, demonstrating the system's versatility and providing invaluable insights into the practical implementation of applied robotics to bridge the theoretical and practical aspects of assistive technology.**

I. INTRODUCTION

In the current era of technological advancement, robots are increasingly commonplace in workplaces, leading to a noticeable decrease in employment opportunities for humans, particularly impacting older individuals and those with disabilities. This trend highlights a critical challenge: better integrating aging populations and people with disabilities into the workforce. The issue is compounded by the global rise in the aging population and the significant number of disabled individuals who are unable to find employment due to physical limitations and the growing preference for automated processes. One way to ease this integration is to offer them the possibility of getting a decent job. However, because of their disability, it can be difficult for them to perform well in that task compared to those without disabilities. As a result, it is crucial to find a way to mitigate this issue by continuously researching for improvement. One way to help this situation is by designing a system that could help them manipulate objects and supervise the work process which will support them in their work so that they can work effectively even with their limited body function. To tackle this issue, the concept of work cells is gaining traction. A work cell is a thoughtfully designed workspace that maximizes efficiency and accommodates the specific needs of all workers, including those with physical limitations. These spaces are equipped with technology, such as robotic assistants, to enhance productivity and support the workforce. There has been much research involving human-robot collaboration in recent years with different approaches and methods such as audio, verbal, sensor, and computer vision. Computer vision with the 3D camera is one of the most popular methods that has been used [1],[2]. However, there is very little research that aims toward older or disabled workers [1]-[5]. Linner et al. have introduced an assisted workstation, derived from the cell production system, aimed at supporting the elderly. They have evaluated this system using design engineering principles. While numerous studies have focused on cell production systems, no production system has been specifically proposed to assist individuals with disabilities [6].

We have been researching and improving the work cell that is more suitable for older or disabled people by introducing a work cell that uses a robotic arm as an assistant worker and creates an environment where tools are placed within a specific zone that can be easily grabbed by the robotic arm or the users and also easy to monitor the 3D printer that has been installed in this work cell. The gripper is also detachable. In this research, there are two types of grippers, one for object handling and the other for soldering assistance. The system will be controlled by computer vision using the camera to read user movement and command the robotic arm to provide assistance based on intention reading. Two types of computer vision have been used in this system : (1) human body pose detection and (2) human hand detection. The work cell is designed to support the process without completing all tasks autonomously, ensuring users remain actively involved. Also because the main focus is not only on older people but also disabled people, the system has been designed based on users that only have one functional arm which is different from the usual work cell. To make a workspace with the assistance of a robot, many factors have to be researched and experimented with to ensure that the robot will cooperate with humans effectively [7]. Safety is also one of the most important factors before making a work cell [8]. The researchers see the opportunity in this human-robot cooperation field as the robotic arm could be used to replace or assist the user's arm; a disabled person could significantly improve his productivity at work. For human motion detection which is a crucial part of the system, some research focuses on human detection. OpenCV and

*This work was not supported by any organization

[1]Natchanon Suppaadirek, Raul Ariel Duran Jimenez and Tomohiro Shibata are with the Department of Life Science and Systems Engineering, Kyushu Institute of Technology, Fukuoka, Japan. suppaadirek.natchanon733@mail.kyutech.jp; jimenez.raul-ariel839@mail.kyutech.jp; tom@brain.kyutech.ac.jp

[2]Maximilien Sonnic is with the Department of Human Intelligence Systems, Kyushu Institute of Technology, Fukuoka, Japan. sonnic.maximilien793@mail.kyutech.jp

mediapipe have been used quite often for human detection as they provide machine learning to estimate the human body [9]-[11]. OpenCV is a program used for computer vision and machine learning, while mediapipe is an open source for performing computer vision inference. So, in this research, computer vision can be built using Python code by combining these two programs to detect the human body and hand. For video capture, a realsense camera is one of the best choices as some researchers use this camera as video feedback data [12],[13]. The realsense camera has tracking depth technologies which are used for systems that require depth perception Also, as an assistive robotic, the system's function should be able to do more than just handle objects [14].

II. RELATED WORK

Much human-robot collaboration research that uses human movement or gestures to control robots has been researched and published. The system's central part is to make an algorithm that reads and collects human gestures and then uses that data to decode human movement in real time. In recent years, there has been much research about recognizing human actions [15]. EL Mehdi Saoudi combines 3D CNNs with the LSTM-Attention network to model subtle video patterns underlying accurate action classification [16]. While this approach has significantly improved over the existing systems, it comes with a lot of limitations when considering videos that have more than one action happening in the same frame. Traditional approaches to interpreting human actions using handcrafted methods focus attention on some specific features for analysis in video frames. These techniques are well studied and used in many instances as benchmarking methods against which newer techniques are evaluated. [17],[18]. This method has been widely used although the downside is that it is error prone and very time-consuming. Then there is machine learning. Machine learning-based approaches have become more popular for reading human actions due to machine learning's ability to learn from data independently and make precise predictions. Although these methods have demonstrated promising results, there remains potential for enhancement in robustness, real-time processing, and adaptability [19]. Next is deep learning. Deep learning methods have overwhelmingly outperformed the traditional approaches for human activity recognition. These advanced approaches employ neural networks to learn features directly from video frames so that actions may be comprehended more sophisticatedly and powerfully [20],[21]. The second part of the human-robot collaboration system is to utilize the human intention recognition of robots.

Yalin Cheng, et al. established a framework for human-robot interaction based on camera images and emphasized the significance of understanding human intentions and actions in these interactions which can be used as a case study for basing human-robot communication through human movement [22]. Robert Codd-Downey and Michael Jenkin explored underwater human-robot communication using divers' hand signals, applying the Faster-RCNN model to recognize divers' gestures in challenging underwater conditions although this system only read hand movement same as Wen Qi, et al. who introduced a robot controlled by hand gestures [23],[24].

Tianqi et al. concentrated on predicting human intentions by employing advanced neural networks to anticipate human actions, which is crucial for seamless human-robot collaboration. However, the results indicate that intentions do not always align with the corresponding action outcomes [25].

Intisar, et al. discussed the application of computer vision in robot control, showcasing a basic robot arm driven by a servo motor [26]. Jin, et al. investigated collaborative work between human teams and robots, employing cameras for object detection and controllers for assistance [27]. As described by the state-of-the-art there are human intention or movement prediction with multiple methods have been used to communicate with the robot.

In this paper, we propose a work cell that facilitates the use of a robotic arm as an assistance worker. The robotic arm is assigned a task by using a depth camera to read user intention with the use of machine learning. The system recognizes a specific array of gestures or actions that correspond to predetermined tasks, so the work process will be more efficacy.

III. SYSTEM ARCHITECTURE

This section introduces the overview system of our research to ensure that it can enable an effective workspace. All of the systems and subsystems have been explained and used as a part of the research.

A. Hardware and Workspace frame

The workspace was developed based on the general office workspace with a computer on the front. An extra frame, such as tool placement, has been added, allowing the robotic arm to assist. The concept of a workspace is shown in Fig. 1. In front of the user and on top of the workspace are depth camera Realsense D435 used for human intention reading and transferring the information to the Ur3e robotic arm. This robotic arm has a safety function that will stop the robotic if it receives external force such as hitting the wall or making contact with the user's body. The 3D printer is Creality CR-30. The computer's operation system is Ubuntu 20.04.6 LTS with Processor Intel Core i7-6700 CPU @ 3.40GHz x 8 and Graphic NVIDIA Corporation GM200 [GeForce GTX TITAN X.

B. Robotic arm station

The Ur3e robotic arm has been placed on the left side of the workspace, enabling it to access the tool placement area and transport tools to the workspace. Three different control modes have been implemented and tested: Manual, Automatic Intention Reading, and Hand Tracking Control. In the Automatic Intention Reading mode, a vision-based system is utilized. Linking camera information to the robotic arm required the development of human intention recognition using machine learning algorithms. For the Hand Tracking mode, algorithms for following hand movements have been developed to move the robotic arm using a camera placed on top of the workspace, alongside manual control facilitated by a control pad.

C. 3D printing station

3D printing can be used to produce custom parts based on 3D modeling. The raw material polylactide (PLA) and Acrylonitrile Butadiene Styrene (ABS) are used for printing. The 3D printer is implemented behind the robotic arm, where The robotic arm has access to grasp ready printed parts out of the 3D printer and carry them to the workspace.

D. Gripper units

The robotic arm's gripper was fabricated using 3D printing and designed for easy detachment and interchangeability with other types of grippers depending on the work situation. Two types of grippers were developed: one for object handling and another for assistive soldering. The use of 3D printing for the grippers means they can be easily replaced or repaired, offering flexibility in adapting new gripper models for various tasks. The grippers are depicted in Fig. 2 and 3. The model for the soldering gripper is made by reference from [28]. The model has been modified to be able to be installed on the robotic arm.

The previously described system outlines the architecture of an assistive workspace, including the functions and details of its various sub-systems. Three critical functional systems have been developed and integrated into a work frame serving as a working platform: workspace, robotic arm station, and gripper unit. These critical systems allowed the assistive human-robot collaboration scenario, while the 3D printing station is an example of the extra function that has been added based on the work situation. Also, there is a headset that provides the user with the ability to control the system to some extent with voice commands in case some users are not comfortable using the keyboard or a control pad to control the system since some users may still have a little bit control over their another arm while some do not have it at all. The chair also has a wheel for easy maneuvering.

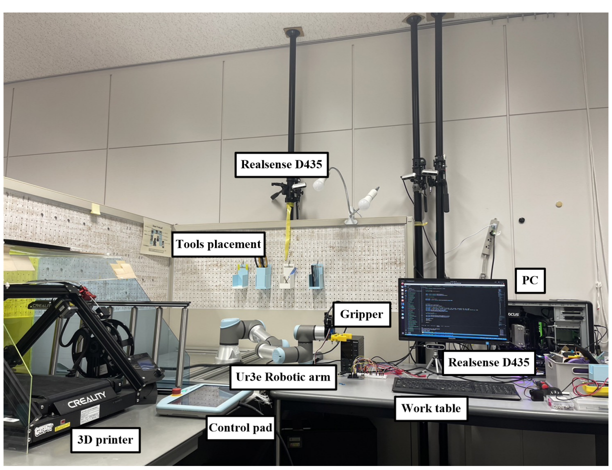

Fig. 1. Work cell

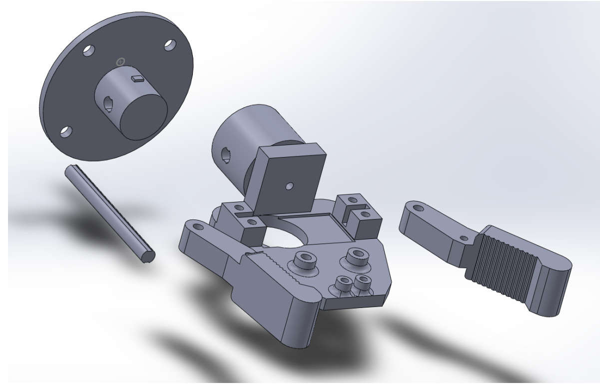

Fig. 2. Gripper model

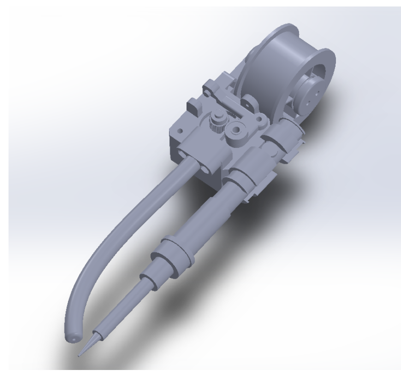

Fig. 3. Gripper model for soldering

IV. HUMAN INTENTION RECOGNITION

Human intention recognition is central to this research, where human body tracking and pose estimation are applied across various domains. These versatile techniques include gaze estimation, head angle positioning, face pose analysis, and hand signal recognition [15]. However, collecting data from a camera perspective and converting it into real-world coordinates is intricate, requiring complex mathematical formulations. Additionally, relying solely on image data for machine learning models may yield more than optimal accuracy, particularly when dealing with intricate scenarios. The RealSense D435 camera (Intel Corp.) is used for video capture in conjunction with Mediapipe Holistic. This combination simultaneously captures human facial expressions, body poses, and hand gestures. Utilizing Mediapipe Holistic to extract body key points alongside image frames enhances the analysis of human poses and actions. Extracted key points offer valuable insights into the spatial configuration of the body, facilitating an improved understanding of human movement dynamics.

Fig. 4 shows the workflow of the assistive robotic arm system. The initial step involves acquiring human gesture data through a depth camera. Following collection, this data undergoes processing and analysis to develop machine-learning models that can decode human gestures. These models assist the robotic arm in discerning human intentions.

In Fig. 5, Mediapipe Holistic outputs 543 landmarks in real-time, comprising 33 pose landmarks, 468 face landmarks, and 21 hand landmarks per hand. These landmarks and connecting lines delineate the body, face, and hand structures.

Given the focus on single-handed interaction in this research, hand tracking is only done with one hand. This system utilizes a set of 522 landmarks for its operations. Data from the depth camera, capturing key point information, is categorized into frames, classes, and coordinates in three dimensions: X, Y, and Z. The 'X' axis details the landmark's horizontal positioning from the camera's perspective, 'Y' outlines its vertical placement, and 'Z' conveys the landmark's depth, indicating its distance from the camera. This data helps construct an accurate three-dimensional model of human poses. The coordinates range from "X1, Y1, Z1" to "X522, Y522, Z522," where, for example, "X1, Y1, Z1" marks the position of the first landmark and so on, up to the 522^{nd} landmark. Additionally, a visibility metric, denoted as "V," is assigned to each landmark to denote its visibility in the camera's frame. A landmark might not be visible due to occlusion or the body, face, or hand's orientation relative to the camera. Visibility is crucial for understanding which body, face, or hand parts are detectable at any given time and can influence the accuracy of hand tracking and gesture recognition accuracy. The visibility values, from "V1" to "V522," identify the visibility status from the first to the 522^{nd} landmark, factoring in potential obstructions or the angle of the hand relative to the camera.

The subsequent phase involves training the collected data using supervised learning, as this method is effective in learning mappings between input features and predefined intention labels. In this research, researchers conduct a series of experiments wherein a set of gestures are performed and analyzed. There are 13 subjects with 26 specific tasks in the data-collecting process. At the end of data collection, there were a total of 80,523 image frames. The analysis of this data is conducted using five distinct machine learning algorithms: logistic regression, ridge classifier, random forest, gradient boosting, and support vector machines. Each algorithm is carefully fine-tuned to achieve optimal performance. After the fine-tuning process, the algorithms are combined into a more cohesive system through a VotingClassifier approach,

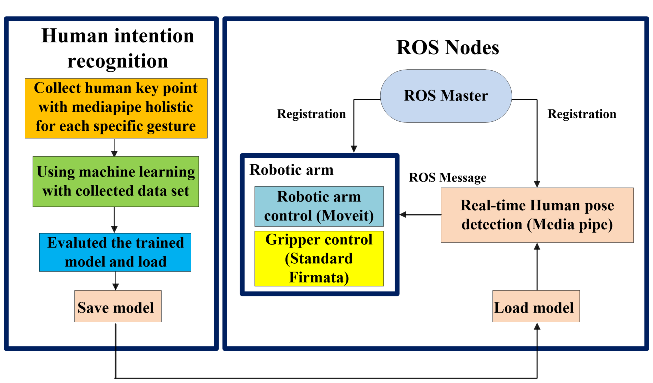

Fig. 4. Workflow of the human intention recognition system

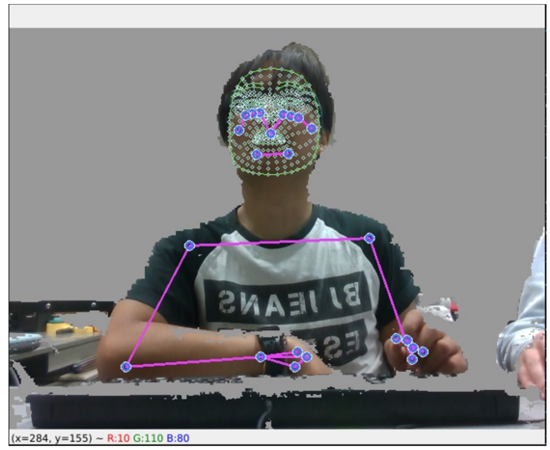

Fig. 5. Mediapipe holistic

significantly improving the system's reliability for gesture recognition. The performance of the integrated system is verified through cross-validation, with the final models being preserved for implementation in a human intention recognition system.

After training and testing, the system is implemented in a camera-equipped robotic arm. This setup enables the robotic arm to interpret human gestures, thereby enhancing its ability to anticipate the subsequent task, as depicted in Fig. 6.

For instance, if the user has a desire to write on paper or a book in Fig. 6, the robotic arm will respond by retrieving a pen from the equipment station and delivering it to the user, as demonstrated in Fig. 7.

V. HAND TRACKING SUBSYSTEM

Before developing the human intention recognition system, researchers have been studying hand-tracking robotic arm control systems [29]. This system has been included in the assistive system as, in some situations, controlling a robotic arm by hand can be helpful. Fig. 8 shows the hand-tracking control system. In this system, "Auto" and "Manual" are recognized as one of the distinct user intentions. The "Auto" mode enables the system to automatically

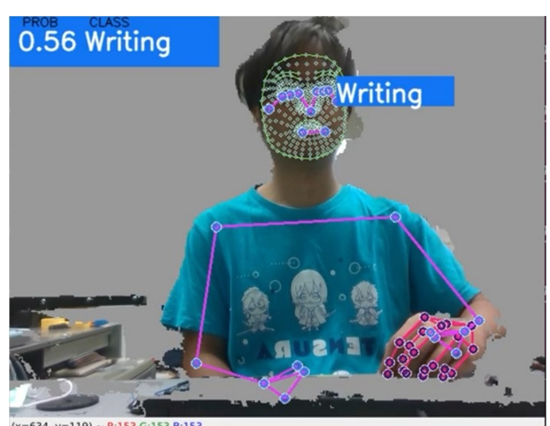

Fig. 6. Example of gesture reading

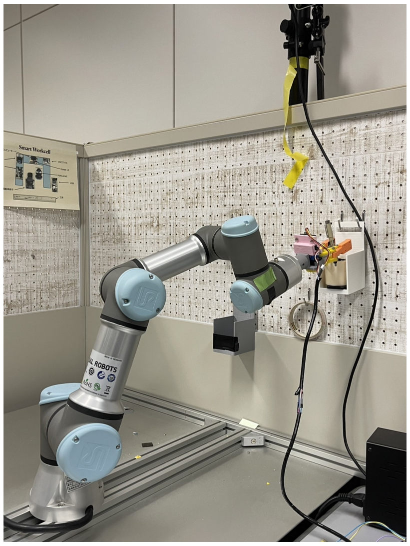

Fig. 7. Robotic arm follow gesture command

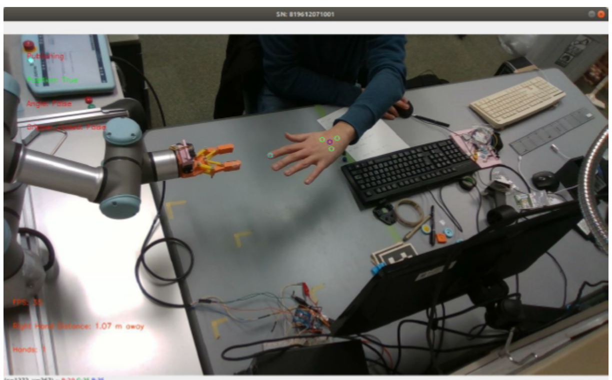

Fig. 8. Teleoperation control

understand and respond to human intentions using human intention recognition. On the other hand, the "Manual" mode allows users to directly control the robotic arm through hand movements, offering a more hands-on approach. This dual-mode capability ensures the system is versatile and can adapt to different user preferences and situations.

This system uses a Realsense camera placed on top of the workspace to track the user's hand with mediapipe hand tracking and create reference points on the center of the backhand, which controls the robotic arm's end effector. The program calculates the distance difference of the reference point when the user moves their hand and sends it to the robotic arm to move based on the value of that difference.

VI. EXPERIMENT

This research involves experiments with a controlled robotic arm that recognizes human intentions. Various functions of the system need testing and evaluation.

The effectiveness of the system in predicting human intentions was evaluated through experiments with healthy participants simulating single-arm functionality by hiding their left hand. The participants were asked to carry on different movements and postures to examine if the system's performance dependency varies with the distance and angle. Nevertheless, participants were healthy individuals, yet their design of the experiment had conditions that approximated, to a great extent, the experience of the target population. The use of healthy participants was justifiable based on the preliminary nature of the study and the need to ensure the basic functionality and safety of the system before trials with the actual target population. A total of five healthy individuals, all males within the age range of 22-29 years, participated in this study. All subjects were either lab members or acquaintances of the experimenter and were fully briefed on the purpose of the experiment, the safety measures taken, and the confidentiality of their data.

The following specific gestures were used while conducting the experiments, as shown in Fig. 9. Table I outlines the probability percentages for these gestures across different participants. Note that, in this experiment, only the human intention reading system is tested, not moving the robotic arm. This experiment only tested the human intention reading system without moving the robotic arm. It is important to note that two participants did not contribute to data collection. Except for Participant 5, who was involved in the programming process, no participants received a prior briefing. This selection aimed to maintain experiment accuracy and evaluate the system's usability by comparing results between informed and novice users. The system was set to capture and display all possible readings without filtration, such that it would even capture the least probable classifications if they had the highest probability. The subjects have also been requested to keep an eye on the screen as well as ensure that the intention written on the screen is accurate as the aim of this experiment is also to determine the optimal working position of each participant. The data summarized in Table I highlights the consistency in accuracy rates across different posture frames, emphasizing the system's adaptability and reliability in understanding human intentions in various scenarios.

Following the gesture recognition tests, the system's integration with a robotic arm was evaluated in practical settings. All the tasks performed in this phase are the same as the first experiment except this time the human intention reading system is connected to the robotic arm control program. This time the participants were asked to focus more on the task than monitor to simulate the real working situation. The system validated 20 consecutive image frames of user intentions, and consistent classification is to be made sure it was the correct intention before the command is passed on to the robotic arm control program. The tasks performed are the basic actions of opening and closing the gripper by showing five fingers and holding gesture, moving the robotic arm to the prescribed positions by showing one to four fingers and offering it upwards, downwards, to the left side, to the right side, forward and backward, and passing it by gestures of hands.

There are also complex tasks that were tested, such as requesting and returning equipment to the tool station. For

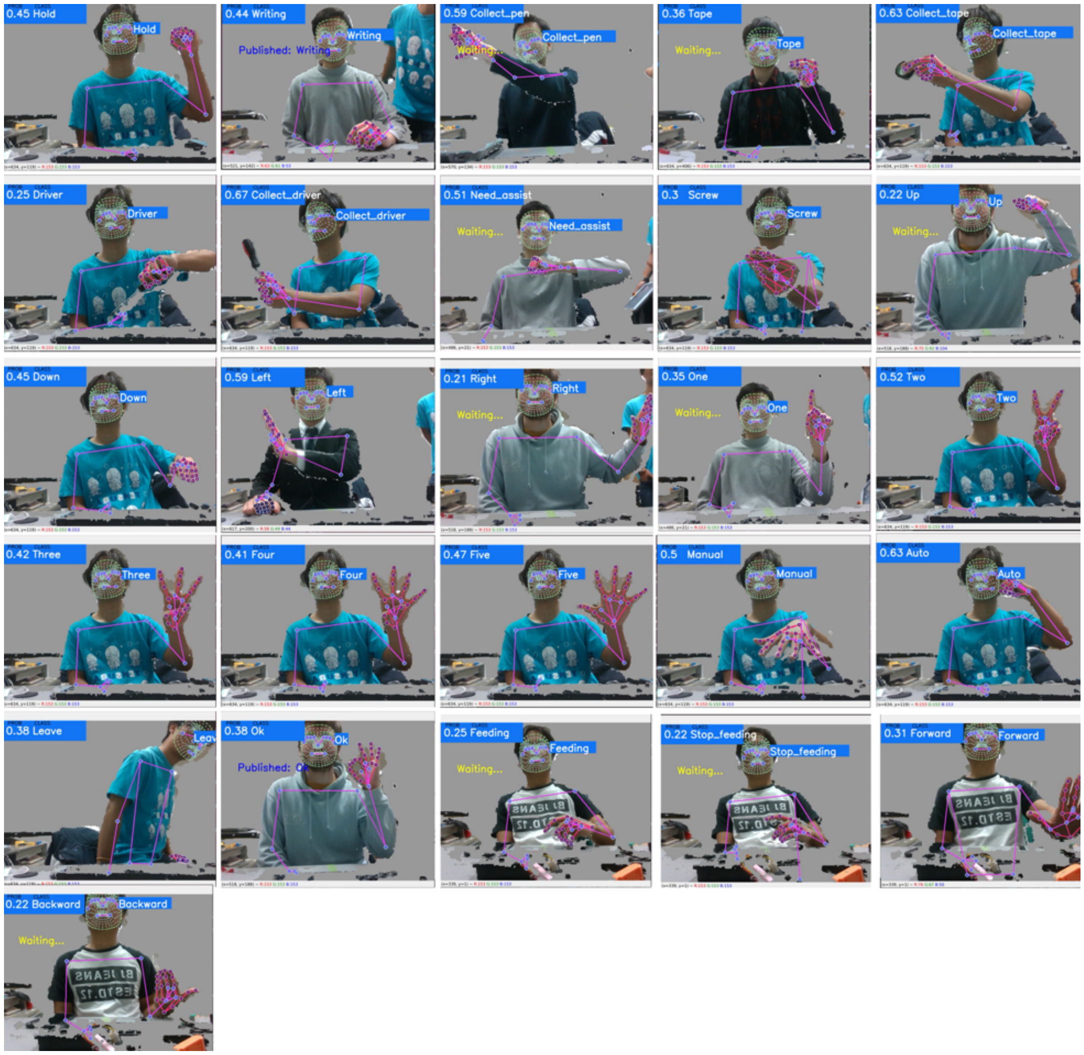

Fig. 9. All human gestures in experiment

example, the robotic arm transferred the box, containing bolts, to the user when the user used a screwdriver to assemble something that matched with the “Driver” command. ”Auto” and ”Manual” are for switching between when the robot recognizes human intention to control and when it recognizes hand tracking to control it, and ”Leave” is the mode that is initiated when the user is leaving the workspace to shut down the program. These procedures were carried out ten times per subject, and the results are shown in Table II.

The soldering machine experiment is also included in this section of the experiment by changing the gripper and studying two intended actions (”Feeding” and ”Stop Feeding”) to control the soldering material feeding. Two methods were evaluated: (1) moving the circuit board while the robotic arm remained stationary, and (2) keeping the circuit board stationary while the robotic arm moved to the desired position. Fig. 10 shows the robotic arm with a soldering assistant gripper.

Table II demonstrates the significant influence of the user’s position relative to the camera on detection accuracy. Notably, Participant 5, an author of this paper, completed all tasks due to a thorough understanding of the system’s

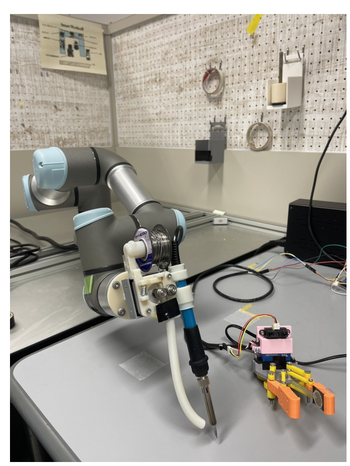

Fig. 10. Solder experiment

TABLE I
HUMAN INTENTION POSSIBILITY READING

Human intention	Test Subject				
	Subject1	Subject2	Subject3	Subject4	Subject5
Hold	0.23-0.49	0.26-0.72	0.26-0.45	0.18-0.54	0.19-0.46
Writing	0.15-0.77	0.33-0.48	0.18-0.35	0.18-0.77	0.19-0.79
Collect pen	0.23-0.64	0.31-0.61	0.29-0.39	0.31-0.61	0.35-0.58
Tape	0.36-0.52	0.29-0.47	0.21-0.36	0.17-0.41	0.25-0.45
Collect tape	0.23-0.46	0.25-0.46	0.46-0.52	0.36-0.55	0.31-0.45
Driver	0.20-0.46	0.37-0.56	0.22-0.38	0.32-0.46	0.28-0.68
Collect driver	0.18-0.55	0.38-0.42	0.34-0.45	0.26-0.45	0.26-0.66
Need assist	0.28-0.69	0.36-0.65	0.36-0.76	0.28-0.45	0.37-0.69
Screw	0.19-0.32	0.23-0.45	0.18-0.30	0.27-0.38	0.22-0.36
Up	0.31-0.56	0.55-0.66	0.36-0.53	0.36-0.56	0.31-0.53
Down	0.20-0.49	0.24-0.37	0.23-0.35	0.31-0.49	0.12-0.48
Left	0.19-0.60	0.31-0.51	0.15-0.33	0.40-0.66	0.31-0.50
Right	0.20-0.65	0.20-0.41	0.12-0.32	0.22-0.51	0.28-0.74
One	0.20-0.38	0.20-0.45	0.14-0.38	0.22-0.36	0.13-0.48
Two	0.20-0.36	0.25-0.41	0.15-0.36	0.22-0.38	0.30-0.52
Three	0.17-0.27	0.18-0.23	0.22-0.40	0.27-0.33	0.16-0.42
Four	0.21-0.30	0.34-0.47	0.28-0.33	0.28-0.35	0.15-0.46
Five	0.15-0.35	0.16-0.47	0.13-0.46	0.22-0.40	0.16-0.50
Manual	0.16-0.45	0.34-0.46	0.17-0.43	0.16-0.42	0.31-0.50
Auto	0.24-0.46	0.25-0.47	0.34-0.82	0.27-0.50	0.24-0.63
Leave	0.15-0.28	0.23-0.34	0.21-0.35	0.23-0.34	0.21-0.39
Ok	0.25-0.27	0.25-0.38	0.20-0.42	0.27-0.42	0.25-0.48
Feeding	0.15-0.34	0.11-0.32	0.17-0.32	0.23-0.42	0.16-0.32
Stop Feeding	0.17-0.36	0.14-0.35	0.12-0.31	0.16-0.29	0.14-0.31
Forward	0.18-0.34	0.18-0.34	0.15-0.35	0.13-0.34	0.15-0.32
Backward	0.18-0.36	0.15-0.32	0.14-0.28	0.14-0.35	0.16-0.34

TABLE II
ROBOTIC ARM TASK SUCCESS RATE

Human intention	Test Subject				
	Subject1	Subject2	Subject3	Subject4	Subject5
Hold	100	100	100	80	100
Writing	100	70	100	80	100
Collect pen	100	90	90	100	100
Tape	90	80	100	100	100
Collect tape	100	80	100	100	100
Driver	100	100	90	80	100
Collect driver	90	100	80	90	100
Need assist	100	100	100	100	100
Screw	80	90	70	100	100
Up	100	100	80	100	100
Down	100	80	80	80	100
Left	100	90	100	100	100
Right	100	90	90	100	100
One	100	90	70	100	100
Two	100	70	50	90	100
Three	100	50	50	80	100
Four	100	90	80	100	100
Five	100	80	100	95	100
Manual	100	100	90	90	100
Auto	90	100	100	100	100
Leave	100	100	100	100	100
Ok	100	100	100	100	100
Feeding	60	70	80	90	100
Stop Feeding	70	70	80	60	100
Forward	70	60	80	70	100
Backward	70	60	70	60	100

requirements and optimal gesture positioning. In contrast, the other participants, including Participant 1, with a notably petite stature and smaller hands, encountered challenges. These were very apparent in the case of distinguishing between similar gestures; thus, misreading occurred several times in the real-time testing. Also, during the soldering machine experiment, it was observed that the gesture recognition had less reading performance than the other because the user needed to move their hand quite often to simulate the real soldering process and solder lead occasionally missed the head of the soldering machine due to its soft and easily bendable nature. Additionally, the robotic arm's range of movement was limited in certain scenarios, which impacted the overall task performance.

VII. CONCLUSIONS

In this research, we developed a vision-based system designed to recognize human intentions and anticipate desired actions, particularly to assist individuals with limited use of one arm in performing work-related tasks. This system was successfully integrated with a robotic arm, demonstrating its potential to enhance workplace functionality for people with physical impairments. Our focus on work-related tasks, rather than everyday activities, highlights the system's application in environments where physical limitations could impede productivity. The integration of a One-Handed Soldering Tool with the robotic arm underscores the system's adaptability and practical applicability across different tasks. The experimental results indicate that users need initial guidance to effectively operate the system, including understanding the necessary movements, positioning, and the range of tasks the robotic arm can perform. Additionally, our findings reveal a gap in the system's performance concerning individuals with smaller body frames, indicating a need for further data collection and refinement to improve accuracy for this demographic. Finally, while the gripper model demonstrated utility, there is room for improvement to enhance its performance further. These insights will guide future developments aimed at making the system more inclusive and efficient for a broader range of users.

REFERENCES

[1] S. Hang, Qi. Wen, C. Jiahao, Y. Chenguang, S. Juan, and L. M. Amine, Recent advancements in multimodal human-robot interaction, Frontiers in Neurorobotics, vol. 17, May 2023.
[2] J. Arent, V. Abolins, J. Judvaitis, O. Vismanis, A. Oraby, and K. Ozols, Human-Robot Collaboration Trends and Safety Aspects: A Systematic

Review, Journal of Sensor and Actuator Networks, vol. 10, no.3, July 2021.
[3] E. Matheson, R. Minto, E. G. G. Zampieri, M. Faccio, and G. Rosati, Human-Robot Collaboration in Manufacturing Applications: A Review, Robotics, vol. 8, issue.4, pp. 100-124, December 2019.
[4] V. V. Unhelkar, P. A. Lasota, Q. Tyroller,R. Buhai, L. Marceau, B. Deml, and J. A. Shah, Human-Aware Robotic Assistant for Collaborative Assembly: Integrating Human Motion Prediction With Planning in Time, IEEE Robotics and Automation Letters, vol. 3, no.3, pp. 2394-2401, July 2018.
[5] A. Guo, B. Wang, and M. Zhao, Design and Application of Control System for Underground Inspection Robot, 7th Asia-Pacific Conference on Intelligent Robot Systems (ACIRS), pp. 18-21, August 2022.
[6] T. Linner, J. Guettler, C. Georgoulas, A. Zirk, E. Schulze, and T. Bock, Development and Evaluation of an Assistive Workstation for Cloud Manufacturing in an Ageing Society, 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), vol. 44, no. 4, pp. 550-551, March 2019.
[7] J. Ender, J. C. Wagner, G. Kunert, R. Larek, T. Pawletta, and F. B. Guo, ‘Design of an Assisting Workplace Cell for Human-Robot Collaboration, International Interdisciplinary PhD Workshop (IIPhDW), pp. 51-56, July 2019.
[8] S. Heydaryan, J. S. Bedolla, and G. Belingardi, Safety Design and Development of a Human-Robot Collaboration Assembly Process in the Automotive Industry, Applied Sciences, vol. 8, February 2018.
[9] C. Xu, Xinyi Yu, Z. Wang, and L. Ou, Multi-View Human Pose Estimation in Human-Robot Interaction, IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society, pp. 4769-4775, November 2020.
[10] P, Padhi and M. Das, Hand Gesture Recognition using DenseNet201-Mediapipe Hybrid Modelling, International Conference on Automation, Computing and Renewable Systems (ICACRS), pp. 995-999, 7 Feb. 2023.
[11] K. Roy and D. Chanda, A Robust Webcam-based Eye Gaze Estimation System for Human-Computer Interaction, International Conference on Innovations in Science, Engineering and Technology (ICISET), pp. 146-151, 23 May 2022.
[12] M. Rabbani, Md. J. Mia, T. Khan, and Md. I. I. Zarif, A Survey on RealSense: in context of Research and Application, 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), October 2020.
[13] A. L. Fielding, A. K. Pandey, Y. Jonmohamadi, R. Via, D. C. Weber, A. J. Lomax and G. Fattori, Preliminary Study of the Intel RealSense D435 Camera for Monitoring Respiratory Like Motion of an Irregular Surface, IEEE Sensors Journal, vol. 21, no. 13, pp. 14443-14453, 1 July 2021.
[14] R. C. Hsu, P. Su, J. Hsu, and C. Wang, Real-Time Interaction System of Human-Robot with Hand Gestures, IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE), pp. 396–398, 29 Dec. 2020.
[15] Y. Kong, and Y. Fu, Human Action Recognition and Prediction: A Survey, International Journal of Computer Vision, vol. 130, pp. 1366-1401, 28 March 2022.
[16] E. M. Saoudi, J. Jaafari, and S. J. Andaloussi, Advancing human action recognition: A hybrid approach using attention-based LSTM and 3D CNN, Scientific African, vol. 21, September 2023.
[17] F. Afza, M. A. Khan, M. Sharif, S. Kadry, G. Manogaran, T. Saba, I. Ashraf, and R. Damaševičius, A framework of human action recognition using length control features fusion and weighted entropy-variances based feature selection, Image and Vision Computing, vol. 106, February 2021.
[18] D. Avola, M. Bernardi, and G. L. Foresti, Fusing depth and colour information for human action recognition, Multimedia Tools and Applications, vol. 78, no. 3, March 2019.
[19] Z. Zhang, Z. Lv, C. Gan and Q. Zhu, Human action recognition using convolutional LSTM and fully-connected LSTM with different attentions, Neurocomputing, vol. 410, pp. 204-316, October 2020.
[20] H. Yang, C. Yuan, B. Li, Y. Du, J. Xing, W. Hu, and S. J. Maybank, Asymmetric 3D Convolutional Neural Networks for action recognition, Pattern Recognition, vol. 85, pp. 1-12, January 2019.
[21] M. A. Khan, M. Sharif, T. Akram, M. Raza, T. Saba, and A. Rehman, Hand-crafted and deep convolutional neural network features fusion and selection strategy: An application to intelligent human action recognition, Applied Soft Computing, vol. 87, February 2020.
[22] Y. Cheng, P. Yi, R. Liu, J. Dong, D. Zhou and Q. Zhang, Human-robot Interaction Method Combining Human Pose Estimation and Motion Intention Recognition, IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD), vol. 4, pp. 958–963, 28 May 2021.
[23] R. Codd-Downey and M. Jenkin, Human Robot Interaction Using Diver Hand Signals, 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), vol. 44, no. 4, pp. 550-551, 25 March 2019.
[24] W. Qi, S. E. Ovur, Z. Li, A. Marzullo and R. Song, Multi-Sensor Guided Hand Gesture Recognition for a Teleoperated Robot Using a Recurrent Neural Network, IEEE Robotics and Automation Letters, vol. 6, no.3, pp. 6039-6045, 16 June 2021.
[25] T. Ma, L. Zhang, X. Diao and O. Ma, Human Intention Prediction Using BIL-SCNN, Active Vision and perception in Human(-Robot) Collaboration (AVHRC 2020), 31 August 2020.
[26] M. Intisar, M. M. Khan, M. Islam and M. Masud, Computer Vision Based Robotic Arm Controlled Using Interactive GUI, Intelligent Automation and Soft Computing, vol. 27, no. 2, Jan. 2021.
[27] Z. Jin and P. R. Pagilla, Human-robot teaming with human intent prediction and shared control, Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications II, 21 April 2020.
[28] RoboticWorx, "Build a One-Handed Soldering Tool — Effortless Soldering," Hackster.io, Aug. 02, 2023. [Online]. Available: https://www.hackster.io/RoboticWorx/build-a-one-handed-soldering-tool-effortless-soldering-b69d2e#cad. [Accessed: Feb. 01, 2024].
[29] T. Batpurev, V. Kumar, M. Anraku, K. Yamasaki and T. Shibata, Real-time Robotic Arm Teleoperation through 3D Hand Pose Estimation, 9th International Symposium on Applied Engineering and Sciences(SAES2021), 2021.