Controlling a 4 Degree of Freedom (4 DoF) Robot Arm with Hand Gestures Using Computer Vision Technology for Manufacturing Processes

1st Royyan Naufal Fauzan
Department Faculty of Engineering Technology and Vocational Education
Universitas Pendidikan Indonesia
Bandung, Indonesia
royyannaufal@upi.edu

2nd Resa Pramudita
Department Faculty of Engineering Technology and Vocational Education
Universitas Pendidikan Indonesia
Bandung, Indonesia
resa.pd@upi.edu

*Corresponding author: resa.pd@upi.edu
Received: July 20, 2024; **Accepted:** November 27, 2024

Abstract— **The increasing complexity of tasks in industrial environments requiring higher efficiency and precision is a major issue for industrial workers. To address these challenges, technology is needed to speed up the production process and improve quality. As part of artificial intelligence systems in industrial environments, robot arms are a solution that can be considered. The development of a robot arm that can be controlled with hand gestures through computer vision technology becomes a feasible solution. 4 Degree of Freedom (DoF) robot arm is chosen as the experimental platform for this research. The research method used in this study is research and development. Utilizing computer vision algorithms to track and detect user hand movements, the user can interact with the robot arm without using additional control devices. Hand object segmentation, feature extraction, and mapping to robot coordinates are parts of the hand gesture recognition process. The results of this research are expected to help develop a responsive and user-friendly robot arm control system. The system was tested under various usage conditions and achieved an average accuracy rate of 93%. With these research findings, it is expected to help industries develop new optimization solutions.**

Keywords— robot arm control, hand gesture recognition, computer vision integration, human-robot interaction

I. INTRODUCTION

Technological development is inseparable from the advancement of robotics. Technological advancements can facilitate human tasks that require concentration and speed, such as the role of robots in assisting human tasks that can be widely used as a replacement for human labor in tasks requiring precision. Robots are increasingly important in various aspects of human life in the modern era full of technological advancements. Robots have transformed many industries. Robots function as personal assistants and automatic service providers in industries, increasing production efficiency in the medical field, and assisting in more precise surgeries. The potential of robots continues to grow despite already producing remarkable results. The future of robots is expected to ©offer more inventive solutions and positive impacts on human life due to ongoing research and development to enhance robot intelligence, adaptive capabilities, and social interaction[1]. Robots have a significant impact on various fields such as medical services and especially industry. One use of robots in industrial environments is the use of

robot arms. Over the past few decades, robot arms or often called robotic arms, have become an essential part of industrial robotics. Generally, a robotic arm is a mechanical device consisting of several arms and joints that can move to perform manipulation and work tasks in various environments. The structure of the robotic arm resembles a human arm and allows each joint to move with several degrees of freedom, providing the necessary flexibility to handle tasks. Robotic arms often have an end effector, which can be a gripper or other tool as needed. [1,2].

Robotic arms are used for various manufacturing processes in industries such as welding, drilling, and material cutting. Typically, robotic arms are controlled using manual control devices like buttons or levers. However, the implementation of robotic arms sometimes faces technical negligence, such as robotic palletizer negligence causing accidents. One intriguing area of research is the control of robotic arms based on 4 Degree of Freedom through human hand gestures. This method uses computer vision technology to identify and translate human hand movements into motion commands that can be interpreted by the robotic arm, resulting in a more natural and user-friendly interface for manufacturing optimization [3].

4 DoF robotic arm was chosen as the main subject of the study due to its flexibility in various applications such as manufacturing optimization. Using image sensors or cameras to capture hand gestures, detection, and understanding of human movements can be carried out with high accuracy. By implementing artificial intelligence in computer vision technology, it is hoped that the robotic arm can become more responsive and intuitive in terms of control and can be integrated more broadly in various application contexts [4].

II. METHODS

The research process began with the selection of the experimental platform. At this stage, a 4 DoF robotic arm was chosen as the research basis, and the image sensor or camera to be used to capture hand gestures was identified. To ensure both can function simultaneously, the sensor and robotic arm were then integrated.

A. Research Stage

To begin the research, the author, as the researcher, identifies the problem of a lack of technology development. The following is a diagram of the stages in this research.

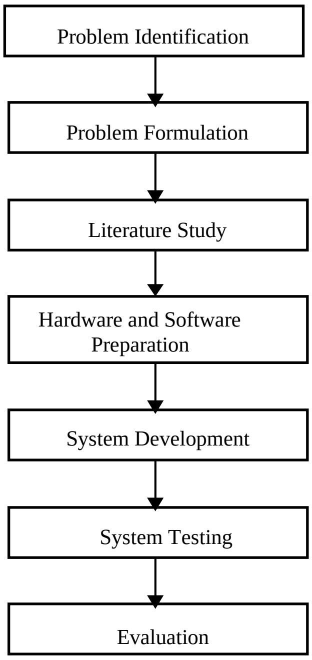

Fig. 1 Research Stage

Next, the third stage is conducting a literature review. This stage is carried out by the researcher with the aim of seeking and obtaining references from various sources regarding the system to be designed. In the fourth and fifth stages, the researcher begins to design the system using the Hand Landmark Model. In designing the system, the researcher uses Python machine learning and OpenCV to process the images.

B. Mediapipe

MediaPipe can recognize and understand hand movements with a very high level of accuracy through a gesture detection module that uses machine learning and computer vision technology. Using a strong pre-trained model, the MediaPipe gesture detection module can accurately identify the position and movement of human hands by mapping the landmarks of fingers and hands[5]. This includes various shapes and variations of hand gestures such as raising a hand, forming certain symbols, or gripping objects.

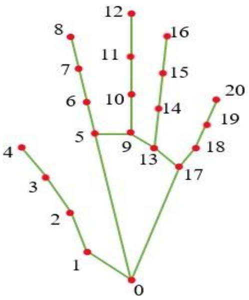

Fig. 2. Hand Landmark Model

Each point on the hand has its own landmark as mapped by MediaPipe:

TABLE I. LANDMARK METAPIPE

0. Wrist	11. Middle_Finger_Dip
1. Thumb_Cmc	12. Middle_Finger_Tip
2. Thumb_Mcp	13. Ring_Finger_Mcp
3. Thumb_Ip	14. Ring_Finger_Pip
4. Thumb_Tip	15. Ring_Finger_Dip
5. Index_Finger_Mcp	16. Ring_Finger_Tip
6. Index_Finger_Pip	17. Pinky_Mcp
7. Index_Finger_Dip	18. Pinky_Pip
8. Index_Finger_Tip	19. Pinky_Dip
9. Middle_Finger_Mcp	20. Pinky_Tip
10. Middle_Finger_Pip	

Looking at Landmark 0 and Landmark 9, both points can be used to determine the approximate orientation of the hand[6]:

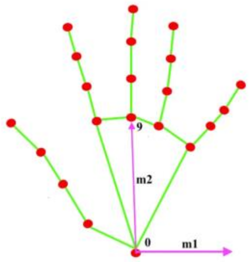

Fig. 3. Hand Landmark axis

The angle made by the line connecting Landmark 0 and Landmark 9 to the horizontal will be calculated as:

$$\tan \theta = (m2 - m1)/(1 - m1m2) \tag{1}$$

Where M1 is horizontal line, which is 0, and m2 is the slope of the line made by Landmark 0 and Landmark 9: [6,7] :

$$\tan \alpha = m2 = \text{abs}\,\frac{(y9 - y0)}{(x9 - x0)} \tag{2}$$

since m1 = 0, than:

$$\tan \theta = m2 \tag{3}$$

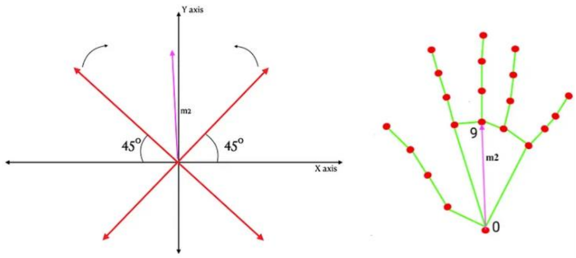

Fig. 4. Hand Landmark orientation : upward

Journal of Electrical, Electronic, Information, and Communication Technology (JEEICT)
Vol. 06 No. 2, October-2024, Pages 48-54 DOI: https://doi.org/10.20961/jeeict.6.2.90823

C. System Flowchart

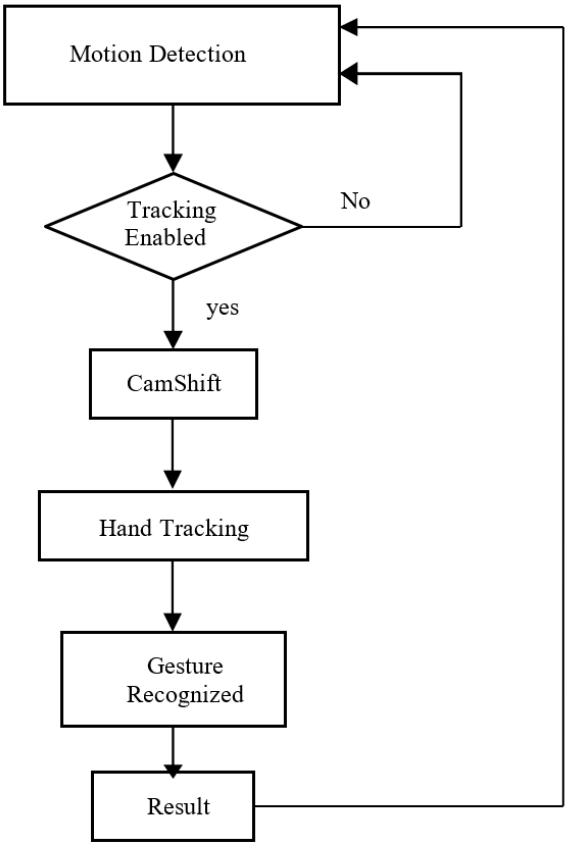

Fig. 5. System Flowchart

The system flowchart is a visualization of the system's operation, designed in accordance with the functioning of a robot arm control system operated using artificial intelligence, specifically computer vision.

D. Analysis precision, recall, f1-score

One of the crucial steps in the process of creating and developing a computer vision model is model evaluation. Model evaluation helps to understand how well the model can produce accurate and relevant predictions on previously unseen data. [8].

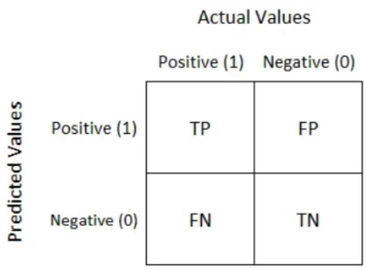

Fig 6. Confussion matriks table

- True Positive (TP):

TP (True Positive) is the number of positive cases that are correctly identified by the model. This is a situation where, in terms of classification, the model correctly recognizes a class as positive, and that class is indeed positive.

- False Negatif (FN):

FN (False Negative) is the number of positive instances that the model should have identified as positive but instead identified as negative. This occurs when the model misses or fails to detect positive instances.

- True Negatif (TN):

In classification, TN (True Negative) indicates the number of negative cases that are correctly identified by the model.

- False Positif (FP):

FP (False Positive) is the number of negative instances that the model should have identified as negative but instead identified as positive. This happens when the model mistakenly includes instances that should be negative in the positive class[9].

These elements are used to calculate various classification evaluation metrics such as precision, recall, and F1-score. They are also used to create a confusion matrix, which provides a comprehensive overview of how the model performs in classifying instances from different classes. [9,10].

1) Precision

Precision is a metric that measures how well the model makes correct predictions for the positive class out of the total positive predictions made [11]. Precision indicates how often the model correctly predicts the positive class from all positive predictions made by the model. Precision is a measure of how well the predictions made by the model are actually positive. Precision can be determined using the formula :

$$\text{precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{3}$$

2) Recall

Sensitivity or Recall is an evaluation metric that describes how well a model can correctly identify the positive class. Recall indicates the model's ability to find true positive values [12]. Recall also measures the extent to which the model can find all positive values that should be found. Recall can be determined by the formula:

$$\text{recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{4}$$

3) F1-Score

F1-score is the harmonic mean between precision and recall. It provides a balance between the two metrics and is useful when class distributions are imbalanced. A higher F1 score indicates better overall model quality. F1-score is an evaluation metric that shows the balance between Sensitivity (Recall) and Precision [13]. F1 Score indicates the model's ability to classify positive and negative values accurately. F1-score can be determined by the formula:

$$\text{recall} = \frac{\text{2.precision.recal}}{\text{precision+recal}} \tag{5}$$

E. Computer Vision

Branch of computer science known as computer vision focuses on developing technologies that enable machines to understand and interpret the world around them visually. Allowing machines to see and comprehend objects, patterns, and visual information like humans is the primary goal of computer vision. This includes the use of algorithms, machine learning models, and image processing techniques for complex visual analysis [14].

Computer vision has numerous applications, such as facial recognition, object detection, autonomous navigation in driverless vehicles, gesture recognition, and medical analysis. This technology plays a crucial role in the industrial revolution, artificial intelligence, and human-machine interactions, enabling innovations in many fields from healthcare to manufacturing and security.

Computer vision continues to evolve alongside rapid advancements in technology and computing capabilities. It has become a critical component in creating smarter systems and applications that are more responsive to the visual environment. [15].

F. Diagram System Overview

The picture below shows the scheme of using a program with its user. In this scheme, the user moves their fingers in front of a webcam, which detects patterns in the user's finger movements.

Fig. 7. Diagram System Overview

The finger patterns will be learned and recognized by OpenCV machine learning. After the finger patterns are recognized, they will be connected to a Python program on the computer. Python will be programmed to recognize user hand gestures, and then the program will be programmed to recognize user hand gestures, and then the program will send data to an Arduino to control a servo motor.

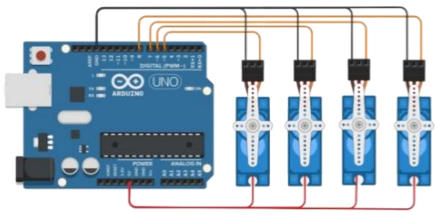

Fig. 8. wiring diagram servo

G. Robot arm 4 DoF

A 4 Degree of Freedom (4 DoF) robot arm is a type of manipulator with four independent degrees of freedom or movements. In robotics, degrees of freedom refer to how much movement or flexibility a robot has, and the number of DoF in a robot arm determines how many joints can be independently controlled [16].

Robot arm has four degrees of freedom and four joints that can be controlled separately. Each joint represents a specific axis of movement that allows the robot to move in three-dimensional space. Some common configurations of a 4 DoF robot arm include combinations of rotation and translation along specific axes [17].

Robot arm can be used for tasks such as handling, lifting, and placing objects. It has limitations in flexibility compared to robot arms with more DoF, but it is often used in various industries for tasks that require relatively straightforward motion control [18].

Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as “3.5-inch disk drive”.

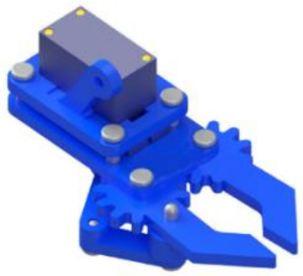

Fig. 9. Gripper design

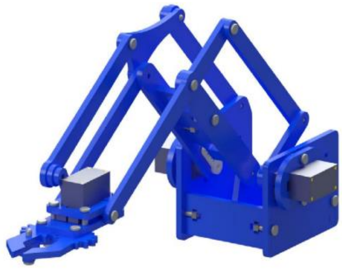

Fig. 10. Frame design

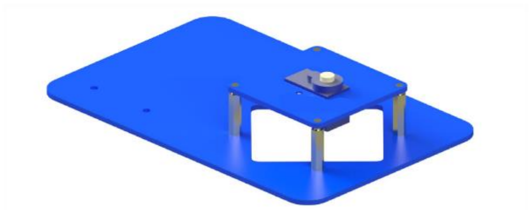

Fig. 11. Base desgin

The main material used to make this robot arm is acrylic. Each component is connected using several bolts, nuts, and spacers.

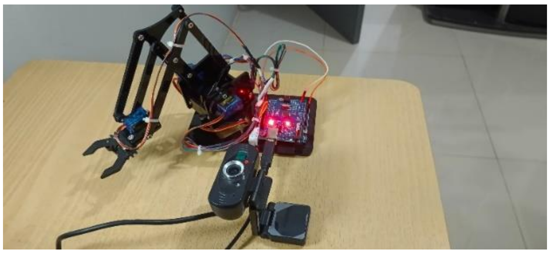

Fig. 12. Robot arm

III. Results and Discussion

A. *Robot Testing*

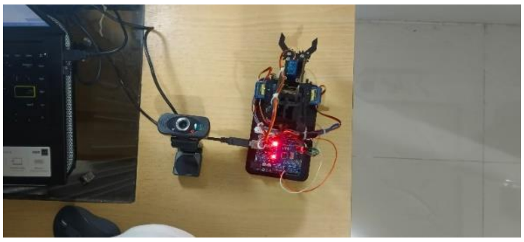

Fig. 13. Robot arm without detected gestures

In experiments without hand gestures, the robot arm will remain stationary at a predetermined position when no gestures are detected.

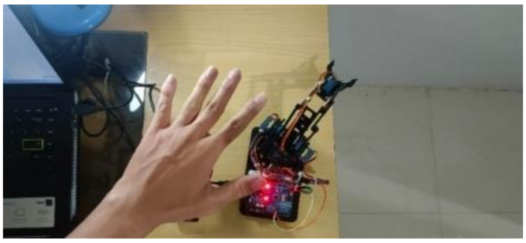

Fig. 14. Robot arm gestures towards the right.

Experiment: When the robot arm detects a gesture of the hand moving to the right (x+ direction), it will rotate accordingly. A hand gesture to the right will command the robot base to rotate to the right.

Fig. 15. Robot arm gestures towards the left.

When the robot arm detects a gesture of the hand moving to the left (x- direction), it will rotate accordingly. A hand gesture to the left will command the robot base to rotate to the left.

When the robot arm detects a gesture of the hand moving upwards (y+ direction), it will move common configurations of a 4 DoF robot arm include combinations of rotation and translation along specific axes [17].

Robot arm can be used for tasks such as handling, lifting, and placing objects. It has limitations in flexibility compared to robot arms with more DoF, but it is often used in various industries for tasks that require relatively straightforward motion control [18].

accordingly. A hand gesture upwards will command the robot arm to rotate upwards.

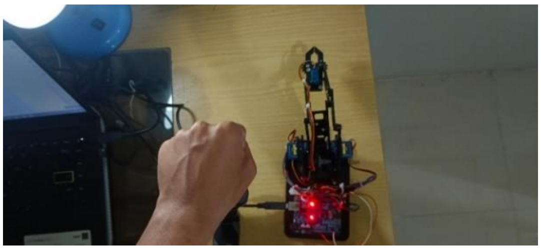

Fig. 17. The robot arm makes a fist gesture.

Experiment when the robot arm detects a fist gesture, it will move accordingly. Making a fist gesture will command the robot gripper to close.

B. *System Testing Result*

In the system testing results, several data points were collected to obtain accuracy values and assess the performance of the designed robot arm.

1) *Accuracy Analysis*

In testing the robot arm using hand gestures, the tests were conducted at distances ranging from 0.6 meters to 2.4 meters. The purpose of these system tests is to determine the accuracy level of the system.

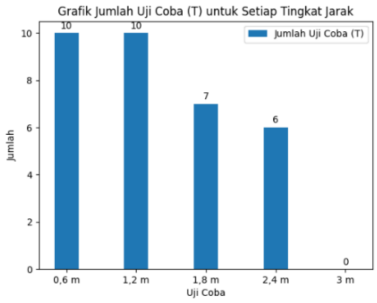

Fig. 18. Robot arm accuracy graph

T: True

F: False

TABLE II. System Testing

N	0,6 m	1,2 m	1,8 m	2,4 m	3 m
1	T	T	T	T	F
2	T	T	T	F	F
3	T	T	T	T	F
4	T	T	F	T	F
5	T	T	T	T	F
6	T	T	T	F	F
7	T	T	F	F	F
8	T	T	T	T	F
9	T	T	T	T	F
10	T	T	F	F	F
Testing (T)	10	10	7	6	0
Total System Accuracy Rate	####	####	70%	70%	0%
Total testing (T)	33				

Journal of Electrical, Electronic, Information, and Communication Technology (JEEICT)
Vol. 06 No. 2, October-2024, Pages 48-54 DOI: https://doi.org/10.20961/jeeict.6.2.90823

From the system testing results presented in Table 2 and the accuracy graph, it is evident that the system's accuracy heavily depends on the testing distance. The closer the testing distance, the higher the accuracy level achievable.

2) F1-score Analysis

TABLE III. ACCURATION ANALYSIS RESULT

No.	Precission	Recall	F1-score
1	0,9	0,9	0,9
2	1	0.91	0,95
3	0,9	0,9	0,9
4	1	1	1
5	0,9	0,9	0,9
accuracy			**0,93**
average	0,94	0,91	0,93

Table III shows a detailed report of the classification results of the implemented computer vision model. The average accuracy achieved by the model is 93%, as indicated in the analysis F1-score of 0.93 indicates excellent performance from the classification model. The F1 score is a metric that combines precision and recall, providing a comprehensive view of the model's ability to correctly identify and differentiate between negative and positive classes.

Specifically, an F1-score value between 0.5 and 1 indicates excellent performance, while an F1-score of 0 indicates poor performance [19]. An F1-score of 0.93 or 93% demonstrates that the model achieves high levels of precision and recall.

3) Comparative Analysis

Analysis comparing the results achieved in this study with previous research in the literature is presented in Table 4. Table 4 provides an analysis of the results of previous works in the context of robot arm control.

TABLE IV. COMPARATIVE ANALYSIS RESULT

Authors	Approach	Results of research	Accuracy
Raheja et al.[6]	Robotic control using hand gesture	The method used is object segmentation to extract hand gestures.	90%
Mardiyanto et al.[20]	Hand Gesture Recognition Sensor	The sensors used in this research are an accelerometer and a gyroscope. Both sensors are attached to the hand to control the robot.	98%
Karuppiah et al.[17]	Robotic Arm using Computer vision Interface	The method used is using Brain Computer Interface (BCI) to select objects from a scene image by creating a visual grid.	83.33%
Wang et al.[3]	Design of a voice control grasping robotic arm	The sensors used are sound and ultrasonic sensors, which provide data to the microcontroller that serves as the control system for the robot arm.	75%

IV. CONCLUSION

The results of this research overall demonstrate that the development of the 4 Degree of Freedom (4 DoF) Robot Arm control system using computer vision technology has been successful. The system can effectively respond to hand gestures through image processing methods and control logic, enabling intuitive and precise control. In evaluating the system's performance, an accuracy of 93% was achieved. This indicates the system's capability to recognize and interpret hand gestures effectively, facilitating seamless interaction between the Robot Arm and the user. However, while this accuracy is satisfactory, further improvements could be made by adjusting parameters and enhancing machine learning models to optimize the system's response to various gestures.

Overall, this research opens opportunities to utilize computer vision technology for enhancing control of 4 DoF Robot Arms through gesture interaction. Its findings provide a solid foundation for developing applications involving object manipulation and human-robot interaction.

ACKNOWLEDGMENTS

Author would like to extend their gratitude to everyone involved in this research for their effort, hard work, and contributions. Author also expresses their appreciation to all who supported this research, including supervisors, research colleagues, and everyone who provided valuable insights and input. Penulis is thankful to all who supported this research. Penulis hopes that the findings of this research will contribute significantly to the advancement of robotics technology and its applications in the future.

REFERENCES

[1] E. Demir, K. Yildiz, O. Demir, and E. E. Ulku, “Computer Vision Based Intelligent 3D CNC Machines,” Proc. - 2020 Innov. Intell. Syst. Appl. Conf. ASYU 2020, 2020, doi: 10.1109/ASYU50717.2020.9259826.

[2] H. Song, “The Application of Computer Vision in Responding to the Emergencies of Autonomous Driving,” Proc. - 2020 Int. Conf. Comput. Vision, Image Deep Learn. CVIDL 2020, no. Cvidl, pp. 1–5, 2020, doi: 10.1109/CVIDL51233.2020.00008.

[3] Z. Wang, D. Chen, and P. Xiao, “Design of a voice control 6DOF grasping robotic arm based on ultrasonic sensor, computer vision and alexa voice assistance,” Proc. - 10th Int. Conf. Inf. Technol. Med. Educ. ITME 2019, pp. 649–654, 2019, doi: 10.1109/ITME.2019.00150.

[4] Y. Li and Y. Zhang, “Application research of computer vision technology in automation,” Proc. - 2020 Int. Conf. Comput. Inf. Big

Data Appl. CIBDA 2020, pp. 374–377, 2020, doi: 10.1109/CIBDA50819.2020.00090.

[5] W. B. Dou, W. H. Chin, and N. Kubota, “Hand gesture communication using deep learning based on relevance theory,” 2020 Jt. 11th Int. Conf. Soft Comput. Intell. Syst. 21st Int. Symp. Adv. Intell. Syst. SCIS-ISIS 2020, pp. 6–10, 2020, doi: 10.1109/SCISISIS50064.2020.9322784.

[6] J. L. Raheja, R. Shyam, U. Kumar, and P. B. Prasad, “Real-time robotic hand control using hand gestures,” ICMLC 2010 - 2nd Int. Conf. Mach. Learn. Comput., pp. 12–16, 2010, doi: 10.1109/ICMLC.2010.12.

[7] H. Rahmani, S. M. Taheri, and E. Yargholi, “From Brain Decoding to Brain-Driven Computer Vision,” Iran. Conf. Mach. Vis. Image Process. MVIP, vol. 2020-Febru, 2020, doi: 10.1109/MVIP49855.2020.9116906.

[8] A. G. Spilling, A. R. Nix, M. A. Beach, and T. J. Harrold, “Electronics and Communication Engineering Journal,” pp. 1–3, 1999, [Online]. Available: citeseer.ist.psu.edu/hong99group.html.

[9] P. Sihombing, M. B. Rifky, H. Herriyance, and E. Elviwani, “Robotic Arm Controlling Based on Fingers and Hand Gesture,” Mecn. 2020 - Int. Conf. Mech. Electron. Comput. Ind. Technol., pp. 40–45, 2020, doi: 10.1109/MECnIT48290.2020.9166592.

[10] C. Orhei, M. Mocofan, S. Vert, and R. Vasiu, “End-to-End Computer Vision Framework,” 2020 14th Int. Symp. Electron. Telecommun. ISETC 2020 - Conf. Proc., 2020, doi: 10.1109/ISETC50328.2020.9301078.

[11] J. Dastur and A. Khawaja, “Robotic arm actuation with 7 DOF using Haar classifier gesture recognition,” 2010 2nd Int. Conf. Comput. Eng. Appl. ICCEA 2010, vol. 1, pp. 27–29, 2010, doi: 10.1109/ICCEA.2010.13.

[12] M. Wongphati, Y. Kanai, H. Osawa, and M. Imai, “Give me a hand - How users ask a robotic arm for help with gestures,” Proc. - 2012 IEEE Int. Conf. Cyber Technol. Autom. Control. Intell. Syst. CYBER 2012, pp. 64–68, 2012, doi: 10.1109/CYBER.2012.6392528.

[13] H. He and Y. Dan, “The research and design of Smart mobile robotic arm based on gesture controlled,” Int. Conf. Adv. Mechatron. Syst. ICAMechS, vol. 2020-Decem, pp. 308–312, 2020, doi: 10.1109/ICAMechS49982.2020.9310156.

[14] J. Sepúlveda and S. A. Velastín, “F1 score assesment of Gaussian mixture background subtraction algorithms using the MuHAVi dataset,” IET Semin. Dig., vol. 2015, no. 5, pp. 1–6, 2015, doi: 10.1049/ic.2015.0106.

[15] R. Mardiyanto et al., “Robotic Arm Controlling Based on Fingers and Hand Gesture,” Proc. - 10th Int. Conf. Inf. Technol. Med. Educ. ITME 2019, vol. 2020-Decem, no. 5, pp. 12–16, 2020, doi: 10.1109/SCISISIS50064.2020.9322784.

[16] X. Zhang and S. Xu, “Research on Image Processing Technology of Computer Vision Algorithm,” Proc. - 2020 Int. Conf. Comput. Vision, Image Deep Learn. CVIDL 2020, no. Cvidl, pp. 122–124, 2020, doi: 10.1109/CVIDL51233.2020.00030.

[17] P. Karuppiah, H. Metalia, and K. George, “Automation of a wheelchair mounted robotic arm using computer vision interface,” I2MTC 2018 - 2018 IEEE Int. Instrum. Meas. Technol. Conf. Discov. New Horizons Instrum. Meas. Proc., pp. 1–5, 2018, doi: 10.1109/I2MTC.2018.8409518.

[18] I. B. Çelik and M. Kuntalp, “Development of a robotic-arm controller by using hand gesture recognition,” INISTA 2012 - Int. Symp. Innov. Intell. Syst. Appl., 2012, doi: 10.1109/INISTA.2012.6246985.

[19] S. Gazzah and O. Bencharef, “A Survey on how computer vision can response to urgent need to contribute in COVID-19 pandemics,” 2020 Int. Conf. Intell. Syst. Comput. Vision, ISCV 2020, 2020, doi: 10.1109/ISCV49265.2020.9204043.

[20] R. Mardiyanto, M. F. R. Utomo, D. Purwanto, and H. Suryoatmojo, “Development of hand gesture recognition sensor based on accelerometer and gyroscope for controlling arm of underwater remotely operated robot,” 2017 Int. Semin. Intell. Technol. Its Appl. Strength. Link Between Univ. Res. Ind. to Support ASEAN Energy Sect. ISITIA 2017 - Proceeding, vol. 2017-Janua, pp. 329–333, 2017, doi: 10.1109/ISITIA.2017.8124104.