Chapter 3

Book Chapter

Gesture-Controlled Robotics: Enhancing Automation and Safety

Venkata Prajwal M *[1], Harish N T †[2], Saurabh Kumar ‡[3], D Cenitta §[4], Gururaj ¶[5], and Neelima B ‖[6]

[1]Student, Mtech in CSE , Manipal Institute Of Technology, Manipal Academy of Higher Education
[2]Student, Mtech in CSE , Manipal Institute Of Technology, Manipal Academy of Higher Education
[3]Student, Mtech in CSE , Manipal Institute Of Technology, Manipal Academy of Higher Education
[4]Professor, Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education
[5]Professor, Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education
[6]Associate professor, comute science and engineering, Manipal Institute of Technology, Manipal Academy of Higher Education

Abstract

Automation has played a critical part in revolutionizing sectors. The usage of gesture based robotic controls, which work without the need of a joystick or buttons, is the most recent breakthrough. The model proposed is trained to read and perform desired action based

*Email: prajwalmv2017@gmail.com Corresponding Author
†Email: harishnt58@gmail.com
‡Email: saurabhk1211@gmail.com
§Email: cenitta.d@manipal.edu
¶Email: gururaj.bijur@manipal.edu
‖Email: neelima.bayyapu@manipal.edu

on hand gestures using Convolutional Neural Network (CNN) technology. The proposed method focuses on controlling any existing RC Robot with the help of hand gestures without reconstructing a new model. This novel technique will allow the robot to maneuver with ease, including forward movement, reverse movement, turning left or right, and halting. Overall, this technology has the potential to improve the efficiency and safety of automated systems, opening the way for a more advanced and sophisticated robotics future.

Keywords: Convolutional Neural Network. Radio control. Receiver. Transmitter.

1 Introduction

For societal and economic advancement, digital capabilities and IT know-how must be integrated with public sector government services.(Mittal & Gautam, 2023). Further, our way of life and our jobs have changed because of the growth of robotics. As automation develops swiftly across areas including the military, healthcare, and industry, the use of robots has gained importance. This is possible due to the recent AI and machine learning technologies.(Gautam & Mittal, 2022). One of the most recent advancements in robotics is the invention of gesture-based remote control robot cars that can be driven via hand gestures or actual physical movements. These robot cars use sensors to keep track of the user's hand or body motions, which are then converted into orders for the robot vehicle to move. There are several applications for this technology, including those in instruction and learning, gaming, and industrial automation. The usage of hand gestures in the real world is contingent upon a number of distinct contextual factors, such as variations in background color, lighting, and hand gesture placement. Unfortunately, there is frequently not enough diversity in the datasets used to train hand gesture recognition algorithms, which makes it difficult to create precise and flexible systems.(Awaluddin, Chao, & Chiou, 2024).

The acoustic background information, content type, expression display strategy (acted vs. authentic), and other variables influence the capacity of deep learning models . (Anusha, Vasumathi, & Mittal, 2023). In this situation, gesture-based remote control robot automobiles are frequently con-trolled by the Convolutional Neural Network (CNN) algorithm. CNN is a deep learning method that excels in applications involving picture identification and classification. CNN can be used as complex patterns in an image that can be analyzed. These images that are given as input to the CNN model are gathered from the webcam. The algorithm in this paper is trained to recognize a wide range of hand gestures associated with different commands, such as moving the robot car forward, backward, left, or right, or stopping it. By being trained in a large collection of hand gesture photos, the algorithm can accurately recognize a variety of motions.

DOI:10.48001/978-81-966500-6-3-3 | ISBN: 978-81-966500-6-3

Moreover, existing models already enable the control of RC Robots using gesture inputs, allowing for the control of these robots with-out the use of an RC remote. By integrating image processing with the current model, it is possible to eliminate the physical interference caused by the remote control. The fact that existing gesture-based remote control robot vehicles rely on Bluetooth technology for connection with the user's smartphone, however, is a significant draw-back. The range of control is constrained by this reliance on Bluetooth, and signal interference is a possibility. Therefore, it is necessary to get rid of this reliance on Bluetooth and create a communication system that is more dependable and effective.

To replace Bluetooth for gesture-based remote control robot automobiles, this publication intends to investigate and recommend substitute communication technologies. The proposed model concentrates on the possibility of radio frequency and infrared technologies taking the place of Bluetooth in this situation. The following are the goals of the proposed model:

- Create a gesture-based remote control robot automobile that can be driven by hand movements that the CNN algorithm can recognize.
- Develop a wide range of hand movements linked with various orders, such as pushing the robot car forward, backward, left, or right, or stopping it, so that the CNN algorithm can properly recognize and understand them.
- To remove the physical interference generated by the remote control, integrate image processing into the present model.
- The RC Model's built with their respective RF Transmitter and Receiver will be directly used during the hand gesture control.

This will enable gesture-based control of robots in a variety of industries, including as industry, medicine, and the military, improving the effectiveness and security of automated systems. By giving us a simpler and more organic approach to operating robots, which will revolutionize the way they are engaged in our daily lives.

2 Literature Review

In the publication Sahoo et al.'s (2022) discusses the development of a real-time hand gesture recognition system using a pre-trained convolutional neural network (CNN) model adapted through fine-tuning and score-level fusion. This approach addresses the challenge of training deep CNNs with a limited number of static hand gesture images. The system's performance is assessed via leave-one-subject-out and standard cross-validation methods on two benchmark datasets, specifically focusing on recognizing American Sign Language gestures.

Zhu et al.'s (2023) presents a novel Hand Gesture Recognition (HGR) method using

DOI:10.48001/978-81-966500-6-3-3 | ISBN: 978-81-966500-6-3

Frequency Modulated Continuous Wave (FMCW) radars, aimed at overcoming challenges related to individual and environmental variations. The approach utilizes a deformable dual-stream fusion network that combines CNN and TCN architectures (DDF-CT). This method processes radar signals to create dynamic maps that effectively capture spatial and temporal gesture features. Incorporating deformable convolutions and inter-frame attention, the system achieves high recognition accuracies of 98.61% and 97.22% in standard and new environments respectively, significantly surpassing existing HGR methods.

Further, the paper Pinto et al.'s (2019) introduces a gesture recognition approach utilizing convolutional neural networks (CNNs). The method incorporates a series of preprocessing steps including morphological filtering, contour generation, polygonal approximation, and segmentation to enhance feature extraction. The system is trained and evaluated using various CNN models and compared against established architectures and methods. The effectiveness of the proposed method is validated through a detailed analysis of performance metrics and convergence graphs during training, confirming its robustness. CNN is used to merge RGB and depth pictures in (Kim, Jeong, & Jung, 2017) to create a reliable gesture detection system. Using a pre-trained CNN, the system extracts feature from RGB and depth pictures independently before fusing the features to enhance recognition performance. The suggested system was tested on a data set of hand gesture photographs, and its performance was compared to that of other systems already in use. The results of the experiments demonstrated that the suggested technique performed better than the other methods and had a high level of recognition accuracy.

A CNN-based hand gesture detection system that incorporates RGB and depth data. The authors first extracted features from the RGB and depth pictures using two different CNNs, and then they integrated the in- formation using a fusion network. The experimental findings demonstrated that the suggested system outperformed previous approaches and achieved excellent recognition accuracy.Adding an attention-based module to the VGG16 architecture is the main way to enable the network to possibly learn differentiating aspects of images.According to experimental findings, characters with hand postures comparable to "m" and "n" have an average recognition accuracy that is higher than that of state-of-the-art networks. Furthermore, for all gesture classes, the suggested approach outperforms the state-of-the-art networks in terms of recognition accuracy, with an approximate 3% advantage. The precision, recall, and F-score metrics have also been used to verify the effectiveness of the suggested architecture.(Barbhuiya, Karsh, & Jain, 2022).

The concept and development of a surveillance robot that can be operated remotely and at a great distance are presented in the study by Hasan et al.'s (2018). An Arduino microcontroller board, a Wi-Fi module, and a Raspberry Pi computer are used to construct

DOI:10.48001/978-81-966500-6-3-3 | ISBN: 978-81-966500-6-3

the robot. While the Raspberry Pi provides the computing power for image processing and object detection, the Wi- Fi module enables wireless communication between the robot and the control station. The control station receives a live video stream from the robot's camera, which is attached to the machine. The control station is a computer that runs a special piece of software that enables the user to direct the robot's movements and see the live video stream. In this proposed methodology can be implemented with any RC Robotic model without separately building it from scratch.

Numerous studies have investigated the integration of various devices and communication modules, highlighting diverse applications and technological advancements. ARVINDH et al.'s (2021) and Hasan et al.'s (2018) explored the combination of Raspberry Pi with Bluetooth modules, demonstrating its potential in IoT projects. Similarly, Jadhav et al.'s (2018) and Shah et al.'s (2020)focused on Arduino devices paired with Bluetooth modules, showcasing their effectiveness in wireless communication. Waskito, Sumaryo, and Setianingsih's (2020) extended this research by examining Arduino devices with ZigBee modules, emphasizing their suitability for low-power, long-range communication.

Gomathy, Niteesh, and Krishna's (2021) and Zagade et al.'s (2018) delved into microcontrollers integrated with Bluetooth modules, presenting innovative solutions for embedded systems. Hemane et al.'s (2022) and Wu, Su, and Wang's (2010) furthered the exploration by combining Raspberry Pi with ZigBee modules, highlighting enhanced connectivity and network efficiency. Deep Shakya et al.'s (2020) and Shukla et al.'s (2019)investigated microcontrollers utilizing Wi-Fi modules, underscoring their importance in wireless networking. Akyol and Canzler's (2000) and Awasthi et al.'s (2023) examined Android devices with Wi-Fi modules, illustrating their versatility and widespread applicability in modern communication systems. The communication method utilized in the studies was restricted to either Bluetooth or Wi-Fi Modules, and the same model could not be applied to any other robot for the intended purpose. In contrast, this paper explores the use of an existing RC robot with a remote control and investigates the possibility of controlling it via hand gestures, which eliminates the need to construct a separate robot for each unique project.

3 Dataset

The dataset acquired for classification of hand gestures using CNN algorithm was obtained with the help of webcam. The data set includes several images of two labels namely forward and backward. This dataset is created by capturing the images of hand gestures using computers webcam. After capturing the images, these can be preprocessed by using image processing techniques and made ready for training the CNN algorithm. The dataset contains images of only two labels because the focus is on how a RC Robot can be controlled

DOI:10.48001/978-81-966500-6-3-3 | ISBN: 978-81-966500-6-3

via hand gestures instead of fixing the RC Robot model.

4 Methodology Used

The created dataset with the help of webcam contains several images of two labels. The images in the dataset are Preprocessed first. Preprocessed images are further converted into binary format which is given as input to the designed CNN model for classification. The parameters needed to control the RC Robot will be decided based on the Classification output. Once the parameter is decided it is passed on to the Arduino which in turn will control RC Robot. The following figure (see figure 1) describes the workflow of the system.

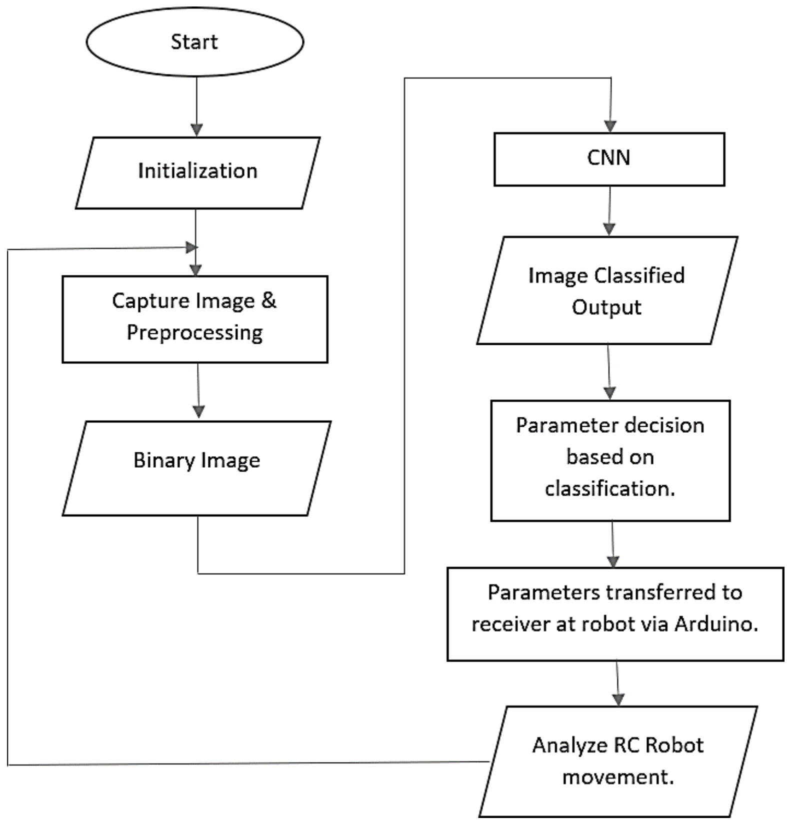

Figure 1. Workflow Overview

DOI:10.48001/978-81-966500-6-3-3 | ISBN: 978-81-966500-6-3

4.1 Data Preprocessing

The preparation of the dataset comes before any analysis. Using MATLAB's Image Acquisition Tool, the image is processed. The picture noise is eliminated using the median filter approach. The main idea behind the median filter is to repeatedly replace each element in the signal with the median of the entries that are closest to it. The neighbor-hood pattern, which moves entries at a time throughout the whole signal, is referred to as the "window". While the window for two-dimensional (or higher-dimensional) data must comprise all entries within a certain radius or elliptical region, the window for one-dimensional signals might be as basic as the first few preceding and following entries. The result of preprocessed image is a threshold image (black and white) which is converted into binary image will be processed again on CNN for Feature extraction and classified. Pre- processing of the image is explained in the following diagram.(see figure 2).

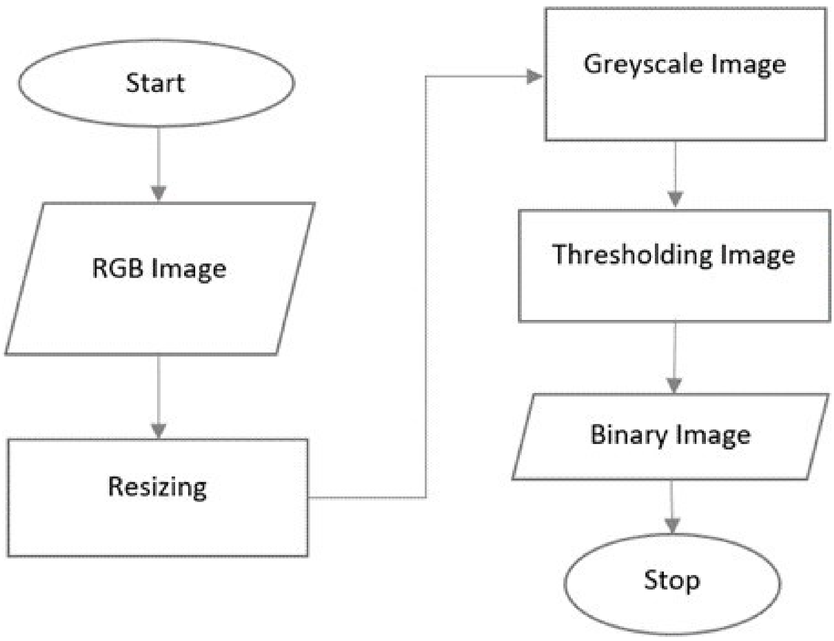

Figure 2. Image Pre-processing Flowchart

DOI:10.48001/978-81-966500-6-3-3 | ISBN: 978-81-966500-6-3

4.2 Convolutional Neural Network

This proposed model uses CNN algorithm for classifying hand gestures. Six 2D convolutional layers are used to make the CNN network and maximum pooling layer algorithm is used for each output of the convolutional layer. The maximum pooling algorithm with stride jump two, gives the maximum value in the pooling layer filter and data points in the pooling layer. ReLu activation function is used which takes the output of convolutional layer as input to activation function and output of activation function is fed into the pooling layer. A completely connected network with two layers receives the output of the sixth convolution layer as the input. Final output layer with two neurons is designed - one for each of the two hand gestures. Each layer has 512 concealed neurons. In the output layer, a sigmoid activation function is applied.

The model built in this paper takes the hyperparameters as convolutional filter with size 5x5, ReLu as the activation function, number of epochs is 30 and maximum pooling filter with stride jump as two. After classifying the image from the CNN algorithm, the result is decided as 1 or 2. This result is transferred to receiver at robot via Arduino. Based on this result, robots are moved forward and backward.

4.3 Hardware Implementation

Typically, a robotic automobile is an electromechanical device that can carry out activities autonomously. A computer interface or a remote control may be used to provide some degree of guiding for some robotic automobiles. Autonomous, semi-autonomous, or remotely operated robot cars are all possible. Robotic vehicles have advanced so far and are now capable of emulating humans that appear to have independent minds. A remote control that uses radio waves to operate a remote-controlled robotic car, also known as an RC car, is self-powered and enables remote control operation of the vehicle from a distance. The transmitter emits a specified number of electrical pulses proportional to our action. The transmitter has a separate power supply, which is often a 9-volt battery. The transmitter won't be able to broadcast radio waves to the receiver without a battery. All operating components, including the motor, get electricity from the power source. The transmitter in the proposed model uses radio waves to provide control while the receiver turns on the motors based on the signals received from the transmitter.

The RC Robot's remote contains a transmitter which transmits certain signals based on the button pressed. Signals which are recognized by the receiver decode the intended action that must be performed. These inferred actions are sent to the motors by the circuit. The quantity of electrical pulses (signals) is translated into action by the circuit board.

DOI:10.48001/978-81-966500-6-3-3 | ISBN: 978-81-966500-6-3

Real-time gestures are categorized into one of two categories (1 or 2) when they are made using the training model. The parameter that is supplied as input to the Arduino is calculated based on this categorization. Additionally, Arduino drives robotic vehicles based on input parameters. This model could achieve control in any direction—forward, backward, left, or right.

5 Results

The intermediate epoch results during the training phase demonstrate to classify 2 gestures it took a total of 7 minutes and 35 seconds.(see figure 3). The training model has been completed and on testing the model with the training data set we achieved 100 percent accuracy. Upon implementing this trained model on real-time gestures, we could classify the image correctly.(see figure 4)

Command Window

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:12	46.00%	32.6954	1.0000e-04
9	50	00:02:18	95.00%	6.6911e+11	2.0000e-05
17	100	00:04:14	100.00%	0.0000e+00	8.0000e-07
25	150	00:06:13	99.00%	4.4898e+11	1.6000e-07
30	180	00:07:35	99.00%	4.4199e+11	3.2000e-08

Training finished: Max epochs completed.

Figure 3. Intermediate Epoch Results

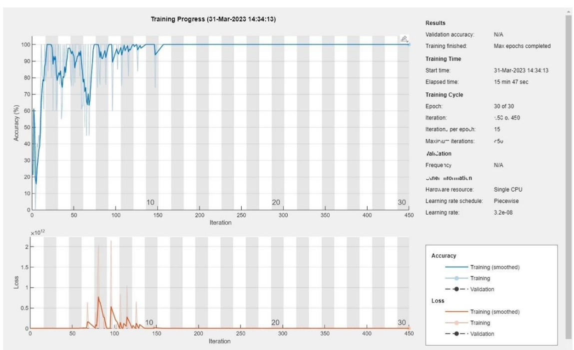

Figure 4. Training Results

Editors: Madhavi. K, P. Venkiah Babu, Bhargava. N
DOI:10.48001/978-81-966500-6-3-3 | ISBN: 978-81-966500-6-3

Based on the classification done the calculated parameters which have been passed to the Arduino as input. It has been observed that parameters decided were correct as RC Robot shows intended action.

6 Conclusion and Future Enhancement

The movement in the remote-controlled car was observed according to the gestures made this project is not a real time project because there was a delay between movement of the car after the gesture was given as input to MATLAB for processing. The model also depicted abnormal behavior when the input gesture was continuously changed without any pause. Between each input gesture there is a time interval when the car does not show any motion, nor does it continue in its previous state. The accuracy of CNN model directly affects the control of robotic car.

Real time refers to a variety of computer activities or other procedures that must ensure response times within a given timeframe (deadline), which is often a brief period. A real-time process is often one that takes place in discrete time steps with maximum length and moves quickly enough to have an impact on its surroundings, such inputs to a computer system. Therefore, the future development can focus on how to reduce the time interval when the RC Robot doesn't show any motion when the input gesture is changed. Further developments can also be made by implementing a better CNN Model to achieve greater control over the RC Robot.

References

Akyol, S., & Canzler, U. (2000). Gesture Control for use in Automobiles. IAPR Workshop on n Machine Vision Applications, 1–4. http://www.cvl.iis.u-tokyo.ac.jp/mva/proceedings/CommemorativeDVD/2000/papers/2000349.pdf

Anusha, K., Vasumathi, D., & Mittal, P. (2023). A Framework to Build and Clean Multilanguage Text Corpus for Emotion Detection using Machine Learning. Journal of Theoretical and Applied Information Technology, 101(3), 1344–1350.

ARVINDH, G., AVINASH, M. V. S., VENKAT, D. H. S., REDDY, M. A. T., & Dutt, V. I. (2021). Gesture controlled robot. International Journal of Creative Research Thoughts (IJCRT), 9(5).

Awaluddin, B. A., Chao, C. T., & Chiou, J. S. (2024). A Hybrid Image Augmentation Technique for User- and Environment-Independent Hand Gesture Recognition Based on Deep Learning. Mathematics, 12(9). https://doi.org/10.3390/math12091393

Awasthi, A., Sharma, N., Gupta, P., & Kushwaha, S. (2023). Hand Gesture Controller Robot Car Using Arduino. International Research Journal of Modernization in

Editors: Madhavi. K, P. Venkiah Babu, Bhargava. N
DOI:10.48001/978-81-966500-6-3-3 | ISBN: 978-81-966500-6-3

Engineering Technology and Science, 5(5), 6625–6629. https://doi.org/10.56726/IRJMETS40229

Barbhuiya, A. A., Karsh, R. K., & Jain, R. (2022). Gesture recognition from RGB images using convolutional neural network-attention based system. Concurrency and Computation: Practice and Experience, 34(24). https://doi.org/10.1002/cpe.7230

Deep Shakya, K., Mudgal, V., Sahu, P. K., & Sahu, N. K. (2020). Hand Gesture Control Electronic Car. International Research Journal of Modernization in Engineering Technology and Science, 2(6).

Gautam, S., & Mittal, P. (2022). Systematic Analysis of Predictive Modeling Methods in Stock Markets. International Research Journal of Computer Science, 9(11), 377–385. https://doi.org/10.26562/irjcs.2022.v0911.01

Gomathy, C., Niteesh, G., & Krishna, K. S. (2021). THE GESTURE CONTROLLED ROBOT. International Research Journal of Engineering and Technology (IRJET), 8(4).

Hasan, S. M., Mamun, S., Rasid, R., Mallik, A., & Rokunuzzaman, M. (2018). Development of a Wireless Surveillance Robot for Controlling from Long Distance. International Journal of Engineering Research And Management (IJERM), 5(9), 2349–2058. https://www.researchgate.net/publication/327833411

Hemane, H. S., Iyer, R., Kumar Mishra, A., & Sangar, A. (2022). Vehicle Controlled by Hand Gesture Using Raspberry pi. International Research Journal of Engineering and Technology (IRJET), 9(7), 650–657.

Jadhav, A., Pawar, D., Pathare, ., Sale, P., & R.Thakare. (2018). Hand Gesture Controlled Robot Using Arduino. International Journal for Research in Applied Science and Engineering Technology, 6(3), 2868–2870. https://doi.org/10.22214/ijraset.2018.3629

Mittal, P., & Gautam, S. (2023). Logistic Regression and Predictive Analysis in Public Services of AI Strategies. TEM Journal, 12(2), 751–756. https://doi.org/10.18421/TEM122-19

Pinto, R. F., Borges, C. D., Almeida, A. M., & Paula, I. C. (2019). Static Hand Gesture Recognition Based on Convolutional Neural Networks. Journal of Electrical and Computer Engineering, 2019. https://doi.org/10.1155/2019/4167890

Sahoo, J., JayaPrakash, A., Pławiak, P., & Samantray, S. (2022). Real-Time Hand Gesture Recognition Using Fine-Tuned Convolutional Neural Network. Sensors, 22(3). https://doi.org/10.3390/s22030706

Shah, R., Mulay, S., Deshmukh, V., Kulkarni, V., & Pote, M. (2020). Hand Gesture Control Car. International Journal of Engineering Research Technology (IJERT).

DOI:10.48001/978-81-966500-6-3-3 | ISBN: 978-81-966500-6-3

Shukla, A., Jain, A., Mishra, P., & Kushwaha, R. (2019). Human Gesture Controlled Car Robot. SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology, 11(02), 115–122. https://doi.org/10.18090/samriddhi.v11i02.5

Waskito, T. B., Sumaryo, S., & Setianingsih, C. (2020). Wheeled Robot Control with Hand Gesture based on Image Processing. Proceedings - 2020 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology, IAICT 2020, 48–54. https://doi.org/10.1109/IAICT50021.2020.9172032

Wu, X. H., Su, M. C., & Wang, P. C. (2010). A hand-gesture-based control interface for a car-robot. IEEE/RSJ 2010 International Conference on Intelligent Robots and Systems, IROS 2010 - Conference Proceedings. https://doi.org/10.1109/IROS.2010.5650294

Zagade, A., Jamkhedkar, V., Dhakane, S., Patankar, V., Kasture, D., & Gaike, P. (2018). a Study on Gesture Control Ardiuno Robot. International Journal of Scientific Development and Research (IJSDR), 3, 385–392.

Zhu, J., Xie, N., Cai, Z., Tang, W., & Chen, X. (2023). A comprehensive review of shared mobility for sustainable transportation systems. International Journal of Sustainable Transportation, 17(5), 527–551. https://doi.org/10.1080/15568318.2022.2054390

DOI:10.48001/978-81-966500-6-3-3 | ISBN: 978-81-966500-6-3