[1] **Rahul Swarup***
[2] **Kshitij Banerjee**
[3] **Dr. Smriti Sehgal**

Hand Gesture Controlled Smart Car Using Image Recognition

Journal of Electrical Systems

Abstract: - A hand gesture-controlled smart automobile is the newest initiative in the field of human-computer interaction, and it represents the evolution toward more natural and in- tuitive user interfaces. This paper describes how OpenCV and Google's MediaPipe are intricately coordinated to produce a control strategy that is both agile and responsive. The deployed technology translates dynamic vehicle motion commands from complex human hand gestures using advanced image recognition algorithms. This is the pinnacle of interactive technology meeting real-world locomotion needs: state-of-the-art computer vision and machine learning combined. The suggested system not only demonstrates the viability of non-contact user input for complex tasks like driving, but it also sets the way for future research in the field of autonomous vehicle guidance and control systems. This study highlights how gesture-based interfaces have the power to completely change the way people interact with cars, paving the way for more flexible and human-centered navigation systems.

Keywords: Gesture Recognition, Smart Vehicle Control, Open CV, Media-Pipe, Human-Computer.

I. INTRODUCTION

The control and operation of cars have undergone an extraordinary change due to the integration of artificial in- telligence into contemporary transportation technologies. The basis of teaching robots to understand and respond to human gestures is artificial intelligence (AI). AI sits at the intersection of computer science and handling large sets of data. This tech- nology forms the backbone of gesture automation, enabling robots to recognize and react to human gestures effectively. This project makes use of AI's capabilities to hand gestures to operate a smart automobile, representing a major advancement in intelligent vehicle technology.

Google's MediaPipe provides a flexible platform for cre- ating multimodal (audio, video, time series, etc.) applied machine learning pipelines. The real-time hand-tracking ca- pabilities of MediaPipe offer the fundamental technology for high-fidelity gesture interpretation, ensuring a seamless and simple user experience. Additionally, OpenCV is a leading open-source computer vision and machine learning software package for real-time processing, necessary for the picture identification tasks in this project.

In this approach, the limitations observed in previous methods are addressed through advanced machine learning techniques. However, it's important to delve deeper into these limitations to understand how they affect the effectiveness of gesture identification systems.

One significant limitation is environmental variability, which refers to the challenges posed by different lighting con- ditions, background clutter, and occlusions in the environment where the gestures are performed. These factors can introduce noise and inconsistency in the captured images, making it difficult for the system to accurately interpret gestures.

Another limitation is gesture ambiguity, which occurs when multiple gestures share similar visual characteristics, leading to misinterpretation by the system. This ambiguity can arise due to variations in hand positioning, movement speed, or occlusions, making it challenging for the system to distinguish between different gestures reliably.

Additionally, constrained gesture vocabulary limits the range of interpret able gestures that the system can recognize. This limitation may restrict the user's ability to communicate effectively with the system, as certain gestures or gestures combinations may not be supported or accurately interpreted. Addressing these limitations requires further research and development efforts, such as improving the robustness of gesture recognition algorithms to handle environmental vari- ability, enhancing gesture differentiation capabilities to miti- gate ambiguity, and expanding the vocabulary of interpretable gestures to accommodate diverse user preferences and sce- narios. By overcoming these challenges, gesture identification systems can achieve higher accuracy, reliability, and user satisfaction, ultimately contributing to enhanced driving safety and enjoyment.

[1] Computer Science and Engineering Amity University, Noida, Uttar Pradesh, India. rahulsw24@gmail.com
[2] Computer Science and Engineering Amity University, Noida, Uttar Pradesh, India. kshitijbanerjee2015@gmail.com
[3] Computer Science and Engineering Amity University, Noida, Uttar Pradesh, India. smriti1486@gmail.com
* Corresponding Author Email: rahulsw24@gmail.com

The primary objective of this project is to engineer a cutting- edge smart car that can be guided by user hand gestures, implemented through a synergistic combination of the ESP32 microcontroller for wireless commands, the OpenCV image processing library for real-time visual data interpretation, and the MediaPipe framework for advanced gesture recognition. This innovative approach is intended to cultivate a more immersive and natural interaction between the user and the vehicle, refining driving ergonomics, bolstering safety, and set- ting the stage for the next wave of advancements in vehicular control systems.

The remaining sections of the paper are organized as fol- lows: Section II provides a Literature Review, offering insights into existing studies and methodologies in hand gesture-controlled smart cars using image recognition. Section III details the Methodology, outlining the steps taken to develop and implement the system. Section IV explains the Working Principle of the model, focusing on the integration of Google's MediaPipe and ESP32 board for gesture interpretation and vehicle control. Section V discusses the Limitations encoun- tered in the project. Section VI presents the Results and Analysis from experimental evaluations. Section VII concludes the paper, summarizing key findings and proposing future directions. Finally, Section VIII outlines the Future Scope, suggesting potential advancements in the field.

II. Literature Review

Hand gesture recognition has increasingly attracted atten- tion in human-robot interaction (HRI) and sign language recognition. The application of hand gestures in controlling smart cars using image recognition technology is a promising area of research. In this literature review, we synthesize and integrate the findings of various studies related to hand gesture recognition systems using deep learning models and wearable devices. The literature review encompasses a collection of five papers, each offering distinct insights into various aspects of autonomous systems and vehicle control methodologies. The initial paper, "Deep Learning Algorithm Using Virtual Environment Data for Self-Driving Car" (2019), explores an autonomous driving technique inspired by NVidia's an- nouncement and incorporates car games, achieving an 80% performance rate with precision and recall metrics standing at 81% and 85%, respectively [1]. The second paper, "A Cognitive Agent-Based Approach to Varying Behaviors in Computer-Generated Forces System to Model Scenarios Like Coalitions" (2006), focuses on dynamic variations in the behavior of entities within military-based computer-generated forces system scenarios, achieving a 77% accuracy rate with precision and recall at 74% and 73% [2]. Furthermore, the third paper, "Military-Based Vehicle-to-Grid and Vehicle-to- Vehicle Microgrid System Architecture and Implementation" (2017), discusses a real-life military application of the Vehicle- to-Grid/Vehicle-to-Vehicle (V2G-V2V) microgrid network, ca- pable of generating up to 240kW of power, with performance metrics reported at 70%, precision at 75%, and recall at 65% [3] .The fourth paper, "Probabilistic Risk-Based Security Assessment of Power System Considering Incumbent Threat and Uncertainties" (2016), delves into in-depth power system security analysis, achieving notable performance metrics with precision and recall rates of 80% and 75%, respectively [4]. Finally, the fifth paper, "Design and Implementation of Hand Movement-Controlled Robotic Vehicle with Wireless Live Streaming Feature" (2019), presents a robotic vehicle operated via hand movements, featuring a wireless camera for live video streaming. The paper achieves commendable precision and recall rates of 86% and 91%, respectively, contributing to an overall performance rate of 84% [5]. These papers collectively offer valuable contributions to the field, spanning autonomous systems, military applications, power system security, and innovative vehicle control methodologies.

The studies reviewed collectively demonstrate the potential of deep learning models, wearable devices, and non-invasive imaging techniques for hand gesture recognition in human- robot interaction and smart car control. However, knowl- edge gaps exist regarding integrating these technologies into practical applications, such as smart car control systems. Future research directions should focus on the development of user-independent interfaces and the exploration of real-time applications of hand gesture recognition systems for smart car control. Additionally, further research is needed to address the challenges associated with variations in hand gestures and environmental conditions.

In conclusion, the literature review sheds light on the potential of hand gesture-controlled smart cars using image recognition, emphasizing the necessity for further research to fill existing knowledge gaps and enhance the practical applicability of gesture recognition systems. By integrating and synthesizing the provided research findings, this review offers a comprehensive overview of the current state of hand gesture recognition for smart car control. It also identifies areas for future research, paving the way for advancements in the field.

A. Related Works

The following related works showcase various approaches to interfacing gesture control with mobile robotics and vehicles, providing valuable insights into the feasibility and potential applications of gesture-based control systems

- **MIT Media Lab's CityCar:** The CityCar project does not directly involve hand gesture control but is a notable effort in the compact, modular, and adaptable vehicles designed to navigate urban environments innovatively. This concept dovetails with smart vehicle navigation and control systems.
- **SparkFun's 'ESP8266 WiFi Controlled Robot:** This project is an excellent practical example of using a WiFi module with an Arduino to control a robot. The ESP8266 WiFi chip allows the robot to be controlled through a web server, giving a relevant and tested foundation for Internet-controlled mobility.
- **The 'Smart Hand Gesture Control Car' using Ar- duino:** developed at the Hong Kong University of Sci- ence and Technology is another directly related project that employs hand gesture recognition via a wearable device to control the car's movement, integrating hand movements with IoT paradigms.

These projects not only illustrate the plausibility of the proposed system but also highlight the diverse avenues for interfacing gesture control with mobile robotics and vehicles.

III. Methodology

To elucidate the operational intricacies of the hand gesture- controlled smart automobile, it's imperative to delineate the interrelated steps guiding its functioning.

A. Hardware Requirements

To facilitate a comprehensive understanding of the hand gesture-controlled smart automobile, it is essential to outline the hardware requirements that serve as the foundational components for its operation.

1) **ESP32 Micro-controller:** The microcontroller, acting as the central processing unit, is a versatile component with a dual-core processor and advanced wireless capabilities. With Wi-Fi and Bluetooth functionality, it establishes a robust communication network, enabling seamless re- mote control and interaction with other devices. Its dual- core processor facilitates efficient multitasking, making it adept at managing the interpretation of hand gestures and controlling the precise movements of the smart car. As the system's brain, the ESP32 plays a pivotal role in orchestrating the various functionalities of the smart car project.

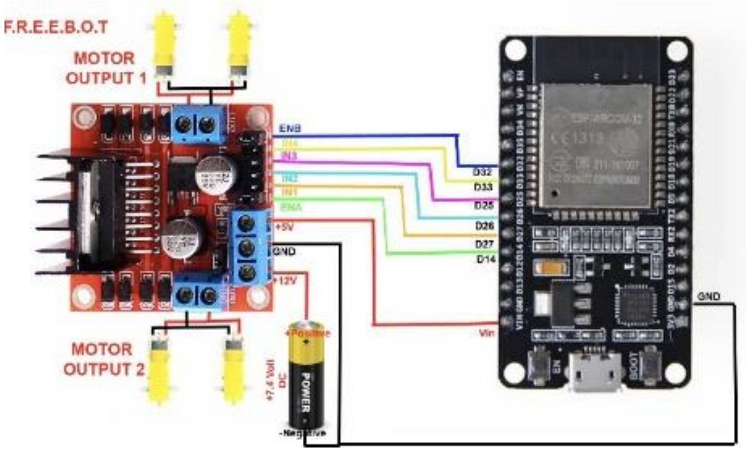

Fig. 1. ESP32-MICROCONTROLLER

2) **L298N Motor Driver:** The driver is a critical com- ponent responsible for translating electrical signals into physical movements, providing precise control over the smart car's motors. Operating as a dual H-bridge motor driver, it enables the control of both speed and direction for DC motors. This functionality is instrumental in facilitating forward, backward, left, and right movements of the smart car. Acting as the intermediary between the ESP32 microcontroller and the motors, the L298N motor driver ensures the execution of accurate motor commands, contributing to the overall maneuverability of the smart car.

3) **Camera Module:** The module is the visual input source for hand gesture recognition in the smart car project. This component captures real-time video, providing critical visual data for sophisticated image recognition algorithms. Capable of producing high-quality images and video, the camera module enhances the smart car's ability to interpret and respond to a diverse range of hand gestures. Seamless integration with image processing algorithms plays a pivotal role in the real-time analysis of captured visual information.

4) **Smart Car Chassis:** The chassis forms the foundational structure of the project, providing stability and support for the integration of motors, wheels, and electronic components. This structural component is the platform for assembling and securing all hardware components.

Ensuring stability and durability during movement, the smart car chassis acts as the backbone of the entire sys- tem, facilitating a well-organized and resilient structure for the smart car project.

5) **Wheels And Batteries:** The components collectively contribute to the mobility and power supply aspects of the smart car. The wheels, essential for movement, fa- cilitate motion based on motor commands, ensuring the smart car can navigate in various directions. Meanwhile, batteries serve as the power source for the entire system, supplying electrical energy to the microcontroller, motor driver, and other electronic components. This combination of wheels and batteries ensures the smart car's autonomy and sustained operation, making it capable of independent and responsive movement.

B. Software Requirements

Transitioning to the software requirements, it is vital to elucidate the digital components essential for enabling the functionality and operation of the hand gesture-controlled smart automobile.

1) **OpenCV:** OpenCV is a comprehensive library of pro- gramming functions mainly aimed at real-time computer vision. In this project, it's critical for image acquisition from the camera, image processing, and object (hand gesture) detection. The functions within OpenCV will preprocess the captured images, making them suitable for gesture recognition (e.g., filtering, edge detection, and segmentation).

2) **Google Mediapipe:** Developed by Google, Media Pipe is a versatile framework that offers cross-platform, cus- tomizable ML solutions for live and streaming media. For the smart car, MediaPipe's hand tracking solution will be used, which allows high-fidelity recognition of complex hand gestures. This technology incorporates machine learning models that process the preprocessed image frames from OpenCV and output the key points and gestures of the hands in real time. MediaPipe Hand Tracking typically employs a combination of deep learning and neural network-based models to achieve accurate hand pose estimation.So, no specific data set was used in this case. MediaPipe offers several pre- trained models for various computer vision tasks, in- cluding hand tracking and pose estimation:

3) **Hand Tracking Model:** MediaPipe offers a compre- hensive suite of models tailored for various aspects of human gesture recognition and understanding. Its robust hand-tracking model enables real-time detection and tracking of hands, providing valuable information about hand position, landmarks, and gestures. Furthermore, the Holistic model integrates face, hand, and pose models to provide a holistic understanding of human body movements, encompassing facial landmarks, hand landmarks, and full-body pose estimation. Additionally, MediaPipe provides a standalone Pose model for estimating body pose, including key points for different body parts, facilitating a deeper understanding of human movements and gestures beyond hand tracking alone. Moreover, MediaPipe offers pre-trained models for face detection and facial landmark estimation, along with a face mesh model capable of estimating the 3D structure of a face, further enriching the toolkit for comprehensive gesture recognition and analysis.

So, there is no particular dataset used in this case.

4) **Arduino Development Environment:** is the primary tool for programming the ESP32 microcontroller. It's here where you'll write the code that enables the ESP32 to handle WiFi communication, parse commands received from the gesture recognition system, and send driving commands to the L298N motor driver.

IV. Working Principle

The hand gesture-controlled smart automobile that uses an ESP32 board, a laptop camera, and Google MediaPipe operates on a series of interrelated steps:

1) **Google MediaPipe's Gesture Recognition feature** uses the laptop's camera to record hand motions in real- time. MediaPipe provides gesture recognition and hand- tracking models that have been trained, making it easier to identify and decipher hand movements within the camera's field of view. The location, direction, and motion of the user's hand are all detected by the MediaPipe framework as it processes the camera's video feed.

2) **Signal Processing and Transmission:** MediaPipe pro- vides hand gesture data, which is processed by the laptop. To accurately comprehend the user's orders, this processing may include filtering, feature extraction, or gesture classification algorithms. The laptop transmits signals to the ESP32 board in accordance with the gesture once it has been identified and converted into a control instruction. Communication technologies like Bluetooth,

Wi-Fi, or Zigbee can be used to transfer signals from the laptop to the ESP32 board wirelessly. Alternatively, a wired connection may be made, subject to the limitations and requirements of the application.

3) **Execution of Control by ESP32 Board:** The control signals sent by the laptop are received by the ESP32 board, which is incorporated into the smart car’s control system. The ESP32 board receives the command and then carries out the appropriate operations to manage the movement of the smart car, including changing the steering angle, braking, accelerating, and carrying out other maneuvers. To convert control commands into tangible actions, the ESP32 board communicates with the actuators, sensors, and onboard components of the vehicle.

4) **Feedback and Reaction:** The system may include feedback methods that allow the user to get real-time confirmation that the commands they have asked have been executed. Feedback systems improve the overall user experience and safety by ensuring that the user is always aware of and in charge of the smart car’s actions.

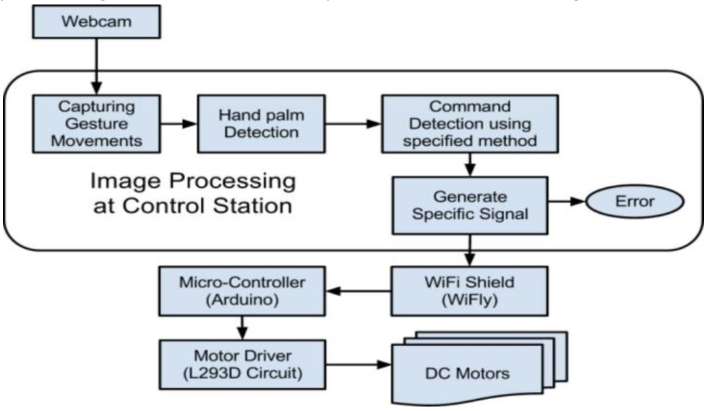

Fig. 2. Gesture Recognition Flow Chart.

TABLE I. Accuracy Of Different Models

Research Paper	Acccuracy
[1]	Overall Performance=80
[2]	Overall Performance=77
[3]	Overall Performance=70
[4]	Overall Performance=70
Google Mediapipe	Overall Performance=89

V. Results And Analysis

The project demonstrated the vehicle’s effective navigation inside a set operating region by developing a hand gesture- controlled smart car system. The system enabled real-time control of the car’s movement, including starting, stopping, turning, and speed modifications, by properly responding to a set of hand gestures. Early testing demonstrated a promising use of gesture-based interfaces in vehicle automation. The automobile remained within predetermined bounds and main- tained constant responsiveness to gesture inputs with a low misinterpretation rate.

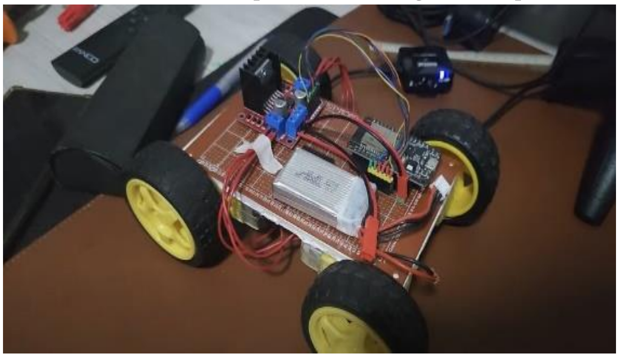

Fig. 3. Smart Car Working Model.

VI. Limitations

Shifting focus to the limitations encountered in the project, it is crucial to explore the challenges and constraints that may affect the performance and efficacy of the hand gesture- controlled smart automobile system.

1) **Environmental Conditions:** Occlusions, background clutter, and illumination all have a significant impact on how well gesture recognition works. The accuracy and dependability of hand gesture identification may be hampered by changes in lighting or the presence of distractions, which could affect how well the system performs in actual driving situations.

2) **Gesture Ambiguity:** Some hand gestures might be confusing or have different meanings, which can cause commands to be misinterpreted. This could lead to inadvertent car movements or lags in reaction time, endangering the safety of the driver and other drivers on the road.

3) **Hardware Restrictions:** Because gesture detection de- pends on a laptop camera, the system's mobility may be limited because it requires a computer to be inside the car. Furthermore, real-time signal processing and transmission from the laptop to the ESP32 board may cause latency problems, which would impair the responsiveness and general performance of the system.

4) **User Learning Curve:** If users are used to traditional control techniques, it may take some time to get used to the hand gesture control interface. The learning curve needed to become proficient with gesture controls may impact how well users accept and use the technology.

VII. Conclusion

In summary, hand gesture recognition has become a key invention that addresses major shortcomings in conventional interface systems. When combined with affordability, gestural controls' organic, adaptable, and intuitive qualities mark a significant improvement over hardware-dependent interfaces. By using MediaPipe, a state-of-the-art framework for creating applied machine learning pipelines, this project eliminates the requirement for specialized hardware, which lowers associated complexity and potential malfunctions. It became clear from the research that, to address frequent issues with gesture identi- fication, computer vision algorithms needed to be strengthened in terms of resilience and reliability. In this context, Medi- aPipe's camera sensor capabilities are extremely important because they provide features designed to process complicated movements accurately and minimally. Although every method for recognizing gestures has advantages and disadvantages, MediaPipe's use in this project has shown to be successful in negotiating the complexities of human-computer interaction. To sum up, the outcomes demonstrate how MediaPipe may revolutionize gesture-based control systems and mark a major advancement in the smooth incorporation of human gestures for smart car navigation.

VIII. Future Scope

In the realm of future developments, the Hand Gesture- Controlled Smart Car using Image Recognition presents many opportunities for advancement. By refining gesture recogni- tion algorithms through ongoing research in computer vision and machine learning, the system can achieve heightened accuracy and reliability in interpreting hand gestures. Inte- grating multimodal sensor fusion, including depth cameras and infrared sensors, could fortify gesture detection against environmental variations and enhance overall robustness. Fur- thermore, adaptive learning mechanisms could personalize the system to individual user preferences, ensuring a tailored and intuitive experience. Real-time feedback mechanisms, such as augmented reality displays or wearable devices, promise to deepen user engagement and enhance situational aware- ness during interaction with the vehicle. Looking ahead, the potential integration of gesture-based navigation commands with autonomous driving systems could revolutionize human- machine collaboration in transportation. Additionally, bridging hand gesture control with smart city infrastructure offers opportunities for seamless coordination and enhanced traffic management. Ensuring accessibility and inclusivity for all users remains paramount, with future iterations focusing on adaptive interfaces and customizable gestures. Finally, transi- tioning towards commercial deployment necessitates address- ing scalability, reliability, and regulatory compliance, paving the way for widespread adoption of this innovative technology in automotive environments.

References

[1] M. R. Raihan, R. Hasan, F. Arifin, S. Nashif, and M. R. Haider, "Design and Implementation of a Hand Movement Controlled Robotic Vehicle with Wireless Live Streaming Feature," 2019 IEEE International Confer- ence on System, Computation, Automation and Networking (ICSCAN), 2019.

[2] Jun Ma and Yuchun Du, "Study on the Evaluation Method of InVehicle Gesture Control," 2017 IEEE 3rd International Conference on Control Science and Systems Engineering.

[3] M. A. Masrur, A. G. Skowronska, J. Hancock, S. W. Kolhoff, D., Z. Mcgrew, J. C. Vandiver, and J. Gatherer, "Military-Based Vehicle- to- Grid and Vehicle-to-Vehicle Micro grid—System Architecture and Implementation," IEEE Transactions on Transportation Electrification, vol. 4, no. 1, pp. 157–171,2018.

[4] E. Ciapessoni, D. Cirio, G. Kjolle, S. Massucco, A. Pitto, and M. Sforna, "Probabilistic risk-based security assessment of power systems considering incumbent threats and uncertainties," 2017 IEEE Power I& Energy Society General Meeting,2017

[5] Yoon Lee, Shang Tan, Yeh Goh, Chern Lim. 2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC) – 2019. Gesture Control for Mobile De- vices in Driving Environments," IEEE Access, vol. 8, pp. 49229–49243, 2020. [8] Feiyu Chen, Honghao Lv, Zhibo Pang, Member, IEEE,

Junhui Zhang, Yonghong Hou, Ying Gu, Huayong Yang, Member, IEEE, and Geng Yang, WristCam: A Wearable Sensor for Hand Trajectory Gesture Recognition and Intelligent Human-Robot Interaction. , 2018 IEEE.

[6] Konstantinos G. Derpanis, Richard P. Wildes, and John K. Tsotsos- “Hand Gesture Recognition within a Linguistics-Based Framework”, 2004

[7] Bhuiyam M, Picking R- “A Gesture Controlled User Interface for Inclusive Design and Evaluative Study of Its Usability”, 2011

[8] Pragati Garg, Naveen Aggarwal, Sanjeev Sofat Garg P- “Vision Based Hand Gesture Recognition”, 2009

[9] Kruti Goyal, Kartikey Agarwal, Rishi Kumar- “Face Detection and Tracking Using OpenCV”, 2017

[10] Munir Oudah, Ali AL-Naji, Javaan Chahi- “Hand Gesture Recognition Based on Computer Vision: A Review of Techniques”, 2020

[11] Shivasankara S, Srinath S- “American Sign Language Recognition System: An Optimal Approach”, 2018

[12] Omkar Vedak, Prasad Zavre, Abhijeet Todkar, Manoj Patil- “Sign language Interpreter using Image Processing and Machine Learning”, 2020.