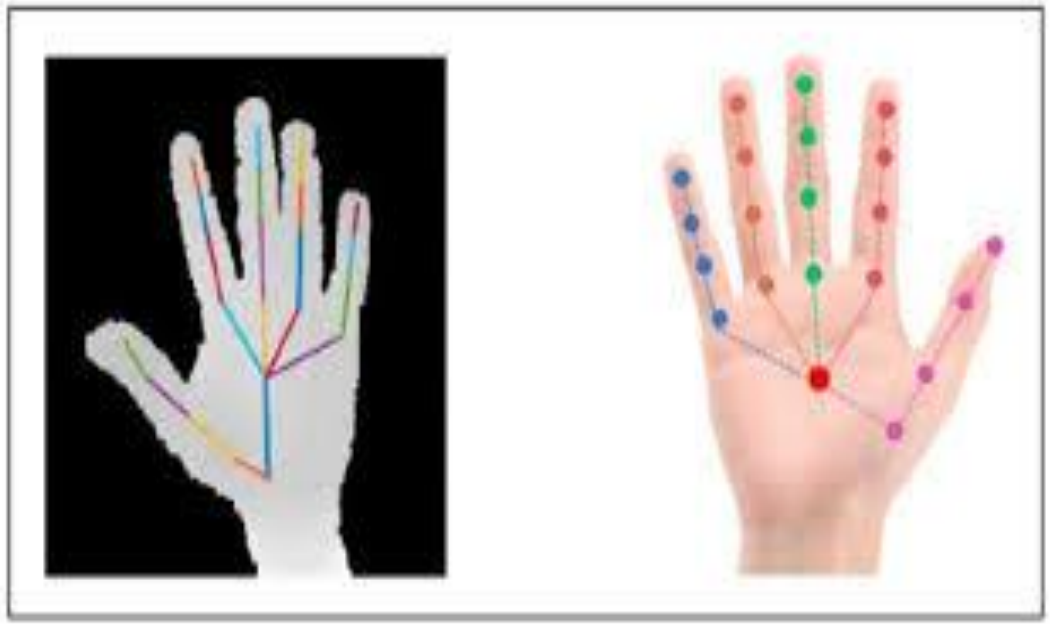
# GEN AI HACKATHON

- **TEAM NAME:** Palm’s Visionaries.
- **PROJECT TITLE**:Gesture-Based Human-Computer Interaction System using OpenCV, MediaPipe and Palm's text-bison-001.




**Presented by:**

K. Sucharitha

G. Sravya

M. Nandini

Y. Nikhitha

# PROBLEM STATEMENT

- The Gesture-Based Human-Computer Interaction System leverages real-time hand gesture recognition to enable users to interact with computers through intuitive hand movements. Utilizing computer vision techniques with OpenCV and MediaPipe, the system detects and interprets a variety of gestures, such as thumbs up, fist, open hand, and more. Integrated with a generative AI model, it provides descriptive narratives for recognized gestures, enhancing user experience. A user-friendly Streamlit interface facilitates easy interaction and visualization, making this system ideal for touchless control, interactive gaming, and assistive technologies.

# ABSTRACT


- The **Gesture-Based Human-Computer Interaction System** combines **computer vision, deep learning, and natural language processing** to provide an intuitive way for users to interact with machines using hand gestures. **OpenCV** is utilized for image preprocessing and feature extraction, while **MediaPipe** offers real-time hand tracking and gesture recognition. The system integrates **Palm's text-bison-001**, a powerful language model, to interpret gestures into meaningful actions or text-based responses. This framework enables applications in **virtual interfaces, accessibility solutions, and interactive systems**. The system ensures accuracy, responsiveness, and adaptability, making it suitable for real-world applications in **smart devices, robotics, and assistive technology**.

# INTRODUCTION

- The integration of the real and digital worlds is the aim of gesture recognition. It is considerably simpler to convey our intents and ideas to the computer via hand gestures. A simple and efficient touchless method of interacting with computer systems is through hand gestures. However, the limited end-user adoption of hand gesture-based systems is mostly caused by the significant technical challenges involved in successfully identifying in-air movements. Image recognition is one of the many ways that a computer may identify a hand gesture. The recognition of human movements is enabled through the implementation of a convolutional neural network (CNN). Within this study, we develop a simple hand tracking method for controlling a surveillance car operating on the Robot Operating System (ROS) by utilizing socket programming. The developed algorithm demonstrates promising implications for individuals with disabilities, including those who are deaf or have speech impairments.

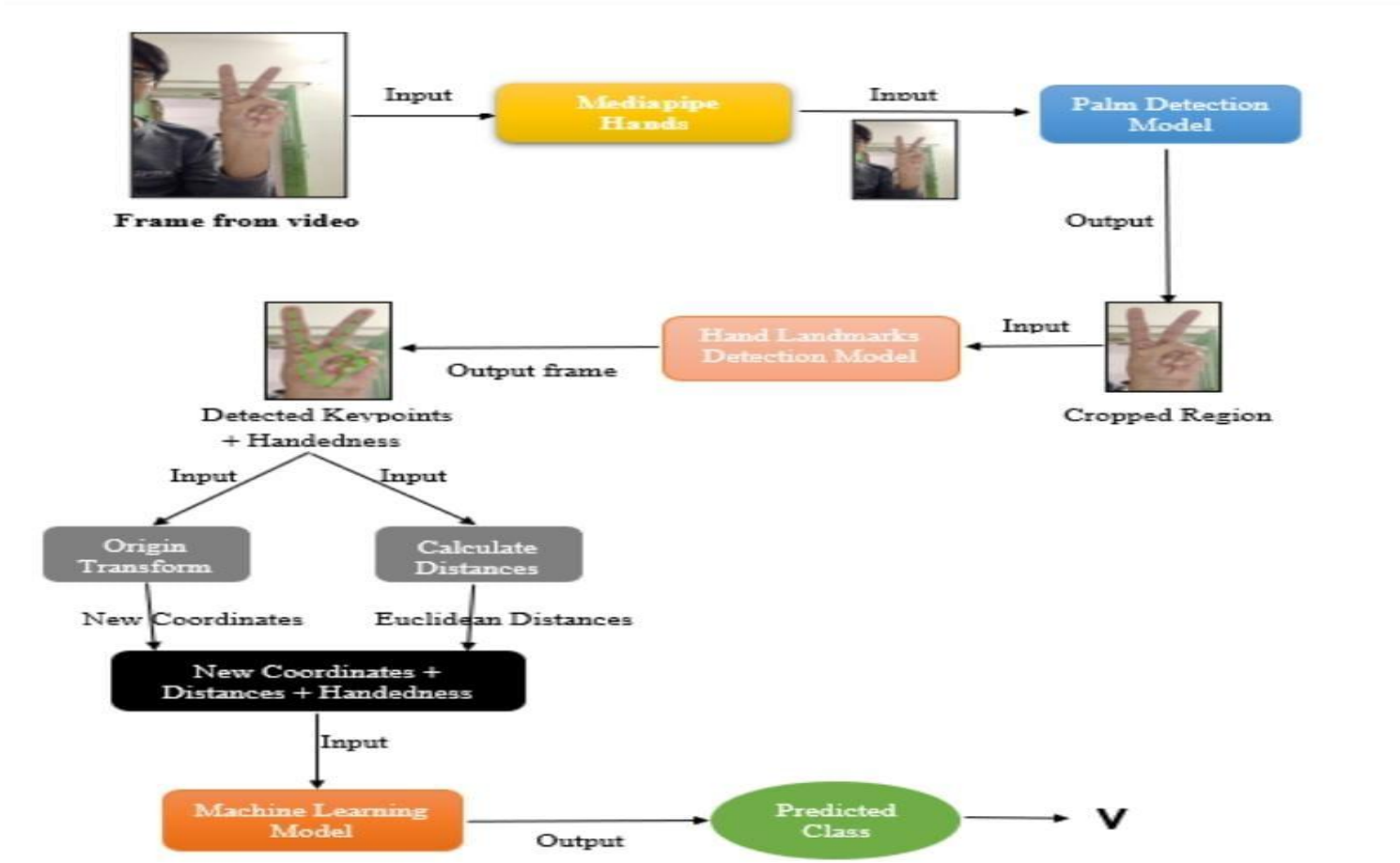
# LITERATURE REVIEW

- Researchers have extensively explored vision-based hand gesture recognition for human-computer interaction. Studies focus on challenges such as image acquisition, segmentation, tracking, and classification across different camera orientations. Various deep learning models and hyperparameter tuning techniques enhance recognition accuracy. MediaPipe, OpenCV, and Support Vector Machines (SVM) are widely used for hand tracking and gesture classification. Innovations like deformable convolutional networks, neuromorphic systems, and linear enhancement training (LET) improve accuracy. Applications range from sign language recognition to virtual interfaces and real-time sketching. Emerging approaches integrate AI, multimodal sensors, and depth estimation to refine gesture-based interaction and user experience.

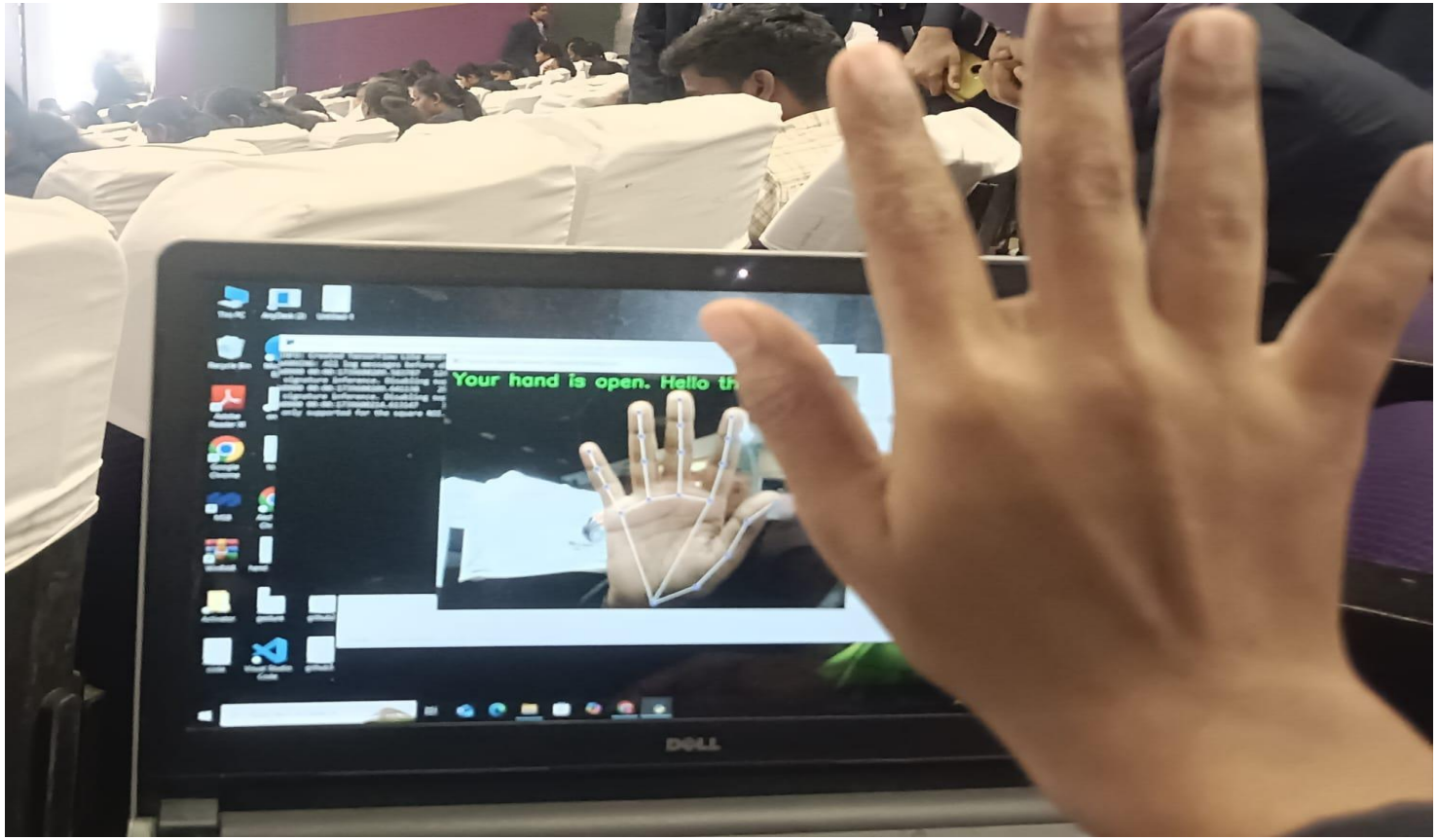
# MATERIAL AND METHODS

- MediaPipe is a versatile cross-platform framework for building machine learning pipelines that process audio, video, and time-series data. It enables seamless integration of ML models across Android, iOS, and web platforms, ensuring efficient data processing and analysis. Widely used in Google products like Gmail, Lens, Maps, and YouTube, MediaPipe offers robust tools for data preprocessing, feature extraction, and model training. Its cross-platform compatibility allows for consistent user experiences across devices. With applications in **gesture recognition, augmented reality, and computer vision**, MediaPipe is a powerful tool for developing high-performance, scalable machine learning solutions for real-world applications.

# BLOCK DIAGRAM



# OUTPUT



# CONCLUSION

- The results obtained from our algorithm, leveraging the capabilities of MediaPipe and OpenCV, demonstrate the ease and accuracy with which our model predicts hand gestures. The model has been trained using a dataset comprising over 3000 images encompassing various configurations of alphabets from A to Z and numbers from 1 to 9. The significance of this algorithm lies in its potential to benefit individuals with disabilities, particularly those who are deaf and mute. Furthermore, this technology has the potential to contribute to the overall improvement in the quality of life for the deaf and mute community.There is a growing opportunity for their widespread adoption and increased benefits. Instead of relying on traditional input devices like keyboards and mice, this application capitalizes on the natural movements of hand gestures.

# REFERANCE


- https://github.com/
- Al Farid, F.; Hashim, N.; Abdullah, J.; Bhuiyan, M.R.; Shahida Mohd Isa, W.N.; Uddin, J.; Haque, M.A.; Husen, M.N. A Structured and Methodological Review on Vision-Based Hand Gesture Recognition System. J. Imaging 2022, 8, 153. https://doi.org/10.3390/jimaging8060153.
- Bhaumik, G., Verma, M., Govil, M.C. et al. HyFiNet: Hybrid feature attention network for hand gesture recognition. Multimed Tools Appl (2022). https://doi.org/10.1007/s11042-021-11623-3.