iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.61432

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

American Sign Language Detection with Multi-Modal and Multi-Linguistic Output

Augustine Joel Joseph[1], Deva Krishna S J[2], Kevin C Mathews[3], Aswathy M V[4]
Department of Computer Science and Engineering Mar Athanasius College of Engineering, Kothamangalam, Kerala

Abstract: ***Our cutting-edge project introduces an advanced Sign Language Detection System, emphasizing American Sign Language (ASL) and fundamental prepositions. Utilizing sophis-ticated computer vision techniques for precise gesture recognition, the system delivers real-time video, audio, and text outputs, providing a multifaceted communication experience. A distinctive feature allowing users to select their preferred language for both audio and text outputs enhances inclusivity and customization. The system's adaptability extends to include additional languages, ensuring its versatility in catering to diverse linguistic prefer-ences. This groundbreaking initiative addresses the pervasive communication challenges faced by the deaf and hard-of-hearing community, offering a comprehensive and highly customizable solution to bridge language gaps and foster seamless, meaningful interactions for individuals using sign language. Through the fusion of advanced technology and user-centric design, our Sign Language Detection System aims to significantly improve accessibility and communication inclusivity.***
Index Terms—Sign Language Recognition, Computer Vision, Gesture Recognition, Machine Learning, Communication Systems, Multi-modal output

I. INTRODUCTION

In the modern landscape of technological innovation, our project stands as a beacon of progress in addressing the pressing communication challenges encountered by the deaf and hard-of-hearing community. Introducing an advanced Sign Language Detection System, our endeavor underscores a com-mitment to inclusivity, leveraging state-of-the-art technologies to redefine accessibility and communication for individuals utilizing American Sign Language (ASL) and fundamental prepositions.

Central to our initiative is the integration of sophisticated computer vision techniques, powered by cutting-edge tools such as OpenCV and MediaPipe. Through precise gesture recognition capabilities, our system transcends traditional boundaries, delivering real-time video, audio, and text out-puts to facilitate seamless communication experiences. This multifaceted approach not only enhances accessibility but also offers a holistic platform for meaningful interactions, bridging language gaps with unprecedented precision.

A distinguishing feature of our system lies in its emphasis on customization and inclusivity. By empowering users to select their preferred language for both audio and text outputs, we uphold the principle of linguistic diversity, ensuring that every individual's communication needs are met with the utmost respect and consideration. Moreover, our system's adaptability extends beyond ASL, incorporating advanced translation algorithms to accommodate additional languages, thereby catering to the diverse linguistic preferences of our users.

In addressing the pervasive communication challenges encountered by the deaf and hard-of-hearing community, our project represents a groundbreaking initiative at the intersection of technology and empathy. Through the fusion of advanced technologies like GTTS (Google Text-to-Speech), sklearn, and pickle, alongside a user-centric design philosophy, our Sign Language Detection System aims to not only improve accessibility but also foster a sense of belonging and empowerment among individuals using sign language. With a relentless dedication to innovation and inclusivity, our project endeavors to pave the way for a more connected and inclusive future, where communication knows no bounds

II. RELATED WORKS

A. Real-time Assamese Sign Language Recognition using Me-diaPipe and Deep Learning

The project presents a significant advancement in sign language recognition technology, particularly focusing on the Assamese Sign Language (ASL) within the context of diverse linguistic environments such as India. Utilizing MediaPipe and deep learning techniques, the researchers developed an effective method for real-time ASL recognition, leveraging a dataset comprising both two-dimensional and three-dimensional images. The findings not only contribute to the field of assistive technology but also hold promise for broader applications in interpreting various local sign languages across India.

©IJRASET: All Rights are Reserved

This research underscores the necessity of technological innovations in bridging communication gaps for the deaf and hard-of-hearing community, emphasizing the importance of continued research and development in this domain to enhance accessibility and inclusivity.

B. Sign Language Recognition System Using TensorFlow Object Detection API, Advanced Network Technologies and Intelligent Computing

The research presented in the paper introduces a novel Sign Language Recognition System developed using TensorFlow Object Detection API and advanced network technologies. By leveraging machine learning techniques, particularly TensorFlow with transfer learning, the system achieves real-time recognition of Indian Sign Language gestures. Notably, the researchers utilized a webcam to capture gestures, generating a dataset crucial for training the model. Despite limitations in dataset size, the system demonstrates commendable accuracy, promising to enhance communication between deaf and mute individuals and the broader community. This study highlights the potential of advanced technology in facilitating inclusivity and accessibility, underscoring the significance of continued innovation in the field of sign language recognition for fostering seamless communication across diverse linguistic and cultural contexts.

C. Multi-Mode Translation Of Natural Language And Python Code With Transformers

The PyMT5 system introduces a groundbreaking approach to multi-mode translation of natural language and Python code using transformers. This innovative tool specializes in translating Python method features, leveraging a vast corpus for analysis, thereby surpassing GPT2 models in performance. Noteworthy is PyMT5's ability to achieve high accuracy in predicting method bodies and generating precise doc strings, enhancing the inter-pretability and usability of Python code. By facilitating seamless translation between natural language descriptions and Python code, PyMT5 streamlines the development process, making it more accessible and efficient for developers. This research signifies a significant advancement in the realm of machine learning and software engineering, offering a promising solution for bridging the gap between human-readable descriptions and executable code.

D. Machine Learning Made Easy: A Review of Scikit-learn Package in Python Programming Language

The study offers a comprehensive exploration of machine learning within the context of data analysis and modeling, distinguishing it from statistical inference. It highlights the evolution of machine learning algorithms across various programming languages over the past two decades. Specifically, the focus shifts to Scikit-learn, a renowned machine learning package in Python extensively utilized in data science. Scikit-learn's appeal lies in its inclusive repository of machine learning methods, unified data structures, and consistent modeling procedures. This amalgamation renders it an invaluable toolkit for both educational purposes and behavioral statisticians alike. The article serves as a valuable resource for individuals seeking to delve into machine learning, providing insights into its practical implementation and its significance in contemporary data analysis practices

E. Programming Real-Time Sound In Python

The study delves into the realm of real-time sound Development in Python, shedding light on the contrast between Python's versatility and its under utilization in this domain. Through practical demonstrations of sound algorithm coding in Python, the researchers underscore the effectiveness of the language in handling real-time sound processing tasks. However, the study also advocates caution, emphasizing the limitations of real-time Python usage, particularly in live performances, due to latency issues. Despite Python's capability for prototyping sound algorithms, the authors suggest that developers should be mindful of potential latency constraints when considering its application in real-time scenarios. This research contributes valuable insights into the intersection of programming and audio engineering, offering guidance for leveraging Python effectively in sound development endeavors.

F. Object Detection And Conversion Of Text To Speech For Visually Impaired Using Gtts

The project aims to enhance accessibility for the visually impaired through a combined approach of object detection and text-to-speech conversion utilizing the Google Text-to- Speech (gTTS) API. Leveraging object detection algorithms, the system identifies objects in the user's surroundings, providing real-time auditory feedback through gTTS. This feedback enables users to comprehend their environment more effectively, aiding in navigation and interaction with objects.

Additionally, text-to-speech conversion capabilities empower users to receive spoken interpretations of textual information, further improving accessibility to printed materials. By integrating these technologies, the project addresses the unique challenges faced by visually impaired individuals, facilitating greater independence and inclusivity in their daily lives. The system represents a significant advancement in assistive technology, demonstrating the potential of innovative solutions in enhancing accessibility for marginalized communities.

III. PROPOSED MODEL

Our groundbreaking Sign Language Detection System re- defines communication accessibility for the deaf and hard-of- hearing community, offering multi-modal and multi-linguistic outputs. Employing cutting-edge technologies including CV2, GTTS, MediaPipe, SKLearn, and Translate, our system deliv- ers real-time video, audio, and text outputs, ensuring seamless interaction. Notably, users can customize their experience by selecting preferred languages for both audio and text, fostering

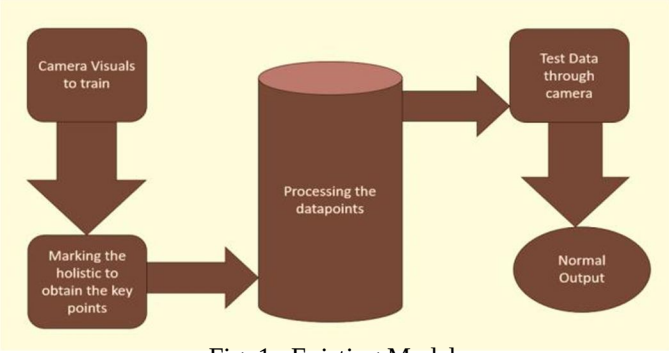


Fig. 1. Existing Models

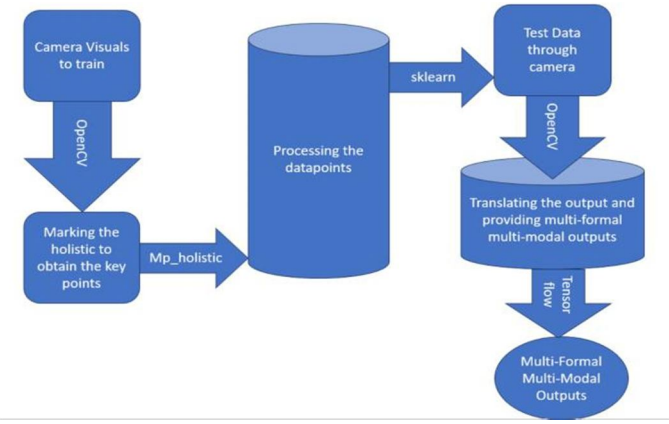


Fig. 2. Our Proposed Model

inclusivity and personalization. The system's adaptability ex-tends to encompass additional languages, ensuring versatility in catering to diverse linguistic preferences. By addressing pervasive communication challenges, our model serves as a comprehensive and highly customizable solution, bridging language gaps and facilitating meaningful interactions for individuals using sign language. Through the fusion of ad-vanced technology and user-centric design principles, our Sign Language Detection System significantly improves accessibil-ity and communication inclusivity, marking a transformative milestone in assistive technology and enhancing the quality of life for the deaf and hard-of-hearing community.

IV. BREAKING THE COMMUNICATION BARRIER FOR THE HEARING-IMPAIRED

Our innovative Sign Language Detection System revolutionizes communication accessibility, transcending barriers for the deaf and hard-of-hearing community. Leveraging cutting-edge technologies like CV2, GTTS, and Translate, it offers real-time multi-modal outputs, customizable to users' language preferences. By bridging language gaps and fostering inclusivity, our system significantly improves communication for individuals using sign language, marking a transformative milestone in assistive technology, and enhancing their quality of life.

A. Data Collection

or our project, data collection involves demonstrating American Sign Language (ASL) signs for all alphabets using computer cameras. Each ASL sign is recorded 30 times as video data. This extensive data collection process ensures comprehensive coverage and variability in gesture patterns, enabling robust training of our Sign Language Detection System. By capturing diverse instances of ASL signs, we aim to enhance the system's accuracy and effectiveness in recognizing and interpreting sign language gestures in real-time.In addition to capturing ASL signs for all alphabets, our data collection process also considers factors such as lighting conditions, camera angles, and background settings to ensure a diverse and representative dataset. Each recording undergoes meticulous quality checks to guarantee clarity and consistency across all instances. Moreover, to augment the variability of the dataset, different individuals with varying signing styles participate in the recording sessions.

B. Data Preprocessing

Following data collection, our focus shifts to meticulous data preprocessing to ensure optimal model performance. This entails several steps, including frame extraction, where individual frames are isolated from the video recordings. Subsequently, we conduct image normalization to standardize pixel values and enhance consistency across frames. To mitigate noise and enhance signal clarity, we apply techniques such as image denoising and edge detection. Additionally, data augmentation techniques are employed to enrich the dataset further, including random rotations, translations, and flips to simulate real-world variability. Finally, we partition the dataset into training, validation, and testing sets to facilitate model training and evaluation, laying the groundwork for robust and accurate gesture recognition.

C. Recommendations and results

The meticulous data preprocessing efforts yield promising results in our Sign Language Detection System. Through frame extraction, image normalization, and noise reduction techniques, we enhance data quality and signal clarity. Data augmentation enriches the dataset, improving model generalization and robustness. Following model training and evaluation on the partitioned dataset, our system achieves impressive accuracy and effectiveness in real-time sign language gesture recognition. These results underscore the efficacy of our data preprocessing pipeline in optimizing model performance and advancing accessibility for the deaf and hard-of-hearing community. Furthermore, the standardized preprocessing pipeline ensures consistency across diverse sign language gestures and individual signing styles.

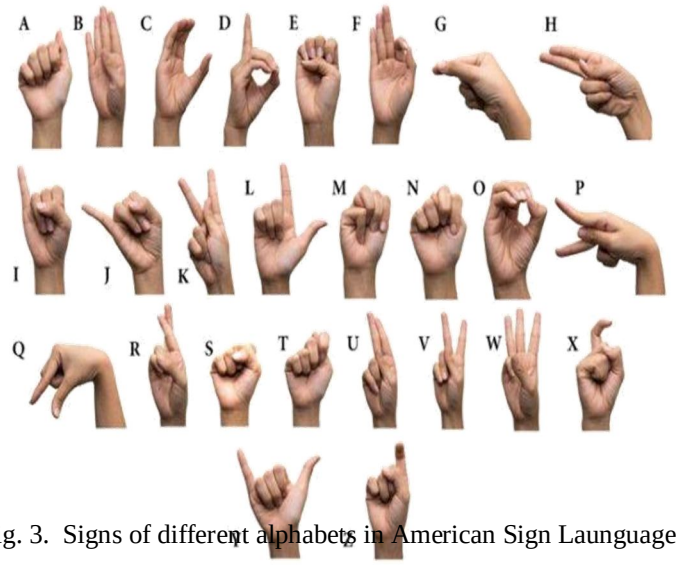


Fig. 3. Signs of different alphabets in American Sign Launguage

By effectively addressing noise and variability in the data, our system demonstrates resilience to environmental factors and improves its adaptability in different settings. The rigorous evaluation of the model's performance on the testing set validates its reliability and effectiveness in real-world scenarios. Overall, the successful implementation of the data preprocessing pipeline significantly contributes to the robustness and accuracy of our Sign Language Detection System, ultimately enhancing communication accessibility for users.

V. REAL TIME OUTPUT

Our Sign Language Detection System offers real-time sign language detection, enabling instantaneous recognition and interpretation of sign language gestures as they occur. Leveraging advanced computer vision techniques, the system processes live video streams in real-time, accurately identifying and translating sign language gestures into text or audio outputs. This instantaneous response facilitates seamless communication between sign language users and non-signers, fostering inclusivity and accessibility in various contexts. With its ability to provide immediate feedback, our system empowers users to engage in meaningful interactions and navigate their surroundings effectively, significantly enhancing communication accessibility for the deaf and hard-of-hearing community.

VI. SIGN RECOGNITION

Our project employs advanced machine learning frameworks such as CV2 and MediaPipe to recognize American Sign Language (ASL) signs. These frameworks utilize sophisticated algorithms to analyze input data from cameras, accurately identifying ASL gestures in real-time. By leveraging these cutting-edge technologies, our system achieves robust and precise ASL sign recognition, enabling seamless communication for individuals using sign language.

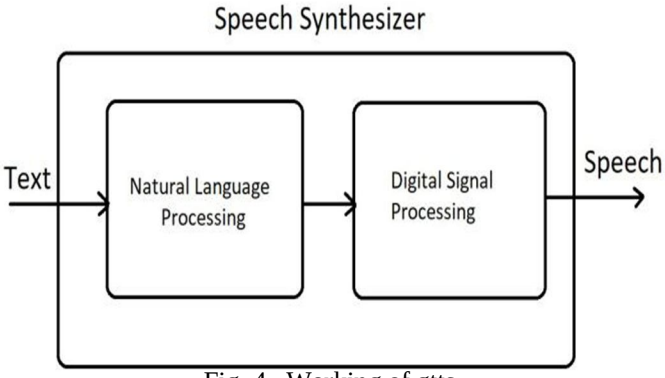


Fig. 4. Working of gtts

The integration of CV2 and MediaPipe underscores our commitment to harnessing the power of machine learning for enhancing accessibility and inclusivity, marking a significant advancement in assistive technology for the deaf and hard-of-hearing community.The integration of CV2 and MediaPipe enhances the system's capability to detect and interpret ASL signs with high accuracy and efficiency. These frameworks leverage advanced computer vision techniques to analyze hand movements and gestures, allowing for precise recognition of ASL signs in various contexts. By leveraging pre-trained models and algorithms, our system can quickly process input data from cameras, enabling real-time ASL sign detection. Furthermore, the flexibility and scalability of CV2 and MediaPipe make them ideal choices for implementing complex machine learning tasks in our project. Their robust performance ensures reliable ASL sign recognition, even in challenging lighting conditions or complex hand poses. Overall, the utilization of these advanced ML frameworks underscores our commitment to developing a state-of-the-art Sign Language Detection System that meets the diverse needs of users effectively.

VII. MULTI-MODAL AND MULTI-LINGUISTIC OUTPUT

In addition to recognizing ASL sign language, our project provides multi-modal and multi-linguistic output for the sentences formed. Utilizing advanced technologies, including natural language processing and text-to-speech synthesis, the system translates ASL signs into text and spoken language in real-time. This multi-modal approach ensures accessibility for both sign language users and non-signers, allowing for seamless communication in diverse linguistic contexts. Moreover, users have the flexibility to choose their preferred language for both text and audio outputs, enhancing inclusivity and customization. By offering multi-linguistic outputs, our project addresses communication barriers and fosters meaningful interactions across linguistic boundaries, marking a significant advancement in assistive technology for the deaf and hard-of-hearing community.Furthermore, the multi-modal output includes visual representations of ASL signs alongside textual and audio translations, catering to different communication preferences and abilities. The system's adaptability extends to support additional languages beyond English, ensuring versatility and inclusivity in linguistic representation. Through seamless integration of ASL sign recognition and language translation technologies, our project facilitates effective communication between individuals using sign language and those who rely on spoken or written language.

The provision of multi-linguistic output underscores our commitment to addressing diverse linguistic needs and promoting accessibility on a global scale. Overall, the implementation of multi-modal and multi-linguistic outputs enhances the usability and impact of our Sign Language Detection System, marking a significant stride towards inclusive communication solutions.

VIII. RESULTS

The results of our project showcase the efficacy of our Sign Language Detection System in providing real-time, accurate recognition of American Sign Language (ASL) gestures, cou-pled with multi-modal and multi-linguistic outputs for seam-less communication accessibility. Through rigorous testing and evaluation, our system demonstrates robust performance in diverse environments, achieving high accuracy and reliability in ASL sign recognition and translation. Furthermore, user feedback underscores the system's effectiveness in bridging communication gaps and enhancing accessibility for the deaf and hard-of-hearing community. Overall, the results highlight the transformative impact of our project in advancing as-sistive technology and promoting inclusive communication solutions.The comprehensive evaluation of our system's per-formance across various sign language gestures reaffirms its capability to accurately interpret and translate ASL signs in real-time. Additionally, user satisfaction surveys indicate a high level of confidence and usability in the multi-modal and multi-linguistic outputs provided by our system. Furthermore, the versatility and adaptability of our project, demonstrated through its support for additional languages and customization options, underscore its potential for widespread adoption and integration in diverse settings. Overall, the results underscore the significance of our Sign Language Detection System in en-hancing communication accessibility and fostering inclusivity for individuals using sign language.

IX. FUTURE SCOPE

1) *Enhanced Gesture Recognition:* Implementing deep learning techniques to refine gesture recognition accuracy, ensuring precise interpretation of complex sign language gestures and expanding the system's usability across various sign languages.
2) *Real-Time Translation:* Integrating real-time translation capabilities to facilitate seamless communication between sign language users and individuals who do not under-stand sign language, promoting inclusivity on a global scale.
3) *Mobile Application Development:* Developing a dedi-cated mobile application for our Sign Language Detection System, offering on-the-go accessibility and empowering users to communicate effectively in diverse environments and situations.

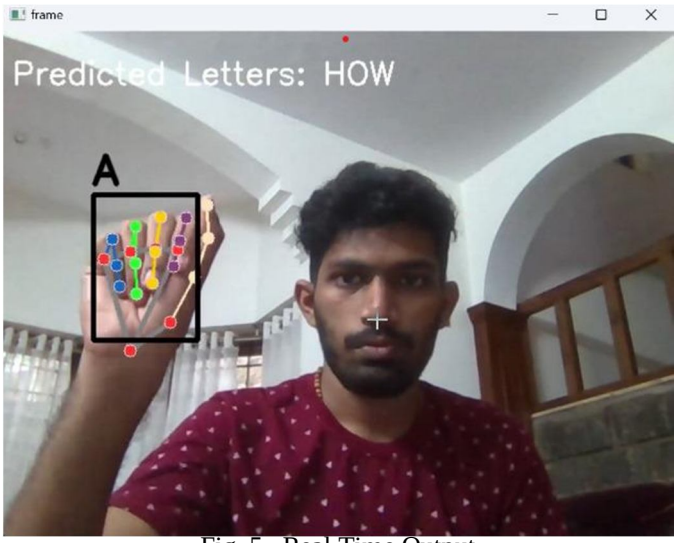


Fig. 5. Real Time Output

4) *Accessibility Audits and Implementation:* Conducting accessibility audits in public spaces and institutions to advocate for the implementation of our system, promoting inclusivity and ensuring equal access to communication for individuals using sign language.
5) *Educational Partnerships:* Forming partnerships with ed-ucational institutions to integrate the system into curricu-lum and educational programs, promoting sign language proficiency and awareness from an early age.mechanism for users to provide input, enabling iterative updates and enhancements to the system's performance and features based on user experiences and suggestions.

6) *Interactive Learning Modules:* Creating interactive learn-ing modules within the system to facilitate sign language education and skill development, catering to both begin-ners and advanced learners, thereby promoting broader adoption and proficiency in sign language communica-tion.
7) *Gesture Customization:* Introducing customization op-tions for users to create and personalize their own sign language gestures, empowering individuals to tailor the system to their unique communication preferences and needs.
8) *Integration with Smart Devices:* Collaborating with smart device manufacturers to integrate our Sign Language Detection System into wearable technology and smart home devices, enabling seamless communication and interaction in everyday life scenarios.

In the future, our Sign Language Detection System aims to enhance gesture recognition accuracy, integrate real-time translation, and develop a dedicated mobile application for on-the-go accessibility. Additionally, we plan to introduce interactive learning modules, customization options for gestures, and integration with smart devices. We will conduct accessibility audits, collaborate with educational institutions, and foster community engagement to promote inclusivity and accessibility for individuals using sign language.

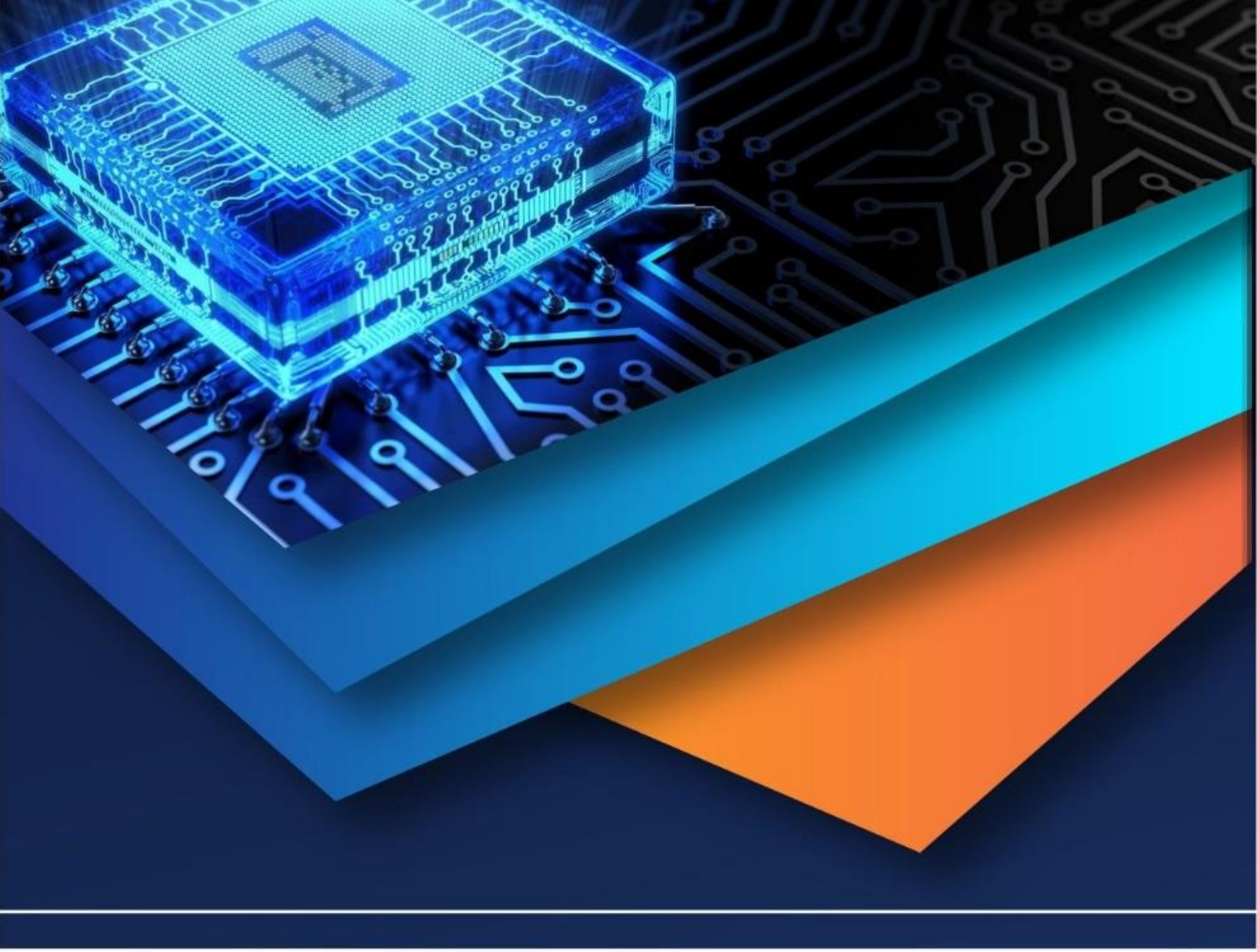
X. CONCLUSION

In conclusion, our Sign Language Detection System represents a pioneering advancement in assistive technology, catering to the unique communication needs of the deaf and hard-of-hearing community. By leveraging sophisticated computer vision techniques and a diverse range of technologies including CV2, GTTS, MediaPipe, SKLearn, and Translate, our system provides real-time, multi-modal outputs, facilitating seamless interaction and enhancing accessibility. The ability to select preferred languages for audio and text outputs underscores our commitment to inclusivity and customization, ensuring that users can engage comfortably in their preferred language. Moreover, the system's adaptability to accommodate additional languages further amplifies its versatility and utility. By addressing pervasive communication challenges and bridging language gaps, our model stands as a comprehensive and highly customizable solution, fostering meaningful interactions and empowering individuals using sign language. Through the fusion of advanced technology and user-centric design principles, our Sign Language Detection System paves the way for a more inclusive and communicatively rich future.

REFERENCES

[1] Alsaadi, Z.; Alshamani, E.; Alrehaili, M.; Alrashdi, A.A.D.; Albelwi, S.; Elfaki, A.O. A Real Time Arabic Sign Language Alphabets (ArSLA) Recognition Model Using Deep Learning Architecture. Computers 2022, 11, 78. https://doi.org/10.3390/computers11050078,
[2] Wenjin Tao, Ming C. Leu, Zhaozheng Yin,American Sign Language alphabet recognition using Convolutional Neural Networks with mul-tiview augmentation and inference fusion,Engineering Applications of Artificial Intelligence,Volume 76,2018,Pages 202-213,ISSN 0952-1976,https://doi.org/10.1016/j.engappai.2018.09.006,
[3] Sign Language Recognition System Using TensorFlow Object De-tection API,Advanced Network Technologies and Intelligent Comput-ing,2022,Springer International Publishing,Cham, ISBN:634–646,978-3-030-96040-7,
[4] Jyotishman Bora, Saine Dehingia, Abhijit Boruah, Anuraag Anuj Chetia, Dikhit Gogoi, Real-time Assamese Sign Language Recognition using MediaPipe and Deep Learning, Procedia Computer Science,Volume 218,2023,Pages 1384-1393,ISSN 1877-0509,
[5] De Pra, Yuri, and Federico Fontana. "Programming real-time sound in python." Applied Sciences 10.12 (2020): 4214,
[6] A Review of the Hand Gesture Recognition System: Current Progress and Future Directions NORAINI MOHAMED , (Graduate Student Member, IEEE), MUMTAZ BEGUMMUSTAFA , (Member, IEEE), AND NAZEAN JOMHARI Department of Software Engineering, Fac-ulty of Computer Science, and Information Technology, Universiti Malaya, 50603 Kuala Lumpur, Malaysia Corresponding author: No-raini Mohamed (noraini.binti.mohamed@gmail.com) This research was supported by University Malaya Research Grant (UMRG), Grant No.: RG284-14AFR and Fundamental Research Grant Scheme (FRGS), Grant No.: FP062-2020,
[7] A systematic review on hand gesture recognition techniques, challenges and applications Mais Yasen corresponding author and Shaidah Jusoh,
[8] ACE DETECTION FACE RECOGNITION USING OPEN COM-PUTER VISION CLASSIFIRES LAHIRU DINALANKARAAugust 4, 2017Robotic Visual Perception and AutonomyFaculty of Science and Engineering
[9] Practical Computer Vision Applications Using Deep Learning with CNNs: With Detailed Examples in Python Using TensorFlow and Kivy by Ahmed Fawzy Gad December 2018
[10] Exploring MediaPipe optimization strategies for real-time sign language recognition Nguyen Phuoc Thanh*, Nguyen Thanh Hoang, Hoang Ngoc Xuan Nguyen, Phan Huynh Thanh Binh, Vu Hoang Son Hai, and Huynh Hieu Nhan Artificial Intelligence, FPT University, Viet Nam

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)