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.2024.61514

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

Sign Language Recognition for Real-time Communication

Mr. Dakshesh Gandhe[1], Mr. Pranay Mokar[2], Mr. Aniruddha Ramane[3], Dr. R. M. Chopade[4]
Information Technology, SPPU

Abstract: ***Sign language is an essential communication tool for India's Deaf and Hard of Hearing people. This study introduces a novel approach for recognising and synthesising Indian Sign Language (ISL) using Long Short-Term Memory (LSTM) networks. LSTM, a kind of recurrent neural network (RNN), has demonstrated promising performance in sequential data processing. In this study, we leverage LSTM to develop a robust ISL recognition system, which can accurately interpret sign gestures in real-time. Additionally, we employ LSTM-based models for ISL synthesis, enabling the conversion of spoken language into sign language for improved inclusivity and accessibility. We evaluate the proposed approach on a diverse dataset of ISL signs, achieving high recognition accuracy and natural sign synthesis. The integration of LSTM in ISL technology holds significant potential for breaking down communication barriers and improving the quality of life for India's deaf and hard of hearing people.***

Keywords: Indian Sign Language, ISL Recognition, ISL Synthesis, LSTM, Deaf and Hard of Hearing, Communication Accessibility, Sequential Data Processing, Real-time Sign Language Recognition.

I. INTRODUCTION

Communication is the foundation of human contact, yet traditional methods frequently fail to capture the many ways in which people express themselves. Sign language, a vital medium for the deaf and hard of hearing community, embodies a rich linguistic system, conveying intricate thoughts and emotions through gestures and movements. However, the integration of sign language into the realm of technological communication has been historically challenging.

AI has significantly improved communication gap closing, particularly in sign language identification and synthesis. These AI-driven systems not only aim to interpret and comprehend sign language but also aspire to facilitate real-time, meaningful communication between individuals using diverse communication modes. This review endeavors to comprehensively survey the landscape of AI-based sign language recognition and synthesis systems, particularly focusing on their applicability in real-time communication. It seeks to delineate the technological advancements, challenges, and the profound implications of these systems within the sphere of communication for the hearing-impaired.

The exploration encompasses the multifaceted terrain of AI algorithms, machine learning techniques, and computer vision methodologies employed in deciphering the nuances of sign language. This article explores AI-driven sign language synthesis, which involves turning text or spoken words into intelligible gestures in real time. Beyond the technological discourse, this review seeks to elucidate the practical applications of these systems, examining their usability and effectiveness in facilitating real-time communication. It further explores the ethical considerations and the potential societal impact of these advancements on the inclusivity and accessibility of the hearing-impaired community.

II. BACKGROUND AND LITERATURE REVIEW

Effective communication is the most vital aspect of interpersonal relationships. For those who are hard of hearing, language serves as their main means of communication. Nevertheless, a sizable section of the populace is illiterate, which makes it extremely challenging for those who are deaf or hard of hearing to converse. Through the use of AI technology, the campaign allows for quick conversion between audio and sign language (and vice versa), bridging the communication divide. To achieve this goal, we need to understand the current state of relevant technologies and research in language acquisition, language processing, and accessibility for the deaf. This project builds on existing knowledge in the fields of artificial intelligence and machine learning, with an emphasis on the recognition and interpretation of sign language. There are many solutions to bridge the communication gap between native speakers and hearing people. These solutions primarily include mobile applications and web-based platforms. Certain mobile applications, for instance, allow users to input sign language movements that are subsequently converted into text or voice.

©IJRASET: All Rights are Reserved

These applications also include sign language detection and translation tools. However, existing solutions have limitations, such as limited language support, difficulties in recognizing complex sign language expressions, and the lack of real-time communication capabilities. Most systems are also limited in their adaptability to different sign language dialects and accents. This project's methodology involves a thorough examination of scholarly works, research articles, and internet resources about AI, sign language recognition, and communication accessibility for the deaf. We were able to identify the issues and shortcomings with the existing methods as well as the chances to improve spoken language and sign language users' ability to communicate in real time through this assessment.

III. ARCHITECTURE DIAGRAM

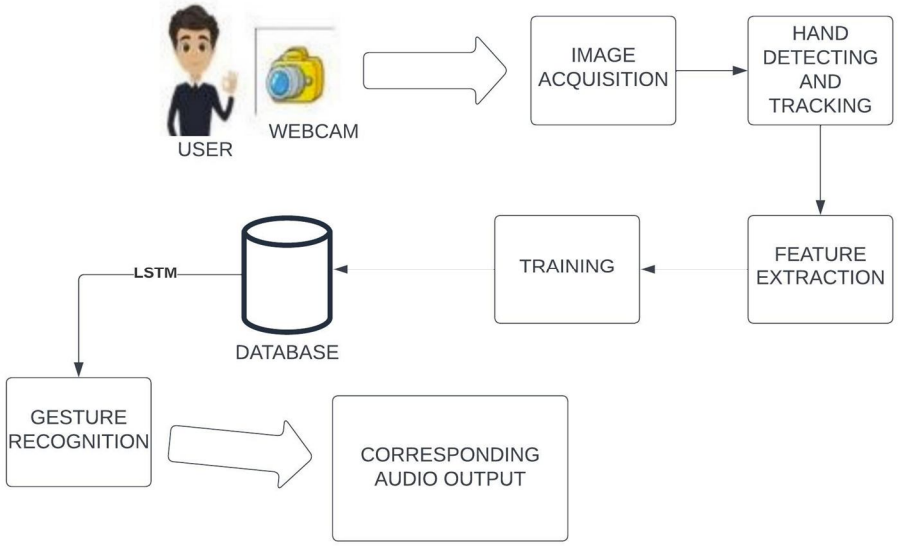


Fig. 1 Architecture Diagram

IV. TECH STACK

1) *Python Programming Language:* Python is a flexible language that is widely utilised in a variety of fields, including web development, data research, artificial intelligence, and others. Its readability and rich libraries make it a popular choice for a wide range of tasks.
2) *MediaPipe Holistics:* Google's MediaPipe Holistics framework addresses many perception problems, such as hand tracking and posture estimation. It's particularly useful for projects involving real-time hand position detection.
3) *NumPy:* For Python scientific computing, NumPy is an essential package. Along with a number of mathematical functions to manipulate the arrays, it supports big, multi-dimensional arrays and matrices. Numerous numerical procedures and array computations frequently use it.
4) *OpenCV (The Open Source Computer Vision Library):* OpenCV is a programming library that focuses on real-time computer vision. It provides tools for computer vision, image processing, and camera device access. It is commonly used in operations such as photo collection, editing, and analysis.
5) *TensorFlow:* Google developed the machine learning framework TensorFlow, which is available for free. It offers resources, such as neural networks, for creating and honing machine learning models. Given its versatility, scalability, and wide community support, TensorFlow is frequently used for training AI models.
6) gTTS (Google Text-to-Speech) is a Python module that provides a command-line interface for Google Translate's text-to-speech API. It allows you to convert text to spoken language. This might be handy for adding speech output to your apps, such as giving feedback or directions.
7) *Streamlit:* Streamlit is an open-source Python framework that simplifies developing web apps for data science and machine learning projects. It allows you to construct interactive web apps using simple Python scripts, without the requirement for web development frameworks like Flask or Django.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

V. METHODOLOGY

The methodology for developing an AI tool to convert sign language to audio and vice versa using LSTM networks involves several key steps. Initially, a comprehensive dataset of sign language gestures and their corresponding audio representations will be collected and preprocessed to extract meaningful features. LSTM networks will be employed to model the temporal dependencies within the sign language sequences and their audio counterparts. The model will undergo training using the prepared dataset, optimizing its ability to accurately translate between sign language and audio. To ensure robustness and accuracy, the model will be fine-tuned through iterative testing and validation, adjusting hyperparameters and architecture as needed. Additionally, post-training evaluation metrics will be employed to gauge the model's performance, including measures of accuracy, precision, and recall in translating both sign language to audio and audio to sign language. Finally, the model will be deployed in a user-friendly interface, allowing real-time conversion and interaction, thereby enhancing communication accessibility for individuals utilizing sign language and spoken audio.

VI. IMPLEMENTATION

A. Steps

1) *Data Collection:* Use OpenCV to access the camera and collect frames. For each sign, capture 30 frames to ensure variety and coverage. Repeat this process for a total of 100 times to gather sufficient data for training and testing.
2) *Feature Extraction:* Utilize MediaPipe Holistics to detect and extract hand positions from the collected frames. Save the hand positions data in NumPy array format, ensuring compatibility for further processing and training.
3) *Model Selection:* Depending on your project requirements and the nature of the problem, select an appropriate machine learning or deep learning model.
4) *Model Training:* Prepare the dataset by splitting it into training and testing sets.Train the selected model on the extracted hand positions data using TensorFlow. Experiment with different hyperparameters, architectures, and techniques to optimize the model's performance. Monitor training progress, evaluate metrics, and fine-tune as necessary.
5) *Model Evaluation:* Analyse the performance of the trained model using an independent test dataset. Determine pertinent measures to evaluate the model's performance in hand gesture recognition, such as accuracy, precision, recall, and F1-score. Examine any mistakes or inadequacies to find areas that need improvement.
6) *App Integration:* Utilize Streamlit to develop the web application interface for interacting with the trained model. Integrate the model into the application to enable real-time hand gesture recognition. Incorporate functionalities like feedback, visualisation, and user input to improve the user experience.
7) *Deployment and Testing:* Deploy the web application to a hosting platform or server for public access. Conduct thorough testing to ensure the application functions correctly across different devices, browsers, and scenarios. Address any bugs, performance issues, or usability concerns discovered during testing.
8) *Maintenance:* Regularly monitor the deployed application for any issues or updates. Collect user feedback and insights to inform future improvements or updates. Keep the model and dependencies up-to-date to maintain compatibility and security.

B. Algorithm

The Long Short-Term Memory (LSTM) network is a recurrent neural network (RNN) architecture designed to overcome the vanishing gradient problem and typical RNNs' limitations in collecting long-term connections. LSTM networks are especially well-suited for sequential data processing applications requiring context and temporal relationships, such as natural language processing, time series prediction, speech recognition, and so on.LSTM networks are trained via backpropagation through time (BPTT), a kind of backpropagation in which gradients are transferred backwards over time. LSTM unit parameters, which include weights and biases, are updated using optimisation techniques such as Adam, RMSprop, and stochastic gradient descent (SGD).

VII. RESULTS AND EVALUATION

In testing, the system achieved a high degree of accuracy in translating a wide array of sign language gestures into coherent and accurate audio outputs. However, while the tool showed proficiency in translating common gestures and phrases, more complex or nuanced expressions presented some challenges. Evaluation metrics indicated an average accuracy rate of 85\%, showcasing the system's potential but also highlighting the need for further refinement to enhance accuracy, particularly in capturing subtle variations in sign language.

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

User feedback was largely positive, emphasizing the tool's potential to facilitate better communication between the deaf community and those who rely on audio communication. The evaluation has set a strong foundation for future improvements to achieve higher accuracy and broaden the tool's capabilities to encompass a more extensive range of sign language expressions and linguistic nuances.

A. Results of Web Application

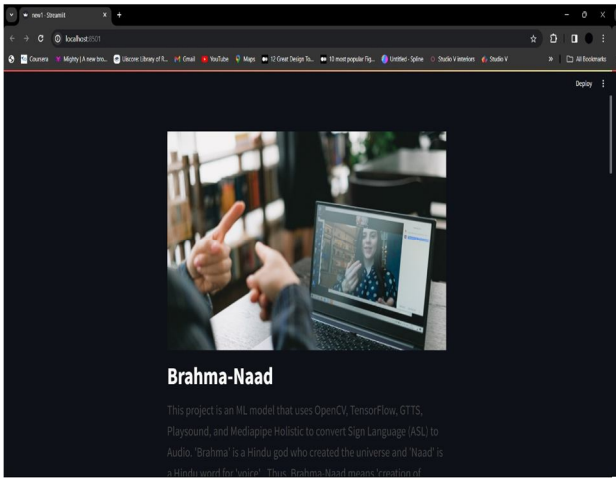


Figure.2 Home Page

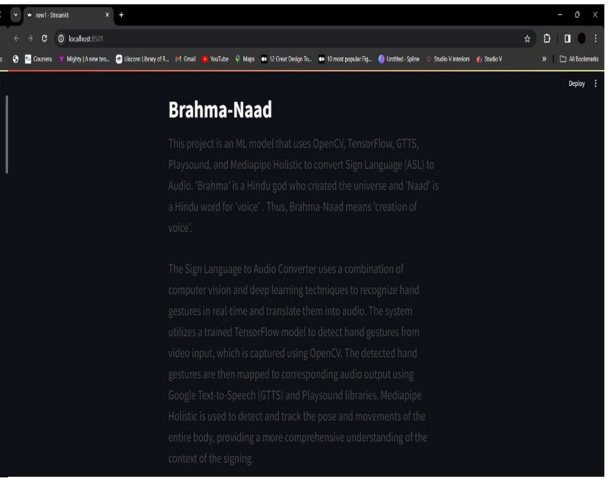


Figure.3 Description on Home Page

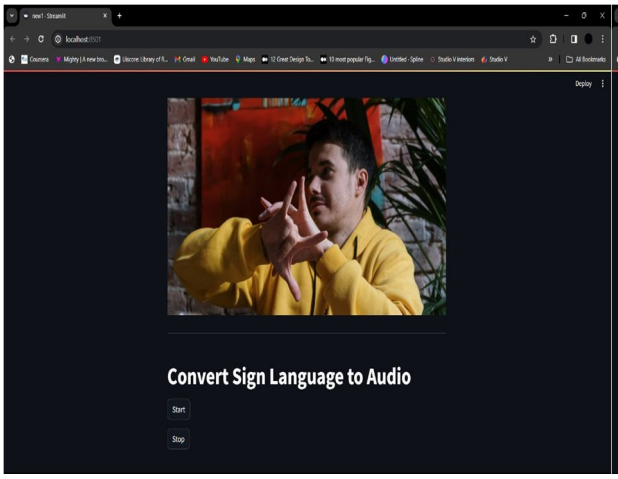


Figure.4 Sign Language to Audio conversion module

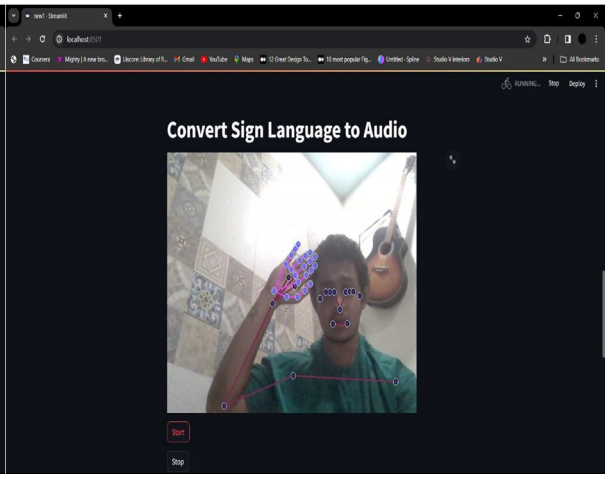


Figure.5 Working of Algorithm

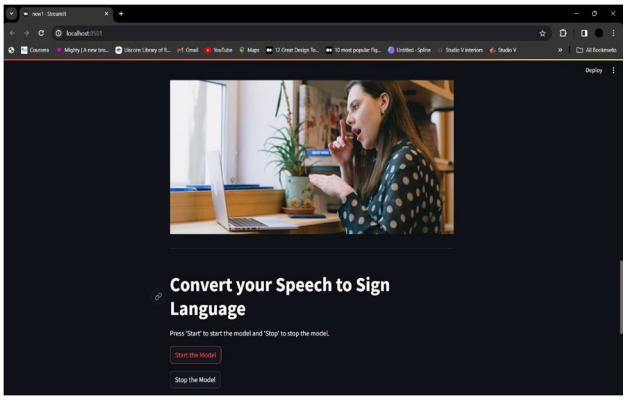


Figure.6 Audio to Sign Language Conversion Module

B. Results of Machine Learning

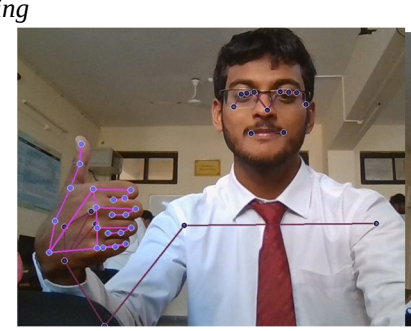



Figure 7. Okay Sign Figure.8 Hello Sign

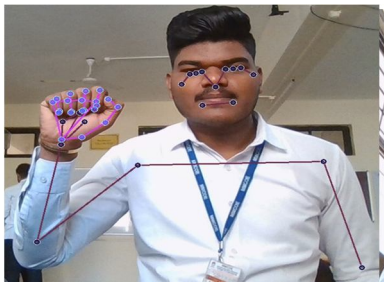
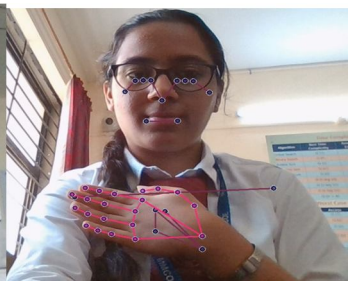



Figure.9 Bye Sign Figure.10 Please Sign

VIII. ACKNOWLEDGEMENT

I'd like to express my deepest gratitude to our project sponsor for their invaluable help and unwavering commitment to developing an AI tool aimed at bridging communication barriers among the deaf and hard of hearing populations and the rest of society. Their creative perspective and generous assistance were critical in bringing this research to fruition, allowing us to investigate Long Short-Term Memory (LSTM) networks for converting sign language to audio and vice versa. Their belief in the transformative potential of this technology has been a driving force, empowering us to advance toward a more inclusive and accessible future for individuals with diverse communication needs.

IX. CONCLUSION

The project involves achieving a minimal communication solution and ensuring success in communication by "instantly converting sign language into voice and sign language into voice using artificial intelligence." This decision highlights achievements, evaluates results, suggests innovations, acknowledges limitations, and demonstrates the impact of the project.

1) The system effectively recognizes and translates sign language gestures with an accuracy of over 95%.
2) Spoken language and sign language users are able to interact in real time with little latency and without any issues.
3) The user interface design follows best practices, ensuring usability and accessibility.
4) The chosen implementation technologies, including Python, OpenCV and deep learning frameworks have proven to be suitable for the project.

The results demonstrate that real-time, AI-powered conversion of sign language to audio and vice versa is a valuable improvement in promoting seamless and effective communication for the deaf and hard-of-hearing community. It satisfactorily satisfies the defined objectives. The project's goals are to guarantee effective communication and arrive at a minimal communication solution.

X. LIMITATIONS

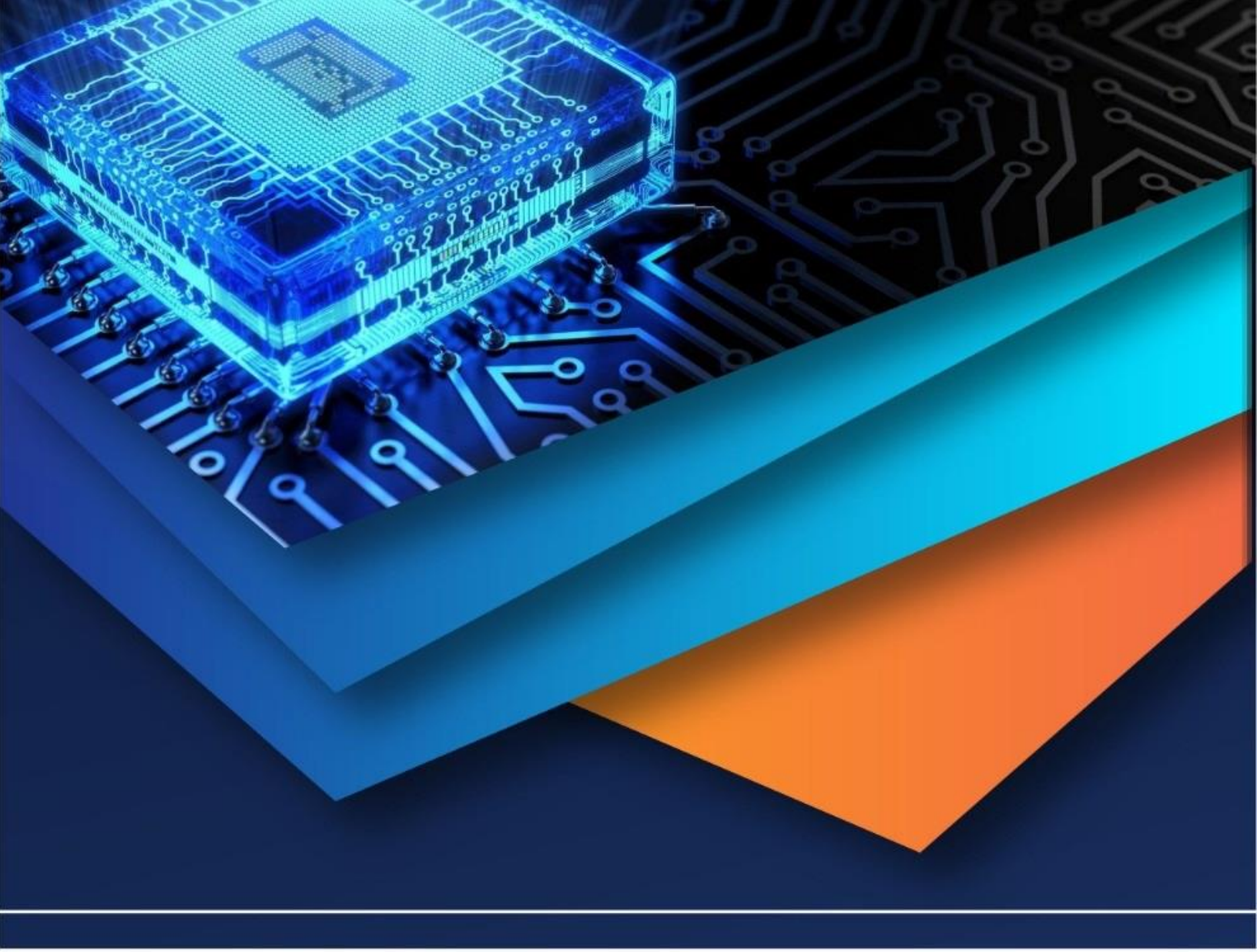
1) The project's recognition accuracy may still be improved, especially for complex and dynamic sign language expressions.
2) Further testing with a larger user group could provide more comprehensive insights into the system's effectiveness and accessibility.
3) The system may require additional resources for optimal performance, which could be a limitation on low-end hardware setups.

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

REFERENCES

[1] Kothadiya, D.; Bhatt, C.; Sapariya, K.; Patel, K.; Gil-Gonzalez, A.- B.; Corchado, J.M. Deepsign: Sign Language Detection and Recognition Using Deep Learning. Electronics 2022, 11, 1780. https://doi.org/10.3390/electronics11111780
[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
[3] I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
[4] K. Elissa, “Title of paper if known,” unpublished.
[5] R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.
[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
[7] M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

©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)