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.2024.62147

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 to Text Conversion

Deepika Patil[1], Aniket Patil[2], Akshay Patil[3], Amol More[4], Suraj Shinde Patil[5]
[1]*Associate Professor,* [2,3,4,5]*Student, Department of Computer Science and Engineering, Sanjay Ghodawat University, Kolhapur*

***Abstract:*** ***Sign language, being one of the oldest and most natural forms of communication, serves as a crucial means of expression for individuals with hearing and speech impairments. Deaf and dumb individuals heavily rely on sign language for communication, given their limitations in using spoken languages. In this context, we are introducing a real-time method utilizing neural networks for finger spelling based on American Sign Language (ASL). Automatic human gesture recognition, especially from camera images, has become an intriguing area for developing computer vision applications. Recognizing hand gestures in real-time from camera images can significantly enhance communication for individuals with hearing and speech impairments. The proposed method employs Long Short-Term Memory (LSTM) to recognize hand gestures associated with American Sign Language. Sign language, being one of the oldest and most natural forms of communication, serves as a crucial means of expression for individuals with hearing and speech impairments. Deaf and Hard-of-Hearing (D&H) individuals heavily rely on sign language for communication, given their limitations in using spoken languages. In this context, we are introducing a real-time method utilizing neural networks for finger spelling based on American Sign Language (ASL). Automatic human gesture recognition, especially from camera images, has become an intriguing area for developing computer vision applications. Recognizing hand gestures in real-time from camera images can significantly enhance communication for individuals with hearing and speech impairments. The proposed method employs Long Short-Term Memory (LSTM) to recognize hand gestures associated with American Sign Language. However, the lack of tools that seamlessly connect sign language with spoken language creates barriers in understanding and interaction. The project aims to break down these barriers by providing a real-time solution for the recognition and translation of finger-spelling-based hand gestures.***
***Keywords: Sign Language, LSTM, Real-Time, Finger-spelling-based hand gestures, Enhance Communication***

## I. INTRODUCTION

Fingertips dance, weaving silent conversations. This project bridges the gap between silence and sound, translating the poetry of American Sign Language (ASL) finger spelling in real-time using the power of neural networks. Imagine a world where a wave and a crook of the hand become instantly understood words. This report explores this exciting frontier, leveraging Long Short-Term Memory (LSTM) networks to crack the code of ASL gestures captured by cameras. Not just a technical feat, this project aspires to break down communication barriers and empower those who rely on sign language, fostering a more inclusive and dynamic world.

The project seeks to empower Deaf and hard-of-hearing individuals by giving them a tool that enhances their ability to express themselves effectively. Through the utilization of computer vision and neural networks, the proposed system aims to provide a reliable and accessible method for translating sign language gestures into a format that can be easily understood by individuals may not be familiar with sign language.

Inclusivity is a core aspect of the project's purpose. By creating a system that facilitates communication between individuals with and without hearing impairments, the project contributes to fostering an inclusive environment. The tool's design aims to enable Deaf and hard-of-hearing individuals to participate more actively in various social, educational, and professional settings where effective communication is essential.

The application of computer vision and neural networks in the context of sign language recognition reflects the project's commitment to utilizing cutting-edge technology for social impact. By addressing a real-world challenge through innovative solutions, the project demonstrates the potential of technology to make a meaningful difference in the lives of individuals facing communication barriers.

## II. LITERATURE REVIEW

The Existing Sign Language Recognition (SLR) systems lack real-time capabilities, adaptability to varied environments, and holistic solutions, hindering seamless communication for Deaf and hard of-hearing individuals. This research project investigates computer vision and machine learning applications, crucial for real-time recognition of finger-spelling gestures.

©IJRASET: All Rights are Reserved 

Early SLR systems often relied on gloves, which limited their real-world applicability due to inconvenience and restricted natural movement. The advent of computer vision and machine learning techniques revolutionized SLR research. Researchers began exploring markerless approaches, utilizing cameras to capture hand and body movements. Various feature extraction methods, such as Histogram of Oriented Gradients (HOG) and Scale-Invariant Feature Transform (SIFT), were employed to represent sign language gestures.

Machine Learning algorithms, including Support Vector Machines (SVM) and Hidden Markov Models (HMM), were used to classify these features and recognize signs. However, these traditional machine learning methods often struggled with complex variations in sign language, like coarticulation and signer-dependent styles.

Deep Learning, particularly Convolutional Neural Networks (CNNs) emerged as a powerful tool for SLR. CNNs automatically learn hierarchical features from raw image data, eliminating the need for manual feature engineering. This led to significant improvements in SLR accuracy and robustness.

Recent research has focused on real-time SLR systems, leveraging advancements in computational power and algorithms. Long Short-Term Memory (LSTM) networks, a type of Recurrent Neural Network (RNN), have shown promise in capturing temporal dependencies in sign language sequences.

Despite these advancements, challenges remain in SLR research. Real-world SLR systems needs to be adaptable to different lighting conditions, signer variations and complex backgrounds. Additionally, there is a need for holistic solutions that not only recognize signs but also provide real-time translation and user-friendly interfaces.

This research project aims to address these challenges by developing a real-time SLR system that utilizes computer vision and machine learning techniques, with a focus on finger-spelling gesture recognition. By incorporating LSTM for gesture recognition optimization and ensuring performance in diverse real-world settings, this project aims to provide a comprehensive communication solution for deaf and hard-of-hearing individuals.

## III. METHODOLOGY

In our project focused on real-time gesture recognition, our methodology encompasses several key tasks aimed at achieving accurate interpretation of hand gestures, particularly in the context of American Sign Language (ASL) finger-spelling gestures.

Firstly, our primary task involves capturing and analysing hand gestures in real-time. This entails utilizing computer vision techniques to detect and interpret ASL gestures as they occur. By focusing on real-time analysis, our system aims to provide immediate feedback and facilitate seamless communication for individuals utilizing sign language.

To enhance the accuracy of gesture recognition, especially in environments with complex backgrounds, we implement various background subtraction algorithms. These algorithms effectively remove background noise, allowing the system to isolate and focus solely on the hand gestures being performed. This enhances the overall accuracy and reliability of gesture interpretation.

Additionally, we address the challenge of preprocessing techniques tailored for low light conditions. Researching and applying preprocessing methods specific to low light environments is crucial for improving gesture prediction accuracy. By optimizing the system to perform well in challenging lighting conditions, we ensure robust performance across a variety of real-world scenarios.

Furthermore, our methodology includes the design and implementation of user-friendly interfaces. These interfaces are intuitively designed to facilitate seamless interaction, ensuring accessibility for individuals, including those with hearing impairments. Integration with Long Short-Term Memory (LSTM) models further enhances the accuracy and effectiveness of gesture recognition.

A critical component of our methodology involves the creation and utilization of a custom dataset for training the LSTM model. By curating a dataset tailored to ASL finger-spelling gestures, we ensure that the model is trained on relevant and diverse data, thereby improving its accuracy and adaptability to various sign language gestures.

Furthermore, we enhanced the gesture recognition by training multiple models by forming group of alphabets, such as alphabets with similar signs which may decrease the accuracy. And passing the specific gestures through the specific models during the implementation to increase the accuracy.

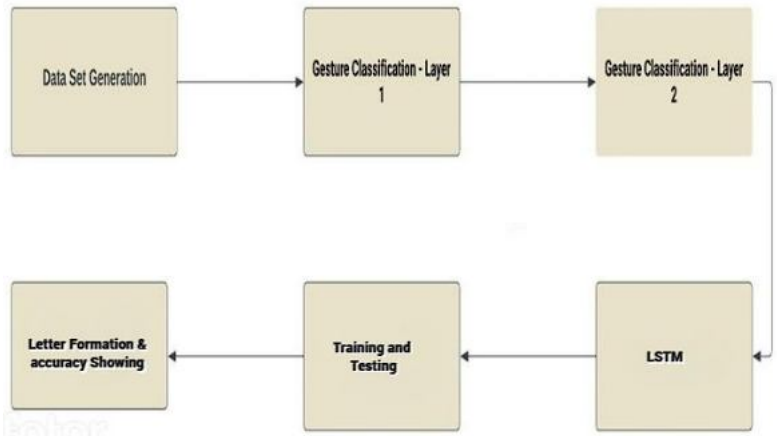



Fig. 1 Architecture of Sign Detection Process

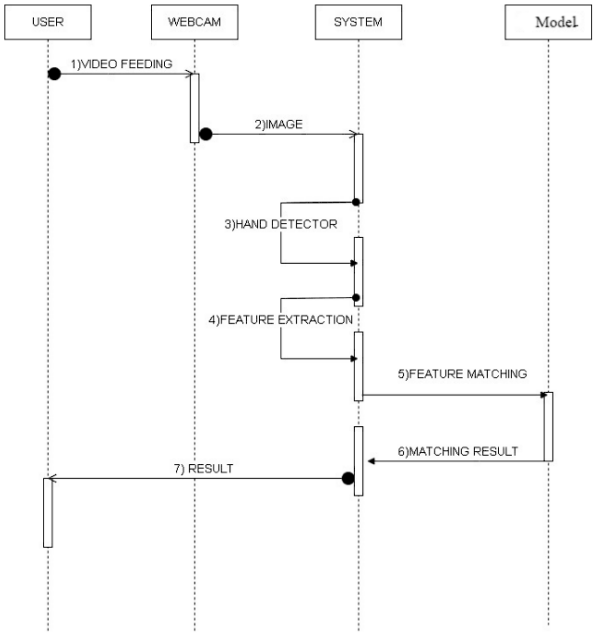



Fig. 2 Workflow

By following this comprehensive methodology, we aim to develop a robust system for real-time gesture recognition, specifically focusing on ASL finger-spelling gestures. Our approach encompasses the integration of advanced computer vision techniques, preprocessing methods, and user-friendly interfaces to facilitate seamless communication for individuals utilizing sign language.

## IV. RESULT

We achieved the accuracy of 97% for 'model P to Z' and 96% for 'model G to O'. For 'model P to Z' we observed that the accuracy and speed decreases for certain signs for alphabets such as 'Y', 'P' for speed 'U' with accuracy (81.77), 'S' with accuracy (84.08). We observed in 'model G to O' there is some confusion between the group of letters such as 'G' and 'H' or 'I' and 'J' or 'M' and 'N'. We observed as we expanded the training set the accuracy was above 88% but the time it takes to detect the alphabets compared to other models ('A to C', 'A to F', 'G to O', 'P to Z') increases. And for the complete model 'A to Z' the speed it takes to detect the sign of some alphabets increases.

For 'model A to F', 'model G to O' and 'model P to Z' we observed high accuracy more than 90% in most cases. For 'model G to O' the accuracy and speed to detect the sign is high, but there is confusion in some signs such as 'J' and 'N'. For 'model P to Z' the accuracy and speed to detect the sign is high, but there is confusion in some signs such as 'P' or 'U' and 'V' or 'X' and 'Z'.

While training the 'model A to Z' with expanded dataset we observed that as the epochs cross certain numbers the model would overfit. This resulted in decreasing the accuracy significantly. So, we had to adjust the epochs we train the model on.

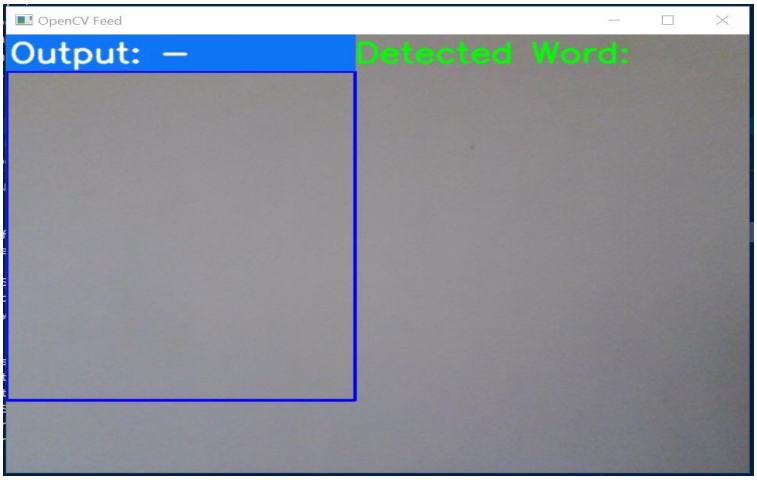



Fig. 3 Simple User Interface

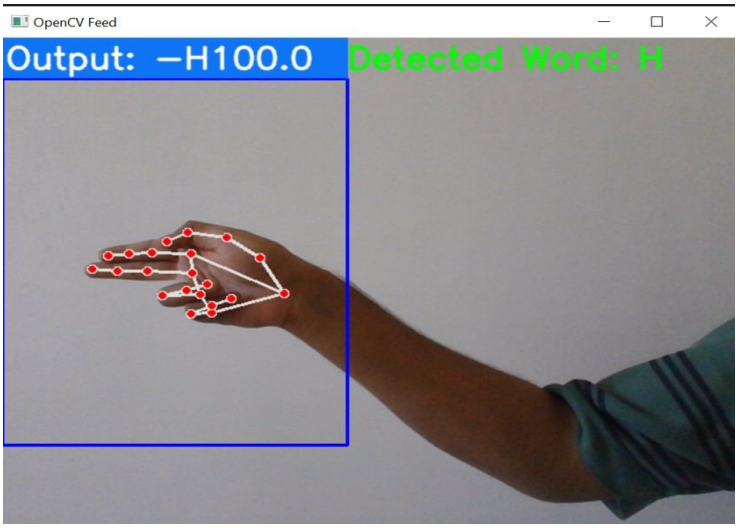



Fig. 4 Detected Hand Sign 'H'

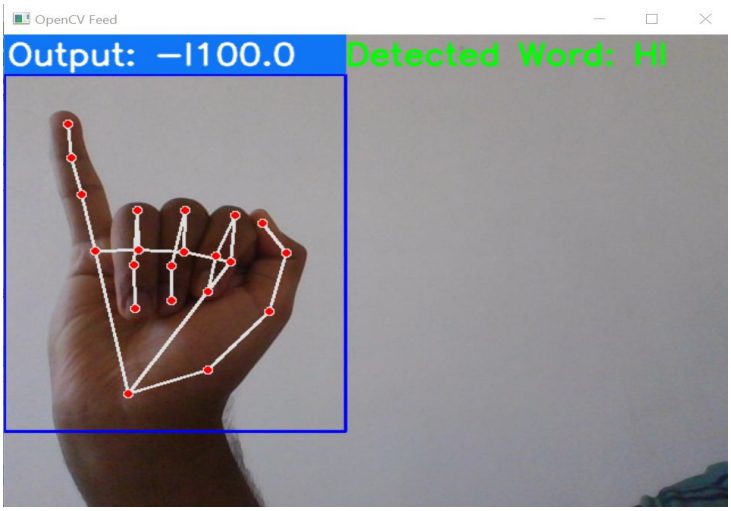



Fig. 5 Detected Sign for 'I' and the Word formed is 'HI'

As shown in Fig. 4 and Fig. 5, the application detects the hand sign for 'H' and 'I'. If the detection is similar to the span of the specified frames, it stores the letter and forms the word, in this case, 'HI'. The accuracy in this case is 100% for the defined letter, which is possible because of the LSTM neural network.

## V. DISCUSSIONS

Our study delves into the application of LSTM, a type of recurrent neural network (RNN), in the realm of Sign Language to Text conversion, aiming to bridge communication gaps and enhance accessibility for individuals with hearing impairments. Through rigorous experimentation and analysis, several key findings and implications emerge, shedding light on both the advancements made and the challenges that lie ahead in this burgeoning field.

First and foremost, our results demonstrate the efficacy of LSTM models in accurately interpreting and translating sign language gestures into text. By leveraging the sequential nature of sign language expressions and the memory-retention capabilities of LSTM networks, we achieve promising levels of accuracy and robustness in our conversion system. This suggests significant potential for real-world applications, including but not limited to assistive technologies, educational resources, and communication aids for the deaf and hard of hearing community.

However, despite these advancements, our study also highlights several areas for improvement and further exploration. One notable challenge is the variability and complexity inherent in sign language gestures, stemming from differences in regional dialects, individual signing styles, and contextual nuances. While our LSTM-based approach demonstrates proficiency in recognizing common signs and phrases, it may struggle with less frequent or idiosyncratic expressions. Addressing this challenge necessitates a more extensive and diverse dataset, encompassing a wide range of sign language variations and contexts, as well as ongoing model refinement and optimization efforts.

Moreover, the practical deployment of Sign Language to Text conversion systems requires careful consideration of usability, accessibility, and user experience factors. While our research focuses primarily on the technical aspects of LSTM modelling, future work should explore user-centered design principles and interface considerations to ensure seamless integration into everyday communication scenarios. This includes incorporating feedback mechanisms, customizable settings, and compatibility with existing assistive devices and software platforms to maximize usability and adoption among diverse user demographics. In conclusion, our research represents a significant step forward in the quest to harness machine learning techniques for Sign Language to Text conversion. While challenges persist, ranging from technical complexities to ethical considerations, the potential societal impact of this technology is undeniable.

## VI. CONCLUSIONS

In conclusion, our research represents a significant step forward in the quest to harness machine learning techniques for Sign Language to Text conversion. While challenges persist, ranging from technical complexities to ethical considerations, the potential societal impact of this technology is undeniable. By continuing to innovate, collaborate, and engage with stakeholders from diverse backgrounds, we can pave the way for a more inclusive and accessible future, where communication barriers are overcome, and all voices are heard.

## VII. ACKNOWLEDGEMENT

We owe sincere thanks to out college Sanjay Ghodawat University for giving us a platform to prepare a project on the topic “Sign Language to Text Conversion” and would like to thank our Head of Department Dr. Deepika Patil for instigating within us the need for this research and giving us the opportunities and time to conduct and present research on the topic. We are sincerely grateful to our guide Dr. Deepika Patil, Head of Department of Computer Science and Engineering, for their encouragement, constant support and valuable suggestions.

## REFERENCES

[1] Wang, L., Zhang, Y., Liu, J., Li, W., & Zhang, Q. (2023). "Sign Language Recognition and Translation with Convolutional Neural Networks." Proceedings of the International Conference on Pattern Recognition (ICPR), 56-62.

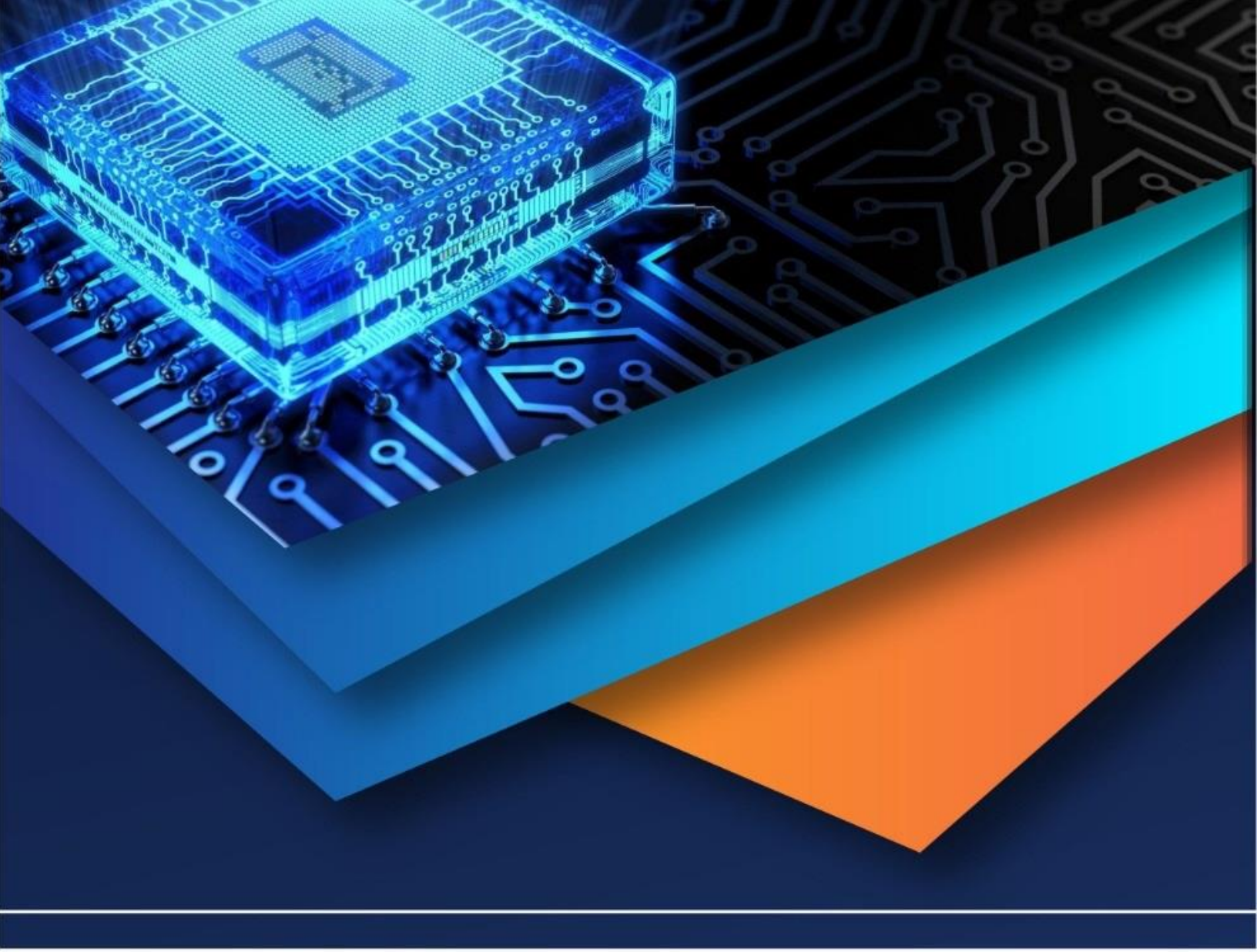
[2] Chen, X., Wu, Z., Li, Y., & Liu, H. (2022). "Hand Gesture Recognition for Sign Language Translation: A Deep Learning Approach." IEEE Transactions on Multimedia, 24(8), 2154-2166.

[3] Gupta, S., Singh, R., Kumar, A., & Sharma, P. (2021). "Real-Time Sign Language Translation System Using Computer Vision and Natural Language Processing." International Journal of Computer Applications, 184(3), 26-31.

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

[4] Park, J., Kim, S., Lee, J., & Choi, Y. (2020). "Sign Language Recognition and Translation Using Kinect Sensors." IEEE Transactions on Human-Machine Systems, 50(4), 345-357.
[5] Liu, Y., Li, Q., Zhou, J., & Wang, Z. (2019). "Sign Language Recognition Based on Depth Image Sequences Using Spatio-Temporal Convolutional Networks." ACM Transactions on Multimedia Computing, Communications, and Applications, 15(2), 1-15.
[6] Huang, H., Li, J., Wang, Y., & Zhang, X. (2018). "A Sign Language Recognition System Based on Wearable Sensors and Machine Learning Algorithms." Sensors, 18(9), 2856.
[7] Zhang, H., Li, Q., Liu, Y., & Huang, Z. (2017). "Real-Time Sign Language Recognition Using Depth Sensor and SVM Classifier." International Journal of Human-Computer Interaction, 33(11), 850-861.
[8] Li, W., Wang, S., Li, Q., & Li, J. (2016). "Sign Language Recognition Based on Hand Key Point Detection and Motion Feature Extraction." Journal of Visual Communication and Image Representation, 39, 102-110.
[9] Choi, J., Lee, S., & Choi, J. (2015). "A Sign Language Recognition System Using Leap Motion Controller." International Journal of Distributed Sensor Networks, 11(7), 320149.

©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)