iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.51981

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 11 Issue V May 2023- Available at www.ijraset.com

Conversion of Sign Language to Text

Akash Kamble[1], Jitendra Musale[2], Rahul Chalavade[3], Rahul Dalvi[4], Shrikar Shriyal[5]
[1, 3, 4, 5]*UG Student,* [2]*Professor, Department of Computer Engineering, Savitribai Phule Pune University, India*

Abstract: ***Sign language is a form of communication that uses hand sign and gestures to convey meaning. we present a new approach to converting sign language into text format. Our system is designed to enable deaf and mute people to communicate with others in a more accessible and convenient way. The proposed method uses computer vision and deep learning methods to recognize hand gestures and translate them into appropriate text. The system was built using a combination of key point detection using MediaPipe, data pre-processing, label, feature generation and LSTM neural network training. This work has the potential to significantly improve communication for deaf and dumb individuals and reduce the barriers to communication with the rest of the world. The system uses key point detection algorithms such as MediaPipe to identify hand gestures and a Lstm model to translate them into corresponding text output. The data collected from the sign language is pre-processed and then used to train an LSTM neural network to accurately recognize gestures and produce text output. This method of conversion not only helps deaf and hard of hearing individuals communicate with the hearing population, but also serves as an assistive tool for individuals who are trying to learn sign language. Overall, the proposed solution has the potential to greatly improve communication and reduce barriers for deaf and hard of hearing individuals***

Keywords: Hand Gesture, Media Pipe, RNN, machine learning, TensorFlow, OpenCV, Python, etc.

I. INTRODUCTION

This Effective communication is essential in all aspects of life, and it is especially important for individuals who are deaf or hard of hearing. With the rising number of people suffering from hearing loss, it is crucial to find ways to bridge the communication gap between the hearing and non-hearing population. To address this issue, we present a new system for converting Sign Language into text format using computer vision and machine learning techniques. This system aims to provide an efficient and accessible solution for deaf and hard of hearing individuals to communicate with the hearing population.In the today's world, Communication is always having a great impact in every domain and how it is considered the meaning of thoughts and expressions that attract the researchers to bridge this gap for normal and deaf people. According to World Health Organization, by 2050 nearly 2.5 billion people are projected to have some degree of hearing loss and at least 700 million will require hearing rehabilitation. Over 1 billion young adults are at the risk of permanent, avoidable hearing loss due to unsafe listening practices. Sign languages vary among regions and countries, with Indian, Chinese, American, and Arabic being some of the major sign languages in use today. This system focuses on Indian Sign Language and utilizes the Media Pipe Holistic Key points for hand gesture recognition. The system uses an action detection model powered by LSTM layers to build a sign language model and predict the Indian Sign Language in real-time. The use of cutting-edge technologies and efficient algorithms makes this system a valuable tool for improving communication between deaf and hard of hearing individuals and the rest of the world. It is difficult to finding a sign language translator for converting sign language every time and everywhere, but electronic devices interaction system for this can be installed anywhere is possible. Computer vision is one of the emerging frameworks in object detection and is widely used in various aspects of research in artificial intelligence. Sign language is categorized in accordance with regions like Indian, Chinese, American and Arabic. This system introduces efficient and fast techniques for identifying the hand gestures representing sign language meaning. In this system we will extract the Media Pipe Holistic Key points, then build a sign language model using an Action detection powered by LSTM layers. Then Predict Indian sign language in real time.

II. LITERATURE REVIEW

1) The paper proposes the importance of sign language as a natural and expressive way for hearing-impaired people to communicate. However, most people who are not deaf do not try to learn sign language, leading to isolation for the deaf community. By developing a system that can translate sign language to text, the difference between normal people and the deaf community can be minimized. It able to recognize various alphabets of ISL accurately, reducing noise. This system provides an opportunity for deaf-dumb people to communicate with non-signing people without the need for an interpreter. It also compares the finger-spelling systems used in ASL and BSL. ASL uses a one-handed finger-spelling system, while BSL uses a two-handed finger-spelling system. The paper notes that many BSL signs are derived from their initialized (English) base, while many ASL signs have been developed without initialization, revealing a cultural value.

©IJRASET: All Rights are Reserved

2) The paper discusses various techniques that have been used for converting sign language into text/speech format and compares their performance. Based on the analysis, the authors select the most effective method and develop an Android application that can convert real-time American Sign Language (ASL) signs into text/speech. This application could potentially facilitate communication between people who use ASL and those who do not, making it easier for them to interact in real-time.
3) The proposed work aims to assist individuals with hearing, speech, or visual impairments by providing a platform for communication with others. The system uses convolutional neural networks (CNN) to recognize hand gestures in American Sign Language (ASL) and convert them into text or speech output. The system offers a high accuracy rate of 88% in identifying the hand gestures. This system offers a user-friendly interface that allows special individuals to communicate more effectively with others who may not be familiar with ASL. This work provides a practical solution to the communication challenges faced by individuals with hearing, speech, or visual impairments, thus promoting inclusivity and accessibility.
4) The system presents a method for recognizing American sign language alphabet and numbers using saliency detection, PCA, LDA, and neural networks. This method can be used for communication with deaf people as well as for connecting with computers. The system uses standard letters in sign language for recognition. The experiments were carried out on a new standard dataset in this field, and the recognition rate of the system was 99.88% using 4-fold cross-validation method in 4 training terms on average. The results of this method show high accuracy and proper performance compared with other methods in the field. This method provides an efficient way for recognizing sign language and can be a useful tool for communication between deaf people and the hearing world.
5) The paper proposes a technique for sign language recognition using principal component analysis (PCA) to recognize static hand postures. The system captures 3 frames per second from a live video stream and compares three continuous frames to identify the frame containing the static posture of the hand. The system then matches this posture with a stored gesture database to determine its meaning. The system has been successfully tested in a real-time environment with an approximate matching rate of 90%. The proposed technique provides a fast and efficient way to recognize sign gestures from a video stream and could be useful in enabling communication with hearing impaired people.
6) This paper proposes a framework for Sign Language Recognition (SLR) based on Hidden Markov Models (HMMs). The proposed framework utilizes trajectories and hand-shape features of sign videos to translate sign language into text or speech. The authors introduce a new trajectory feature called "enhanced shape context" to capture spatio-temporal information and fetch hand regions using Kinect mapping functions, which are then described by HOG (pre-processed by PCA).
7) This paper proposes an adaptive GMM-based HMMs framework for vision-based sign language recognition (SLR) which aims to improve the recognition precision. The complexity of signs and limited data collection make SLR a challenging task. The authors discovered that the inherent latent states in HMMs are related to the number of key gestures and body poses, as well as their translation relationships. They propose adaptive HMMs and obtain the hidden state number for each sign with affinity propagation clustering. To enrich the training dataset, a data augmentation strategy is also proposed by adding Gaussian random disturbances. The experiments were conducted on a vocabulary of 370 signs and demonstrated the effectiveness of the proposed method over comparison algorithms.
8) This paper describes a system for recognizing isolated signs in video-based sign language recognition. The system focuses on the manual parameters of sign language and aims to achieve signer-dependent recognition of 262 different signs. The system uses hidden Markov modeling to represent sign language as a doubly stochastic process with an unobservable state sequence. The observations emitted by the states are represented as feature vectors extracted from video frames. The system achieves high recognition rates, up to 94%, indicating that the proposed approach is effective for recognizing isolated signs in video-based sign language recognition.
9) The proposed system uses surface Electromyography data acquired from the subject's right forearm to recognize twenty-six American Sign Language gestures in real-time. The raw surface Electromyography data is first filtered and then feature extracted to obtain useful information about the hand movements associated with each sign language gesture.
10) The paper describes a new image preprocessing and feature extraction approach for Sign Language Recognition (SLR) based on Hidden Markov Models (HMMs). The approach uses a multi-layer Neural Network to build an approximate skin model using the Cb and Cr color components of sample pixels. Gesture videos are split into image sequences and converted into the YCbCr color space. By using a multi-layer Neural Network to build an approximate skin model, this approach can accurately identify and extract the hand area in each image. By using a multi-layer Neural Network to build an approximate skin model, this approach can accurately identify and extract the hand area in each image.

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

III.METHODOLOGY

A. Data Collection

To develop the Sign Language to Text Conversion System, a large and diverse dataset of hand gestures representing Indian Sign Language is required.

This dataset is collected with the help of a webcam and the Media Pipe library. The Media Pipe library provides the tools to track the hand gestures in real-time and place key points on the user's hand. The webcam captures the hand gestures and stores them as data samples for the dataset.

The collected data is used to train and test the machine learning model, which is responsible for recognizing the hand gestures and converting them into text. To ensure the robustness and accuracy of the system, it is important to collect a diverse and representative dataset, covering a wide range of hand gestures and variations in hand movements. The data collection process is ongoing, and the dataset is continually updated to ensure that it accurately reflects the Indian Sign Language. With the help of the Media Pipe library and webcam, we can collect high-quality data samples to build a robust and accurate Sign Language to Text Conversion System.

B. Data Pre-Processing

Pre-processing the hand gestures images is an important step in the development of the Sign Language to Text Conversion System. The purpose of pre-processing the images is to prepare them for the machine learning model, making it easier for the model to recognize the hand gestures and translate them into text.

During the pre-processing step, the images of hand gestures are resized, normalized, and transformed to make them suitable for input into the machine learning model. The images are resized to a consistent size, so that the model can easily process them. The normalization step is performed to remove any inconsistencies in the lighting, background, or colour of the images, which can negatively impact the performance of the model.

In addition to resizing and normalization, the images may also undergo a transformation process, such as cropping or rotation, to ensure that the model has a consistent view of the hand gestures. This helps to reduce the variance in the data and makes it easier for the model to recognize the hand gestures.

Once the pre-processing step is complete, the images are ready to be used for training and testing the machine learning model. The pre-processed images provide the model with the information it needs to learn the relationship between the hand gestures and the corresponding text, allowing it to recognize and translate the hand gestures into text with high accuracy. With the help of pre-processing, the Sign Language to Text Conversion System becomes a powerful tool for helping deaf and dumb persons to communicate with others.

C. Labelling Text Data

Labelling the hand gestures is an important step in the development of the Sign Language to Text Conversion System. In this step, each hand gesture in the dataset is assigned a label representing the word or phrase it represents. This labelling process is crucial as it provides the machine learning model with the information it needs to recognize and translate the hand gestures into text.

The labels for the hand gestures are based on the Indian Sign Language and are created in accordance with the standard terminology and grammar used in the language. The labels are assigned by an expert in Indian Sign Language, who ensures that the labelling is consistent and accurate. The labelling process is performed manually, but with the help of computer vision techniques, it can also be automated to a certain extent. Once the hand gestures are labelled, they are ready to be used for training and testing the machine learning model.

The labelled data provides the model with the information it needs to learn the relationship between the hand gestures and the corresponding text, allowing it to recognize and translate the hand gestures into text with high accuracy. With the help of appropriate labelling, the Sign Language to Text Conversion System becomes a powerful tool for helping deaf and dumb persons to communicate with others.

D. Training and Testing

In the training phase, the model is fed the pre-processed images of hand gestures along with the corresponding text labels. The model uses this information to learn the relationship between the hand gestures and the text, updating its parameters as it processes more data.

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

The goal of the training phase is to train the model to accurately recognize and translate the hand gestures into text.

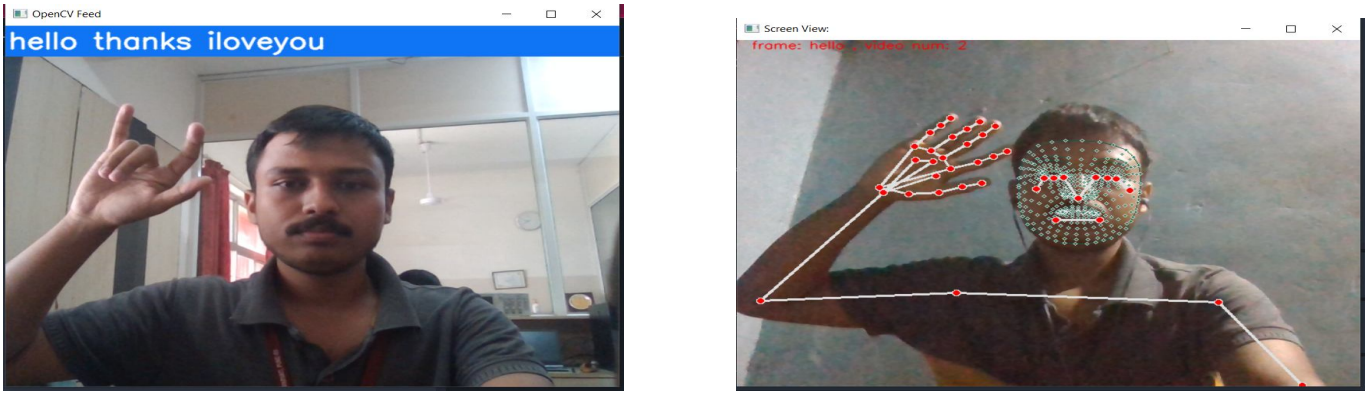


Fig. 1 Training and testing

In our project on "Conversion of Sign Language to Text," we have implemented a multi-layered LSTM (Long Short-Term Memory) model to effectively convert sign language gestures into textual representations. The LSTM model consists of three LSTM layers and three Dense layers having activation function as RELU and the output layer having activation function as SoftMax, each contributing to the understanding and interpretation of the sequential nature of sign language.

The addition of multiple LSTM layers allows for the model to learn and capture increasingly complex patterns and dependencies present in sign language gestures. Each LSTM layer in the network takes in a sequence of inputs, processes them through its memory cells, and outputs a hidden state that carries information forward to the next layer.

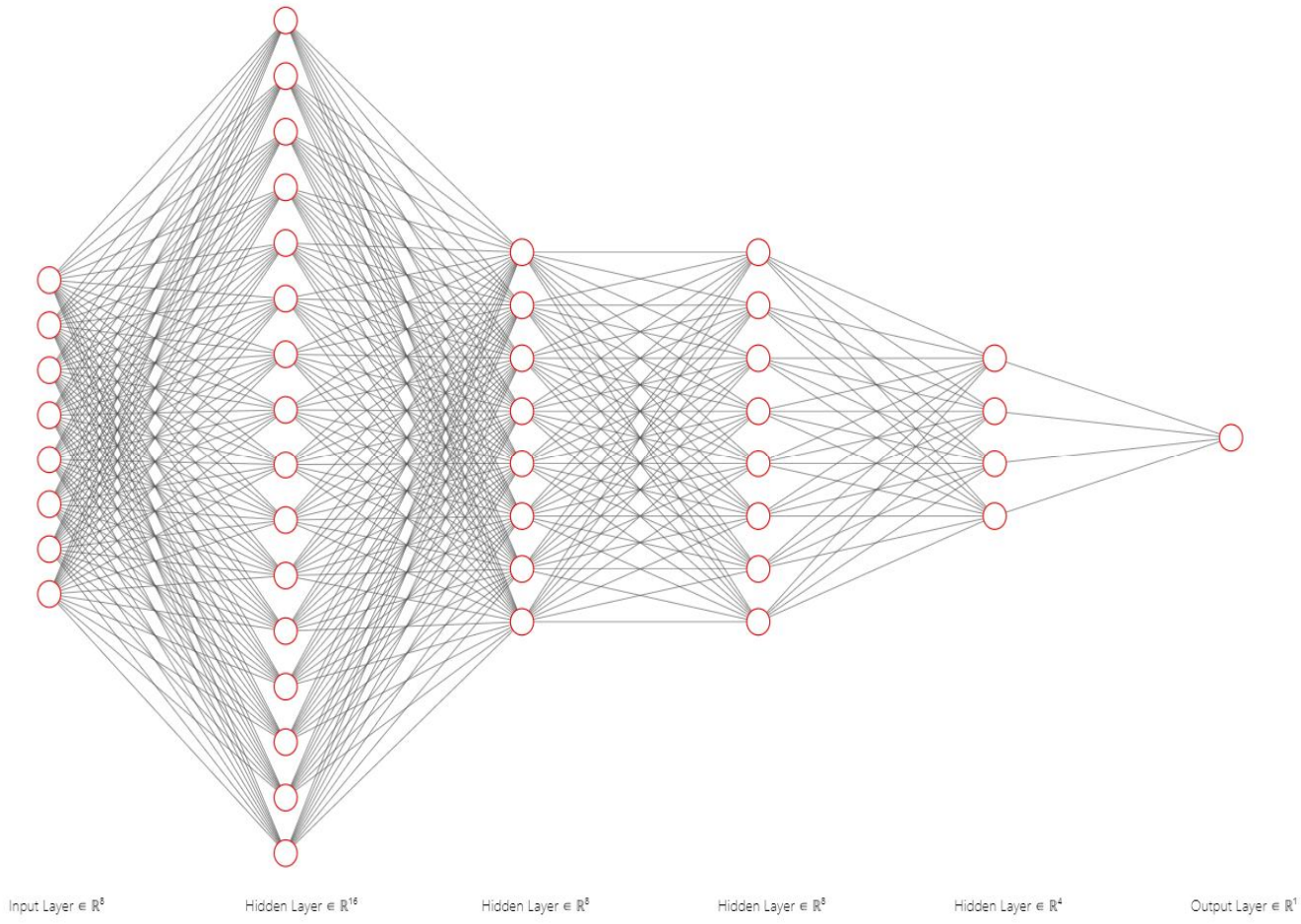


Fig. 2 layers used representation

IV. ARCHITECTURE DIAGRAM

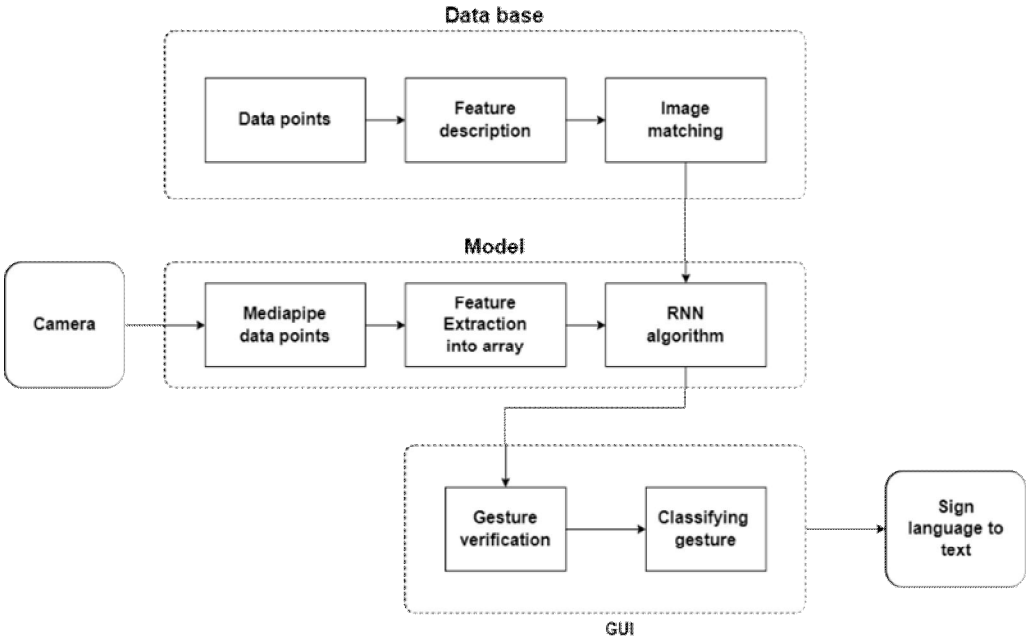


Fig. 3 Architecture Diagram

The conversion of sign language to text involves several steps and technologies, including the use of a camera, the Mediapipe library, feature extraction, data points, image matching, RNN algorithm, gesture verification, and gesture classification. Here's a breakdown of how each component works together in the process:

1) *Camera:* A camera is used to capture the sign language gestures performed by the user. The camera captures a video stream of the user's hand movements, which is then processed by the software
2) *Mediapipe Library:* The Mediapipe library is a computer vision library that is used to detect and track hand movements in the video stream captured by the camera. It uses machine learning models to identify the key points on the user's hand, such as the position of the fingers, palm, and wrist.
3) *Feature Extraction into Array:* The key points identified by the Mediapipe library are used to extract features that describe the shape and movement of the user's hand. These features are then converted into an array of numerical values that can be processed by the software.
4) *Data Points:* The array of features extracted from the hand movements is treated as a sequence of data points. Each data point represents the position of the hand at a specific point in time.
5) *Image Matching:* The sequence of data points is compared to a database of known sign language gestures using image matching algorithms. The image matching algorithm compares the extracted features with the features of known gestures to identify the closest match.
6) *RNN Algorithm:* A Recurrent Neural Network (RNN) algorithm is used to process the sequence of data points and predict the sign language gesture being performed. The RNN algorithm is trained using a dataset of labeled sign language gestures and can predict the gesture being performed with high accuracy.
7) *Gesture Verification:* The predicted gesture is verified using a gesture verification algorithm. This algorithm checks whether the predicted gesture matches the expected gesture based on the context of the conversation.
8) *Gesture Classification:* The final step is to classify the verified gesture into a text message. This is done using a gesture classification algorithm that maps each sign language gesture to a corresponding text message. The resulting text message is then displayed to the user or transmitted to another user in the conversation.
9) *Sign to Text:* This is the final step of the system which gives the actual output of the sign language to text.

Overall, the process of converting sign language to text requires a combination of computer vision, machine learning, and natural language processing technologies to accurately detect, recognize, and translate sign language gestures into text messages and the developed model was able to detect various hand gestures and signes with an accuracy of 96.66%.

V. CONCLUSIONS

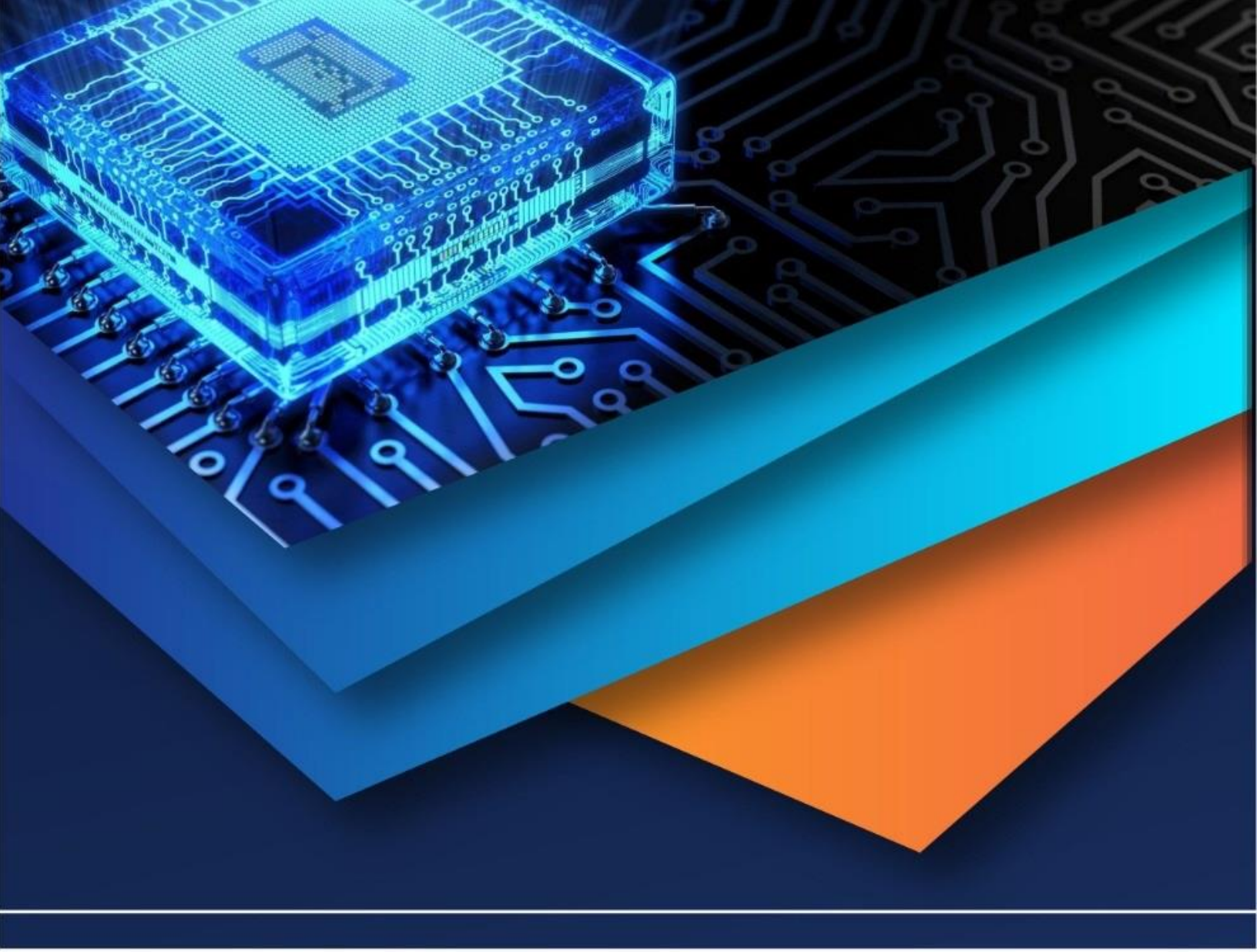
The project is focused on solving the problem of deaf and dumb people. This system will automate the hectic task of recognizing sign language, which is difficult to understand for a normal person, thus it reduces the efforts and increases time efficiency and accuracy. Using various concepts and libraries of image processing and fundamental properties of image we trying to develop this system. This paper represented a visioned based system able to interpret hand gestures from the sign language and convert them into text. The proposed system is tested in the real-time scenario, where it was possible to prove that obtained RNN models were able to recognize hand gestures. As future work is to keep improving the system and make experiments with complete language datasets.

VI. ACKNOWLEDGMENT

We are delighted to present the paper on "Conversion of Sign Language to Text." We would like to seize this moment to express our heartfelt gratitude to Prof. Jitendra Musale, our internal guide, for his unwavering assistance and invaluable guidance throughout the project. His support has been instrumental in our progress, and we are truly thankful for his contributions. We would also like to extend our deepest appreciation to Dr. Sunil Thakare, the principal of ABMSP's Anantrao Pawar College of Engineering Research, for his continuous support and encouragement. Additionally, we are grateful to Prof. Rama Gaikwad, the project head at ABMSP's Anantrao Pawar College of Engineering & Research, for their indispensable guidance, insightful suggestions, and for providing us with the necessary infrastructure to carry out our project effectively.

REFERENCES

[1] S. M Mahesh Kumar, “Conversion od Sign Language into Text,” International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 9, 2018.
[2] Kohsheen Tiku, Jayshree Maloo, Aishwarya Ramesh, Indra R, “Real-time Conversion of Sign Language to Text and Speech,” 2020 Second International Conferenceon Inventive Research in Computing Applications, Coimbatore, India, 2020, pp. 346-351.
[3] C. Uma Bharti, G. Ragavi, K. Karthika "Signtalk: Sign Language to Text and Speech Conversion," 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), Coimbatore, India, 2021, pp. 1-4, doi: 10.1109/ICAECA52838.2021.9675751.
[4] M. Zamani and H. R. Kanan, "Saliency based alphabet and numbers of American sign language recognition using linear feature extraction," 2014 4th International Conference on Computer and Knowledge Engineering (ICCKE), Mashhad, Iran, 2014, pp. 398-403, doi: 10.1109/ICCKE.2014.6993442.
[5] A. Saxena, D. K. Jain and A. Singhal, "Sign Language Recognition Using Principal Component Analysis," 2014 Fourth International Conference on Communication Systems and Network Technologies, Bhopal, India, 2014, pp. 810-813, doi: 10.1109/CSNT.2014.168.
[6] J. Zhang, W. Zhou, C. Xie, J. Pu and H. Li, "Chinese sign language recognition with adaptive HMM," 2016 IEEE International Conference on Multimedia and Expo (ICME), Seattle, WA, USA, 2016, pp. 1-6, doi: 10.1109/ICME.2016.7552950.
[7] D. Guo, W. Zhou, M. Wang and H. Li, "Sign language recognition based on adaptive HMMS with data augmentation," 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 2016, pp. 2876-2880, doi: 10.1109/ICIP.2016.7532885.
[8] K. Grobel and M. Assan, "Isolated sign language recognition using hidden Markov models," 1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation, Orlando, FL, USA, 1997, pp. 162-167 vol.1, doi: 10.1109/ICSMC.1997.625742.
[9] D. Guo, W. Zhou, M. Wang and H. Li, "Sign language recognition based on adaptive HMMS with data augmentation," 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 2016, pp. 2876-2880, doi: 10.1109/ICIP.2016.7532885.
[10] D. Van Hieu and S. Nitsuwat, "Image Preprocessing and Trajectory Feature Extraction based on Hidden Markov Models for Sign Language Recognition," 2008 Ninth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, Phuket, Thailand, 2008, pp. 501-506, doi: 10.1109/SNPD.2008.80.



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)