iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IV **Month of publication:** April 2023

DOI: https://doi.org/10.22214/ijraset.2023.51125

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 IV Apr 2023- Available at www.ijraset.com

Two Way Sign Language for Deaf and Dumb using Deep Convolution Neural Network

Vishnu Priya A. K[1], Jayashri. S[2], Sivanjali. V[3], Sneha. V[4], Thamaraiselvi.K[5]
[1, 2, 3, 4]*UG Scholars,* [5]*Assistant Professor, Department of Computer Science and Engineering , Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, India*

Abstract: ***The goal is to build a bridge between the hearing and hearing-impaired communities and to start two-way communication. Here we propose a two-way sign language translator that translates speech in real time into sign language and vice versa. Each video frame is processed using Python's Open CV Library. Additionally, we use the Gaussian Mixture-based Background/Foreground Segmentation Algorithm to remove the background from each frame. The hearing impaired community will then be more effectively integrated into regular schools thanks to this technology, making education more affordable and available to them.***

Keywords: Gesture recognition , Deep Convolution Neural Network (DCNN), Speech recognition, Hand Signs.

I. INTRODUCTION

Since so few people outside of the deaf-mute population are familiar with sign languages, it presents a significant obstacle to communication between the speaking and deaf-mute communities and, consequently, to the advancement of these cultures as a whole. Voice to sign language translators and sign language to voice translators are the two categories into which the currently available sign language translators fall. To recognise the ArSL signs generated by the deaf person, they have demonstrated a method that makes use of gloves with sensors integrated in them. The use of deep neural networks to analyse the spatiotemporal aspects for sign language identification has gained popularity in recent years. For example, classify movements using the DCNN architecture. The DCNN architecture requires RGB-D inputs with a depth dimension, which are produced by devices with specialised depth sensors, despite the fact that these inputs are highly accurate. The issue persists when deaf persons use hand signs to communicate, which makes it difficult for other people to understand their language from the signs they employ. A system that can identify various indications and provide information to common people was required. The suggested system can recognise hand movements from the complete image in real-time without the use of any image region selection framework. This is accomplished through the careful construction of a deep CNN, which can assist in recognising hand motions made in any place within the image under the condition that the hand making the gesture only occupies a small portion of the image relative to the entire image. The image is further overloaded by background elements, including upper bodies and faces of people. As a consequence, there are two fundamental guidelines that have oriented the design of the developed deep network.

II. RELATED WORK

Dina A. Alabbad,Nouha O. Alsaleh,Naimah A. Alaqeel,Yara A. Alshehri,Nashwa A. Alzahrani,Maha K. Alhobaishi, etal.[1](2017) discussed about “A Robot-based Arabic Sign Language Translating System”where services for the deaf in the Eastern province of Saudi Arabia were assessed, confirming a critical need for deaf community support. The process of translating Arabic sign language (ArSL) into written text is accomplished in the study using the Pepper robot, which aids in the robot's ability to distinguish static hand motions for each letter in ArSL from each frame of the input video.

Rishi K,PrarthanaA,Pravena K S,S. Sasikala,S. Arunkumaretal.[2](2018)explains about “Two-Way Sign Language Conversion for Assisting Deaf-Mutes Using Neural Network”By creating an assistive tool that converts sign language into a legible format will enable deaf-mutes to easily communicate with the general public. Convolution Neural Network (CNN) technology is used in the suggested method to translate sign language into voice. The accuracy provided by the CNN model is 95.5%.

Abdul-Latif Hamdan, JadJabb Wuyang Qin, XueMei,Yuming Chen, Qihang Zhang, Yanyin Yao, ShiHuouret al.[3] (2013) discussed about “Sign Language Recognition and Translation Method based on VTN”where they create a simple sign language translation network using VTN (Video Transformer Net). Create the CSL BS (Chinese Sign Language-Bank and Station) dataset and two-way VTN to compare the isolated sign language with I3D. (Inflated three Dimension)

©IJRASET: All Rights are Reserved

SK Nahid Hasan, Md. Jahid Hasan, Kazi Saeed Alametal.[4] (2016) explains about "Comprehensive Multipurpose Dataset for Bangla Sign Language Detection"They employ a dataset to recognise Bangla Sign Language, which is crucial for these persons with disabilities. It is incredibly difficult to have a Bangla Sign Language dataset, and very little study has been done in this area. They have produced a comprehensive dataset of Bangla sign language that include all letters and numerals in order to overcome this problem, which they hope to make available for use in future studies.

Niels Martínez-Guevara, José-Rafael Rojano-Cáceres, Arturo Curiel[5](2017)discussed about "Detection of Phonetic Units of the Mexican Sign Language"provides a method for the unsupervised identification of likely phonetic units in raw SL videos using the expectation maximisation approach. The objective is to identify the smallest linguistic units that may hold important information in order to discover new features for SL-based NLP

Frank M. Shipman,Caio D.D. Monteiroet al.[6] (20) discussed about " Crawling and Classification Strategies for Generating a Multi-Language Corpus" In order to apply multimodal sign language detection and identification classifiers to winnow the videos collected to those thought to be in a certain sign language, SLaDL employs a variety of crawling techniques to gather prospective sign language content. This model contrasts the quantity and variety of sign language films found using targeted, breadth-first, and depth-first crawling techniques.

The 3-way classification task of identifying videos in American Sign Language uses them to compare the precision of various methods by combining textual metadata and video features (ASL).

Gaby Abou Haidar, Roger Achkar, Dian Salhab, Antoine Sayah, Fadi Jobran, et al.[7] discussed about "Sign Language Translator using the Back Propagation Algorithm of an MLP" The best option for enabling deaf/mute people to communicate fluently through technology in different languages is a sign language translator. Combining two gloves with the necessary sensors and a smartphone with a mobile app that uses Lebanese Sign Language to transform hand movements into spoken words, artificial neural networks play a significant role in determining the precise output.

Ariya Thongtawee, Onamon Pinsanoh, Yuttana Kitjaidureet al.[8] discussed about " A Novel Feature Extraction for American Sign Language Recognition Using Webcam" provides a quick and effective feature extraction approach to identify the alphabets of American Sign Language in both static and moving gestures. The suggested algorithm uses four different approaches: White pixel count at the image's edge (NwE), finger length from the centroid point (Fcen), angles between fingers (AngF), and differences between first- and last-frame angles between fingers (delAng). An Artificial Neural Network (ANN) is used to categorise the signs after features are extracted from video images.

Ashish S. NikamAarti G. Ambekar, et al.[7] (20) discussed about " Sign Language Detection from Hand Gesture Images" The best option for enabling deaf/mute people to communicate fluently through technology in different languages is a sign language translator. It consists of two gloves equipped with the necessary sensors and a smartphone with a mobile app that, using Lebanese Sign Language, translates hand gestures into spoken words. Artificial Neural Networks are important in providing the precise output.

III. SYSTEM MODEL

The suggested method can recognise hand gestures right away and helps to extract motions from the complete image without the usage of any image region selection framework.

This is accomplished through the careful construction of a deep CNN, which can assist in recognising hand motions made in any place within the image under the condition that the hand making the gesture only occupies a small portion of the image relative to the entire image.

The image is further overloaded by background elements, including upper bodies and faces of people. As a result, the design of the created deep has been guided by two essential principles.

With the help of the Open CV framework, the suggested methodology is able to recognise hand gestures from a specific area of a picture quickly.

This is carried out using a (Deep Neural Network)DNN that was carefully designed. Following gesture recognition, text is used to replace the appropriate gestures and is then transformed to sound using Python script. The suggested system can also translate spoken words into gestures, so it first records spoken words as sound, transforms them to text using a script, separates the text, and then activates a callback script that sequentially displays each gesture image for each

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

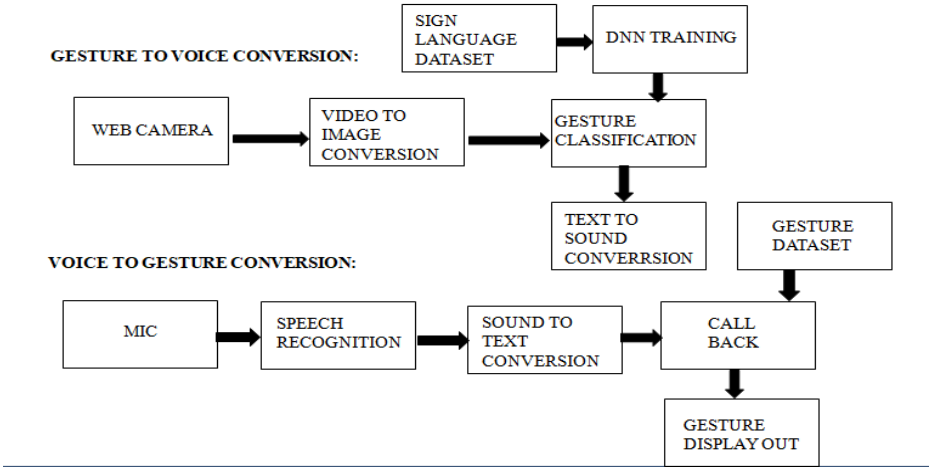


Fig. 1 Block Diagram of Two way Sign language

A. Convolution Neural Network

A convolution neural network (CNN) is a sort of deep neural network used in deep learning that works with a set of data to extract information about that input. Similar to photos, sounds, or videos, CNN can use these for data extraction. In CNN, three things predominate. Local receptive field comes first, followed by shared weight and biases, and finally activation and pooling. To enable CNN to extract the feature of an input, neural networks are first trained using a large amount of data.When an input is received, it undergoes picture pre-processing, feature extraction based on a set of stored data, data categorization, and output before being shown as the outcome.

Only the input used to train the neural network and save data is handled by the CNN. They are utilised in natural language processing, picture classification, recommender systems, medical image analysis, and image and video recognition

IV. SYSTEM IMPLEMENTATION

The system has the following five modules

- Training the dataset using DCNN
- Capturing the hand gesture
- Classifying the hand gestures
- Converting text to audio
- Converting sound to text

A. Training the Dataset using DCNN

One of the key benefits of using CNNs over NNs is that the technique helps the machine, capable of handling picture data in 2D, so there is no need to flatten the input images to 1D. This aids in keeping an image's "spatial" qualities. As a result, this article uses the Kaggle Accident data set, which includes the categories Accident and Normal. The dataset of different hand signs is collected and pre-processed and trained using Deep convolution neural network to create the model file. The model file is used to train the CNN algorithm to classify the gesture images that are obtained by the open cv framework.

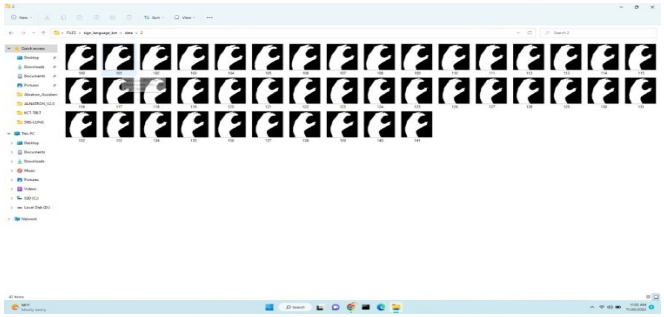


Fig.2 Pre processed dataset

©IJRASET: All Rights are Reserved

B. Capturing the Hand Gesture

Computer Vision is antechniques that helps to deploy the camera to capture real-time hand gestures. The image was converted to grey -scale. Canny edge detection algorithm is a gray scale method that helps to convert the RGB images into machine readable format for the machine and they are applied here. Computer vision is the way of teaching intelligence to the machines and making them see the images or videos and extract the data from them just like humans. By using Open CV framework, a python library that helps to capture the input image from the webcam of the system.

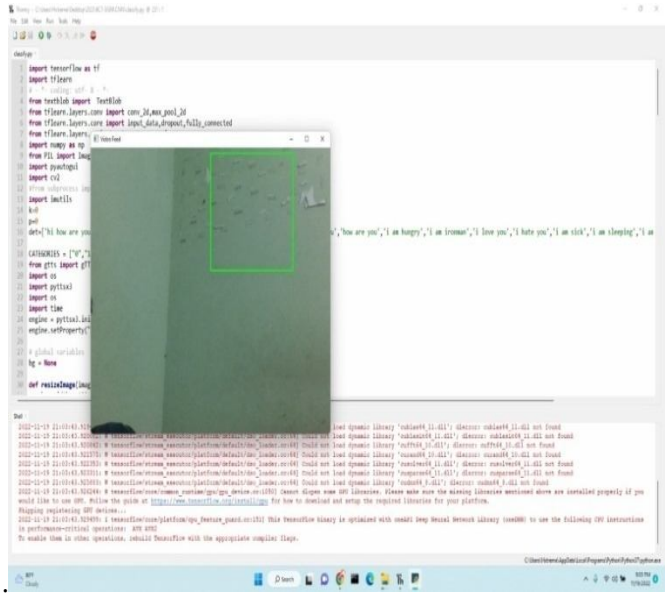


Fig. 3 Capturing the Hand Gesture

C. Classifying the Hand Gestures

Finally the image taken from the webcam is classified using trained model and the gesture is converted in to respective text. The image obtained using Open CV is compared with the model file that contains the pre-processed images of the gestures taken from the kaggle website. CNN algorithm is used to classify the difference between the obtained image and the dataset images. This algorithm helps to test the obtained image to the model file.

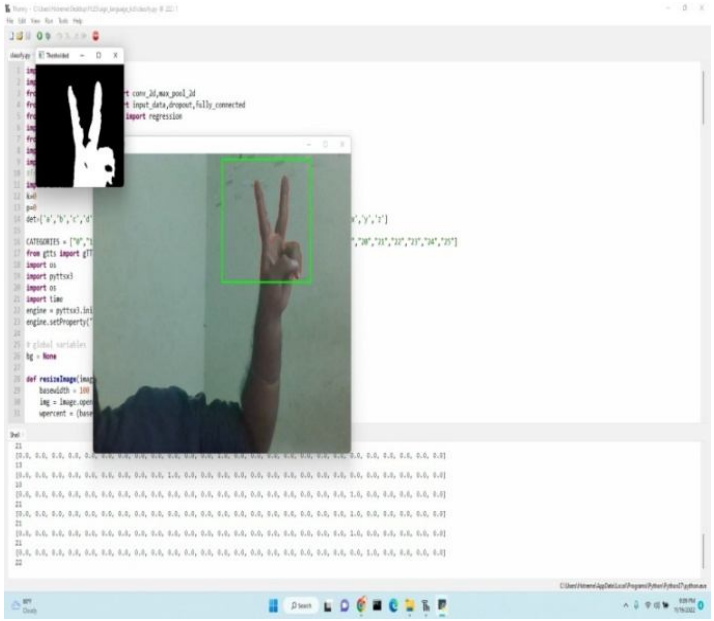


Fig.4 Classifying the Image Captured

D. Converting Text to Audio

Text is converted in to speech using Gtts. It is a service used for speech recognition which is developed by Google for screen reading purpose application for the operating system. It provides applications to read the text aloud from the screen for many languages with many support. The sentence that are displayed is converted to voice for easy understanding.

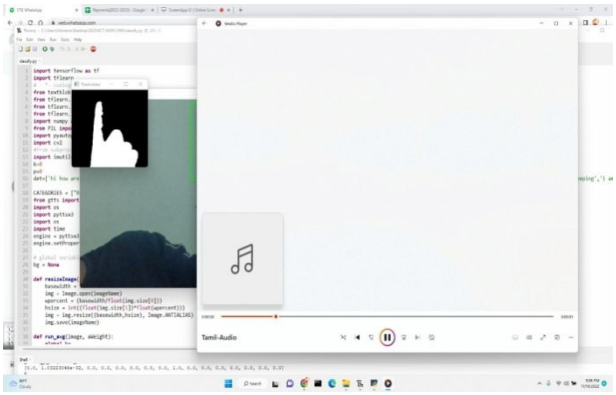


Fig.5 Voice Output of Hand Gesture

E. Converting Audio to Hand Sign

Speech recognition is a library that helps the computer software to identify words and phrases in the language that is spoken so they are converted to human-readable text this is done using python speech recognition library and finally the text is mapped with respective gestures using python script.

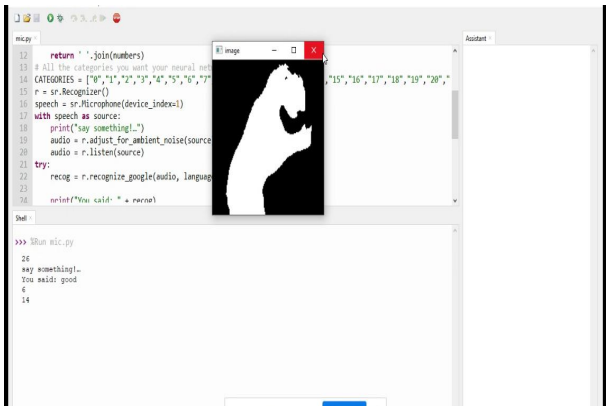


Fig.6 Conversion of Audio to HandSign

V. ADVANTAGES

Bidirectional sign language, sometimes referred to as two-way sign language or symmetrical communication, is a form of sign language that both hearing and deaf people can use to communicate with one another. Two-way sign language has the following benefits:

1) Improved communication: Two-way sign language makes it possible for hearing and deaf people to communicate more effectively. This can be especially helpful in circumstances when speaking is impractical or challenging to comprehend.
2) Inclusivity: Hearing people can converse with deaf people in their preferred language by using two-way sign language. This encourages equality of participation and inclusivity in society.
3) Improved socialization: Deaf and hearing people who use two-way sign language can have better social interactions. It can assist in removing obstacles and fostering appreciation for one another's uniqueness.

VI. DISADVANTAGES

1) Limited Accessibility: Because two-way sign language requires that both participants understand and use sign, it may be inaccessible to persons who are deaf or have limited sign language skills. For others who rely on other modes of communication, this may result in marginalization and communication hurdles.

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

2) Sign Language Variations: There are several differences between the sign languages used in various nations and areas, which can make it challenging for people who do not speak the same sign language to effectively communicate. Miscommunication and frustration may result from this.

3) Limited Expression: Sign language is a limited form of communication that mainly relies on facial expressions and body language to transmit meaning, which can be difficult for people who are not conversant in its subtleties. In order to communicate complicated emotions or thoughts, this can restrict the spectrum of expression.

4) A lack of resources may make it difficult to enable two-way sign language communication. Examples of such resources include qualified interpreters and assistive technology. As a result, successful communication between hearing and deaf people in a number of contexts may be challenging.

5) Social stigma: The use of sign language may be subject to social stigma, especially among hearing people who may perceive it as a "lesser" mode of communication. When a deaf person uses sign language as their only form of communication, it might cause them to feel excluded and alone.

VII. CONCLUSION & FUTUREWORK

By implementing a 2-way sign language translator that can run on any processor device, the suggested system closes the communication gap between the hearing and deaf communities. Because it operates with an RGB input from a typical camera, the DNN architecture used for sign to text translation is particularly well-suited for this purpose. However, our proposed approach is quite effective at performing a speedier conversion for sign language to text.

In the future, this technology could be made available as an Android application for mobile devices, enabling deaf and mute persons to connect and conduct conversations with hearing people.

REFERENCES

[1] Dina A. Alabbad,Nouha O. Alsaleh,Naimah A. Alaqeel,Yara A. Alshehri,Nashwa A. Alzahrani,Maha K. Alhobaishi ,A Robot-based Arabic Sign Language Translating System ,2022 7th International Conference on Data Science and Machine Learning Applications (CDMA)

[2] Rishi K,Prarthana A,Pravena K S,S. Sasikala,S. Arunkumar,Two-Way Sign Language Conversion for Assisting Deaf-Mutes Using Neural Network,2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS)

[3] Debashish Sau, Swapnanil Dhol, Meenakshi K, Kayalvizhi Jayavel, A Review on Real-Time Sign Language Recognition, 2022 International Conference on Computer Communication and Informatics (ICCCI)

[4] Wuyang Qin,Xue Mei,Yuming Chen,Qihang Zhang,Yanyin Yao,Shi Hu,Sign Language Recognition and Translation Method based on VTN,2021 International Conference on Digital Society and Intelligent Systems (DSInS)

[5] L. Priya,A. Sathya,S. Kanaga Suba Raja,Indian and English Language to Sign Language Translator- an Automated Portable Two Way Communicator for Bridging Normal and Deprived Ones,2020 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS)

[6] Gaby Abou Haidar,Roger Achkar,Dian Salhab,Antoine Sayah,Fadi Jobran,Sign Language Translator using the Back Propagation Algorithm of an MLP,2019 7th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)

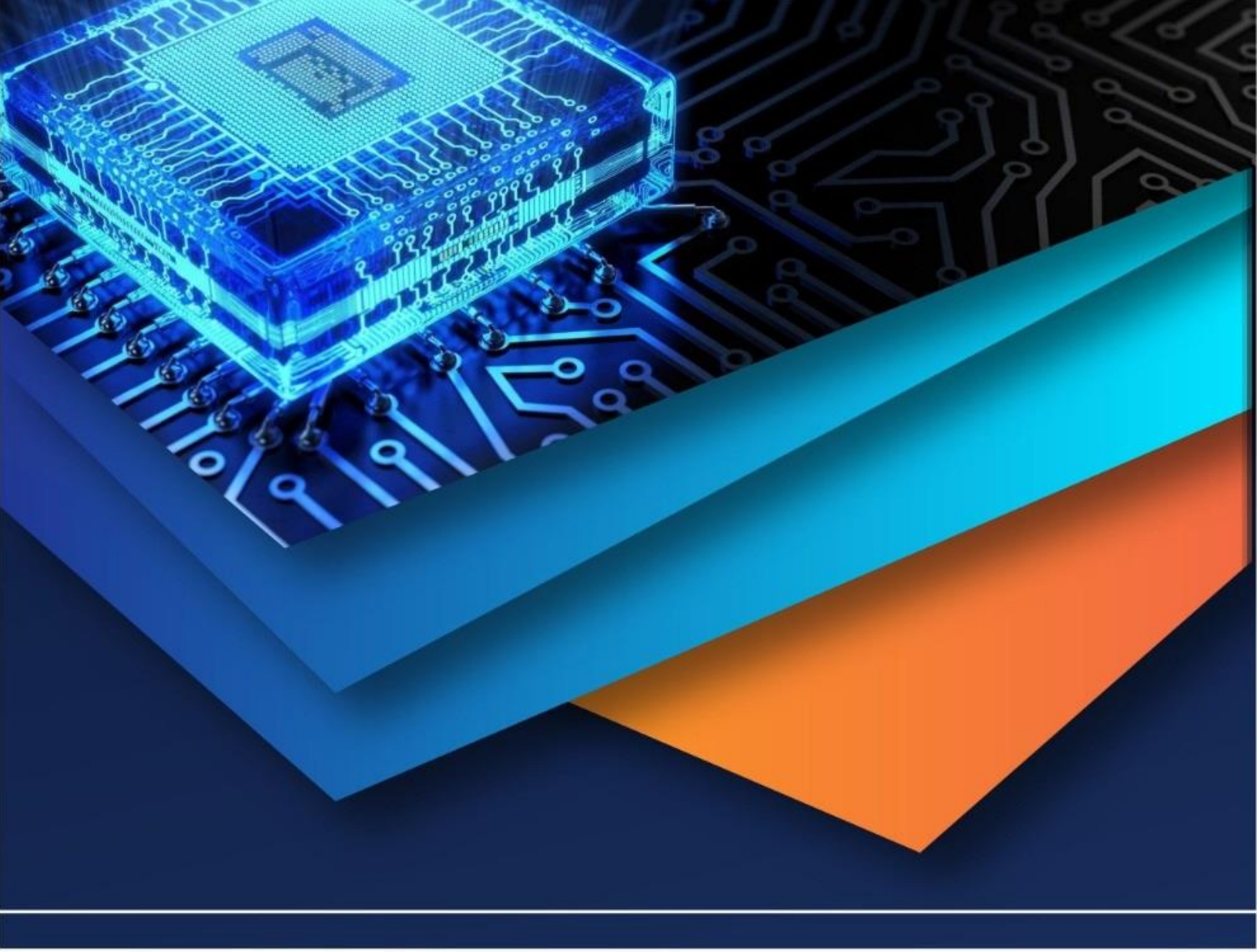
[7] Walaa Aly, Saleh Aly, Sultan Almotairi, User-Independent American Sign Language Alphabet Recognition Based on Depth Image and PCANet Features, IEEE , 2019

[8] R. Rastgoo, K. Kiani, and S. Escalera, Multi-modal deep hand sign language recognition in still images using restricted boltzmann machine, Entropy, vol. 20, no. 11, p. 809, 2018.

[9] Hasan, M., Sajib, T. H., Dey, M.A machine learning based approach for the detection and recognition of Bangla sign, 2016, International Conference on Medical Engineering, Health Informatics and Technology(MediTec).

[10] Rajaganapathy, S., Aravind, B., Keerthana, B., Sivagami, M, Conversation of Sign Language to Speech with Human Gestures , 2015

©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)