iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** II **Month of publication:** February 2023

DOI: https://doi.org/10.22214/ijraset.2023.49007

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 II Feb 2023- Available at www.ijraset.com

A Survey Paper on Sign Language Recognition System using OpenCV and Convolutional Neural Network

Himanshu Tambuskar[1], Gaurav Khopde[2], Snehal Ghode[3], Sushrut Deogirkar[4], Er. Manisha Vaidya[5]
[1, 2, 3, 4]*Students,* [5]*Assistant Professor, Department of Computer Science and Engineering, Priyadarshini JL College of Engineering, Nagpur-440027, India*

Abstract: ***Communication is a very important part of our Human life to express feelings and thoughts. People like the Deaf & Dumb always face difficulty as they cannot speak in their regional languages. Language performs a very important role in communication, it can be verbal i.e. using words to speak, read and write or non-verbal using facial expressions and sign language. So, people like the Deaf and Dumb have the only choice to speak in Sign language means non-verbal. However, Sign language is a very important mode of their community. But it is difficult for people who are unaware of sign language. Hence, here is a system "Sign Language Recognition System Using Open CV and Convolutional Neural Network". We proposed a system that converts sign language to their appropriate alphabet, words in a standard language to make easily understood by all. We also make some default gestures that we daily use in our day-to-day life. The project works on a learning algorithm, it requires the collection of datasets which includes images of each alphabet, and digits to train the model. For the classification of the image convolutional neural network is used. Also for accessing the camera and taking input we used an open cv.***
Keywords: Data collection, Image Pre-Processing, Data Training, Hand Gesture Recognitions.

I. INTRODUCTION

The Deaf and dumb which not able to speak and hear properly for such people only have one mode of communication which is non-verbal. It can be with the help of sign language, facial expressions, gestures, and electronic devices. It would be difficult for them to explain what they want to convey to normal people. It is difficult and expensive to find an experienced interpreter on a regular basis.
We were aiming to develop a system that converts sign language into text format with the help of a vision-based approach so it becomes cost-effective. Sign language consists of a variety of hand transformations, orientations, facial expressions, and hand movements that are used to transmit messages. Every sign is allocated to a particular alphabet and meaning. Some languages are found globally such as American Sign Language (ASL), British Sign Language (BSL), Japanese sign language (JSL), and so on.
Normal people never try to learn sign language to communicate with deaf and dumb people. This leads to the isolation of deaf and dumb people this isolation can be removed with the help of a computer. If a computer can be programmed in such a manner to translate sign language into text format. From this paper, you will get information about, how we create a system, what the requirement, and what kind of data we used for the training and testing of a system, it gives information about the previous research done on sign language, In the end, it contains a conclusion.

II. OBJECTIVE

The goal of the Sign Language Recognition Project, is a real-time vision-based system, is to determine the American Sign Language represented by the alphabet shown in Fig. 1. The prototype's goals were to assess the feasibility of a vision-based system for sign language recognition and, concurrently, to evaluate and choose hand features that could be used to machine learning algorithms to enable real-time sign language recognition systems.

The adopted approach simply makes use of one camera and is based on the preceding notions:

1) The user must remain in front of the camera limitations, within a specified boundary.
2) Due to the camera, the user must be within a certain distance range.
3) A bare palm is used to define the hand posture. andand not occluded by other objects.
4) The system must be used indoors

©IJRASET: All Rights are Reserved

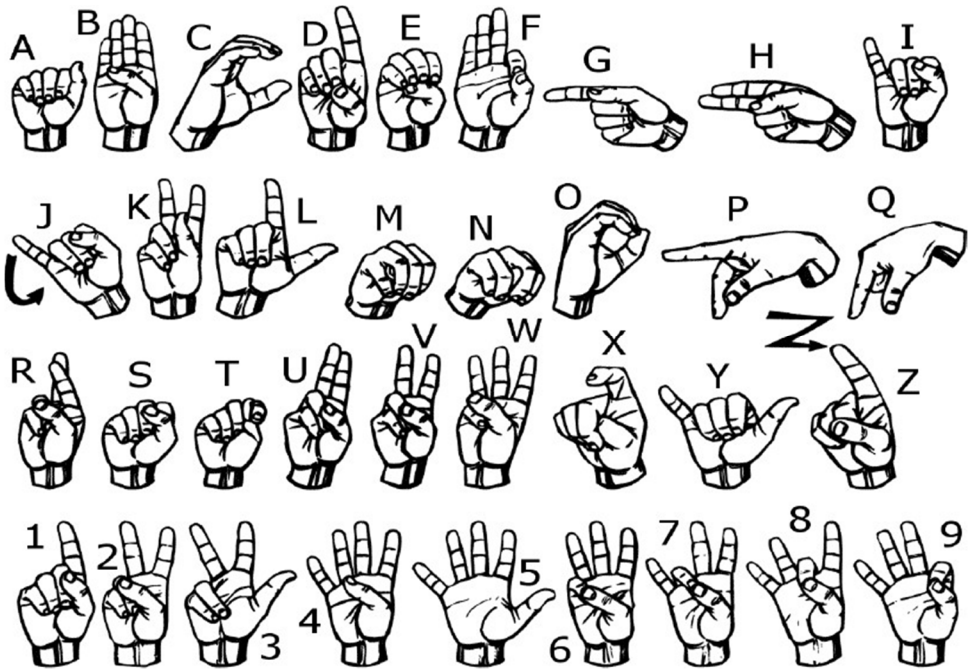


Fig .1 American Sign Language Alphabets and Number

III. LITERATURE SURVEY

1) *A CNN based human-computer interface for American Sign Language recognition for hearing-impaired individuals:* The author develops an interface using a convolutional neural network [1] [2] [3] [5], based on sign language to interpret gestures and hand poses of signs. They used a deep learning algorithm and open cv for accessing the camera. The author developed their own dataset so it helps in the prediction of hand pose and it helps to increase the accuracy. The dataset may be used to develop SLR systems. The proposed system may give solutions in the medical field that use deep learning.
2) *Bangla Sign Language (BdSL) Alphabets and Numerals Classification Using a Deep Learning Model:* According to the author, their model which is a real-time Bangla Sign Language interpreter can enable more than 200 k hearing and speech-impaired people to the mainstream workforce in Bangladesh. It is a challenging topic in computer vision and deep learning research because sign language recognition accuracy may vary on skin tone, hand orientation, and background. The dataset prepared in this study comprises the largest image database for BdSL Alphabets and Numerals in order to reduce interclass similarity while dealing with diverse image data, which comprises various backgrounds and skin tones. This model works on the CNN algorithm, The CNN [1] [2] [3] [5], model trained with the images that had a background was found to be more effective than those without a background. This dataset which is being provided in this study comprises the biggest accessible dataset for BdSL Alphabets and Numerals in order to reduce inter-class similarity while dealing with diverse image data, which comprises various backgrounds and skin tones.
3) *Sign Language Recognition System:* In this paper American Sign Language translator is created on a web application subject to a CNN [1][2][3][5], classifier. A strong model for letters a-e, and a modest one for letters a-k (excepting j) is the consequence of this work. Since it was impractical to discover a variety in our datasets, the validation correction percentage checked during planning were not clearly reproducible subsequent to testing on the web application.
4) *An integrated mediapipe-optimized GRU model for Indian sign language recognition:* According to the author, sign language recognition is challenged by many problems such as recognizing gestures accurately, and occlusion of hand and it becomes costly. The author proposed an integrated mediapipe optimized gated recurrent unit model for Indian sign language recognition. For data preprocessing and feature extraction from the image, they used mediapipe framework, the mediapipe holistic handled individual models for the hands, face, and pose components using an image resolution. Thus their model achieved better prediction accuracy, high efficiency to learn, and capability to process information and becomes faster than other models by replacing standard GRU (gated recurrent unit) with exponential linear activation and softmax with softening activation in the output layer of GRU cell.

5) *Sign Language Recognition:* Sign Language is mainly used by deaf (hard hearing) and dumb people to exchange information between their own community and with other people. Sign Language Recognition (SLR) deals with recognizing the hand gestures acquisition and continues till text or speech is generated for corresponding hand gestures. Computer Vision is a field of Artificial Intelligence that focuses on problems related to images and videos. CNN [1][2][3][5] combined with Computer vision is capable of performing complex problems. To develop a practical and meaningful system that can able to understand sign language and translate that to the corresponding text. There are still many shortages of our system this system can detect 0-9 digits and A-Z alphabets hand gestures but doesn't cover body gestures and other dynamic gestures. We are sure it can be improved and optimized in the future.
6) *Isolated Sign Recognition from RGB Video using Pose Flow and Self-Attention:* In This Proposed Work Author uses The Video Transformer Network (VTN), which is originally proposed by Kozlov et al. The work was done in the context of the ChaLearn 2021 Watching large People Scale Signer Independent Isolated SLR [6][8] CVPR Challenge. In this work, the author introduces the recently released AUTSL dataset for isolated sign recognition and obtains 92.92% accuracy on the test set using only RGB data which is captured using a Kinect camera. This dataset consists of 36,302 samples. Each sample corresponds to one of 226 signs and is performed by one of 43 different persons. To increase the Accuracy they Crop out images of the hands as input for the network as opposed to using full frames of the video, which include irrelevant information and possibly background noise. By visualizing and interpreting both spatial salience maps and attention masks in the top-performing model, they carry out a qualitative analysis of the model. This analysis sheds light on how multi-head attention functions in the context of sign language recognition.
7) *Real-time Vernacular Sign Language Recognition using Mediapipe and Machine Learning:* According to the author, their model is lightweight and can be used in smart devices also used American, Italian, Indian, and turkey sign language to train the model and used support vector machine algorithm without any use of sensors that it becomes easy and comfortable. The preprocessing is done by mediapipe , and prediction is done using a machine learning algorithm, for analyzing each dataset they used a performance matrix and quantitative analysis. Also, the author said they achieve average 99% accuracy. The advantages of the model are that it is adaptable to any regional language and cost-effective, and maximum accuracy can be obtained.
8) *Sign Language Recognition via Skeleton-Aware Multi-Model Ensemble:* Sign language is commonly used by deaf or mute people to communicate but requires extensive effort to master. It is usually performed with the fast yet delicate movement of hand gestures, body posture, and even facial expressions. Current Sign Language Recognition (SLR) methods usually extract features via deep neural networks and suffer overfitting due to limited and noisy data. They propose a novel SAM-SLR-v2 framework to learn multi-modal feature representations from RGBD videos toward more effective and robust isolated SLR [6][8]. Among those modalities, these proposed skeleton-based methods are the most effective in modelling motion dynamics due to their signer-independent and background-independent characteristics. we construct novel 2D and 3D spatiotemporal skeleton graphs using pre-trained whole-body key point estimators and propose a multi-stream SL-GCN to model the embedded motion dynamics. our proposed SAM-SLR-v2 framework achieves state-of-the-art performance on three challenging datasets for isolated SLR (i.e., AUTSL, SLR500, and WLASL2000) as well as won the championships in both RGB and RGB-D tracks during the CVPR 2021 challenge on isolated SLR.
9) *Sign Language Recognition For Sentence-Level Continues Signing:* In this Work, the Author proposes a method to bridge the communication gap between hearing-impaired people and Uses Myo armbands for gesture-capturing [9], signal processing, and supervised learning based on a vocabulary of 49 words and 346 sentences for training with a single signer. The sign language of Sri Lanka was chosen by the author as the sign language for this research project. there are About 2000 characters of Sri Lanka Sign Language 49 of which were used in the study. The signs selected are those that are common and useful in daily life. 49 characters Include only nouns, pronouns, nouns, and verbs. Project Work Achieves the average recognition accuracy was up to 75-80% word-level accuracy and 45-50% sentence-level accuracy using gestural (EMG) and spatial (IMU) features for the signer-dependent experiment.

From the above literature survey, the author uses different techniques to implement and develop the model which is based on a vision-based approach, sensors, MOPGRU (Mediapipe optimized gated recurrent unit) [4], CNN(Convolutional Neural Networks) [1][2][3] and [5], which is used for image recognition and tasks that involve the processing of pixel data. LSTM is used to learn, process, and classify sequential data because these networks can learn long-term dependencies between time steps of data. It is observed that CNN is the most frequently used algorithm in the above papers since is it used for model building.

IV. COMPARATIVE STUDY

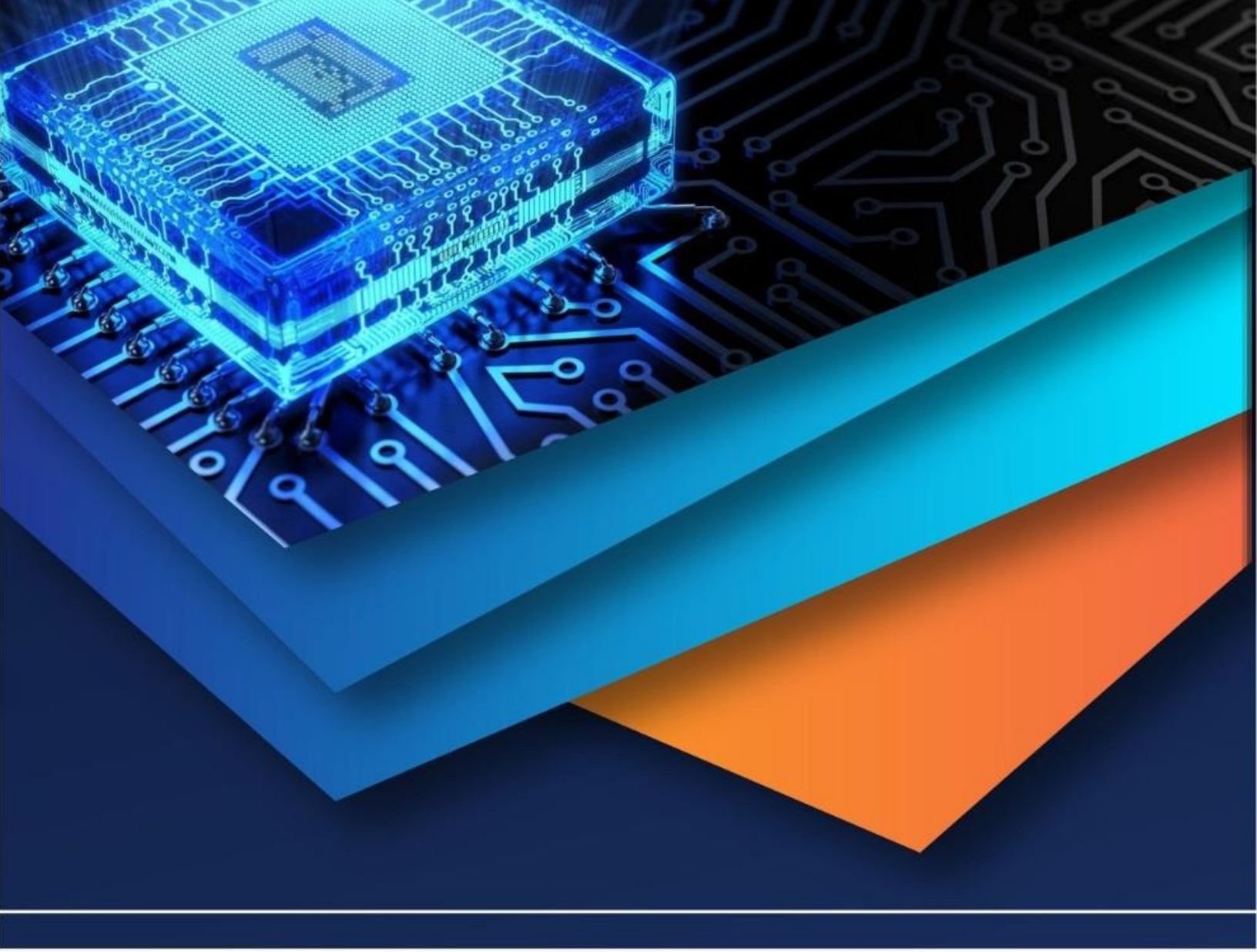
Sr.no	Paper Name	Author Name	Methodology	Advantage	Disadvantage
[1]	A CNN based human computer interface for American Sign Language recognition for hearing-impaired individuals. [2022]	Ahmed Kasapbasi Ahmed Eltayeb Ahmed Elbushra Omar Al-Hardanee Arif Yilmaz	CNN(Convolutional Neural Networks)	The system gives 95.50% accuracy due to the large dataset.	Background affects accuracy, so it was necessary to have the right background.
[2]	Bangla Sign Language (Bdsl) Alphabets and Numerals Classification Using A Deep Learning Model.[2022]	Kanchon K. Podder, Muhammad E.H. Chowdhury, Anas M. Tahir, Zaid Bin Mahbub, Md Shafayet Hossain, Muhammad Abdul Kadir	CNN(Convolutional Neural Networks)	The CNN model performed better when trained on background-free images so the auther uses binary images for training the model.	Background affects accuracy.
[3]	Sign Language Recognition System [2022]	Pooja M.R, Meghana M, Praful Koppalkar, Bopanna M J, Harshith Bhaskar, Anusha Hullali	ANN(Artificial Neural Networks),HOG(Histogram Oriented Gradient)	Clear Visualization	It does not predict some alphabet.
[4]	An integrated mediapipe-optimized GRU model for Indian sign language recognition.[2022]	Barathi Subramanian, Bekhzod Olimov, Shraddha M. Naik, Sangchul Kim, Kil-Houm Park, Jeonghong Kim	MOPGRU(Mediapipe optimized gated recurrent unit)	MOPGRU captured the full information dependency in time series data with a high prediction accuracy of an 95%.	It is complex process.
[5]	Sign Language Recognition[2021]	Satwik Ram Kodandaram., Sunil GI, N.Pavan Kumar	Django Rest Framework, CNN(Convolutional Neural Networks)	It gives 97% Accutacy.	It dosen't Cover body gestures and other dynamic gesture
[6]	Isolated Sign Recognition from RGB video using Pose Flow and Self-Attention[2021]	Mathieu De Coster . Mieke Van Herreweghe. Joni Dambre.	Pose Flow-VTN(video transformer network)	Pose flow is allows better distinguish between similar hand shape.	The model does not predict some signs.
[7]	Real-time Vernacular Sign Language Recognition using MediaPipe and Machine Learning[2021]	Arpita Haldera , Akshit Tayadeb	SVM(Support Vector Machine),MediaPipe framework	By the use of MediaPipe it detect complex hand gesture precisely. It is adaptable in smart devise.	It takes more time to trained model due to large size of dataset. The memory usage is higher.
[8]	Sign Language Recognition via Skeleton-Aware Multi-Model Ensemble [2021]	Songyao Jiang, Bin Sun, Lichen Wang, Yue Bai, Kunpeng Li, Yun Fu	SSTCN(Separable Spatial-Temporal Convolution Network),	Skeleton-based methods are smarter choices over RGB-based methods.	The offline full-body Pose estimator may fail due to off-screen or occlusion especially for finger.
[9]	Sign Language Recognition For Sentence-Level Continues Signing	Ishika Godge, Ruvan Weerasignhe, Damitha Sandaruwan	EMG(electromyography)sensors.	accuracy was up to 75-80% word-level	For sentence recognition accuracy was poor.

V. CONCLUSION

The above article demonstrates some of the techniques listed in building a sign language recognition model that converts hand signs into their corresponding alphabets and digits based on standard languages such as American Sign Language, Indian Sign Language, Japanese Sign Language, and Turkish Sign Language. After a Closer look at the above research Paper of Sign Language Recognition System it is observed that the most widely used data acquisition component were camera and Kinect. Most of the work on sign language recognition systems has been performed for static characters that have been already captured and isolated sign respectively.it has been observed that the majority of work has been performed using single handed signs for different sign language systems. It has been found that the most of the work has been performed using Convolutional neural networks which is used for image recognition and tasks that involve for image processing.

REFERENCES

[1] Ahmed Kasapbasi ,Ahmed Eltayeb ,Ahmed Elbushra, Omar Al-Hardanee , Arif Yilmaz "A CNN based human computer interface for American Sign Language recognition for hearing-impairedindividuals.",2022.https://www.sciencedirect.com/science/article/pii/S2666990021000471?via%3Dihub
[2] Kanchon K. Podder Muhammad E.H. Chowdhury Anas M. Tahir Zaid Bin Mahbub Md Shafayet Hossain Muhammad Abdul Kadir, "Bangla Sign Language (BdSL) Alphabets and Numerals Classification Using a Deep Learning Model",2022 https://www.mdpi.com/1424-8220/22/2/574
[3] Pooja M.R Meghana M Praful Koppalkar Bopanna M J Harshith Bhaskar Anusha Hullali, "Sign Language Recognition System", 2022. ijsepm.C9011011322.
[4] Bekhzod Olimov, Shraddha M. Naik, Sangchul Kim, Kil-Houm Park & Jeonghong Kim "An integrated mediapipe-optimized GRU model for Indian sign language recognition", 2022. https://www.nature.com/articles/s41598-022-15998-7
[5] Satwik Ram Kodandaram, N. Pavan Kumar, Sunil Gl,"Sign Language Recognition",2021. https://www.researchgate.net/publication/354066737_Sign_Language_Recognition
[6] Mathieu De Coster, Mieke Van Herreweghe, Joni Dambre, "Isolated Sign Recognition from RGB Video using Pose Flow and Self-Attention",2021. CVPRW_2021
[7] Arpita Haldera , Akshit Tayadeb, "Real-time Vernacular Sign Language Recognition using MediaPipe and Machine Learning", 2021. IJRPR462.
[8] Songyao Jiang, Bin Sun, Lichen Wang, Yue Bai, Kunpeng Li and Yun Fu, "Sign Language Recognition via Skeleton-Aware Multi-Model Ensemble", 2021. 2110.06161v
[9] Ishika Godage, Ruvan Weerasignhe and Damitha Sandaruwan "Sign Language Recognition For Sentence-Level Continues Signing", 2021. csit112305
[10] N. Mukai, N. Harada, and Y. Chang, "Japanese Fingerspelling Recognition Based on Classification Tree and Machine Learning," 2017 Nicograph International (NicoInt), Kyoto, Japan, 2017, pp. 19-24.doi:10.1109/NICOInt.2017
[11] Jayshree R. Pansare, Maya Ingle, "Vision-Based Approach for American Sign Language Recognition Using Edge Orientation Histogram", International Conference on Image, Vision and Computing, pp.86-90, 2016.
[12] Nagaraj N. Bhat, Y V Venkatesh, Ujjwal Karn, Dhruva Vig, "Hand Gesture Recognition using Self Organizing Map for Human-Computer Interaction", International Conference on Advances in Computing, Communications, and Informatics, pp.734-738, 2013.



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)