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.2024.62528

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 Detection using CNN Model

Prof. Kirti Patil[1], Shejal Kale[2], Sakshi Devkar[3], Prachi Dhamane[3], Vaishnavi Khandagale[5]
[1]*Assistant Professor,* [2, 3, 4, 5]*UG Students, Department of Computer Engineering, Alard College of Engineering and Management Pune*

Abstract: ***Sign Language is mainly used by deaf (hard hearing) and dumb people to exchange information between their own community and with other people. It is a language where people use their hand gestures to communicate as they can't speak or hear. Sign Language Recognition (SLR) deals with recognizing the hand gestures acquisition and continues till text or speech is generated for corresponding hand gestures. Here hand gestures for sign language can be classified as static and dynamic. Deep Learning Computer Vision is used to recognize the hand gestures by building Deep Neural Network architectures (Convolution Neural Network Architectures) where the model will learn to recognize the hand gestures images over an epoch.***
Keywords: Sign Language, CNN, Deep Learning, Dumb, Communication.

I. INTRODUCTION

People with impaired speech and hearing uses Sign language as a form of communication. Disabled People use this sign language gestures as a tool of non-verbal communication to express their own emotions and thoughts to other common people. But these common people find it difficult to understand their expression. To address this issue, Convolutional neural network of 11 layers is constructed, four Convolution layers, three Max-Pooling Layers, two dense layers, one flattening layer and one dropout layer is used. The American Sign Language Dataset from MNIST to train the model to identify the gesture. The dataset contains the features of different augmented gestures. Introduced a custom CNN (Convolutional Neural Network) model to identify of the sign from a video frame using Open-CV. MNIST dataset is a pre-processed dataset of RGB images that are converted to grayscale. The converted image is then scaled in respect to the size of the images with which the model was trained. The image is fed into the pre-trained custom CNN model post scaling and transformation. The gesture prediction from the CNN model is obtained and post that, it is classified based on the categorical label. The classified gesture is displayed as text. A custom CNN model with multi layers is used. The gesture image segmented from the video frame is then converted to a grayscale image.



Figure 1. Predicted signs

©IJRASET: All Rights are Reserved

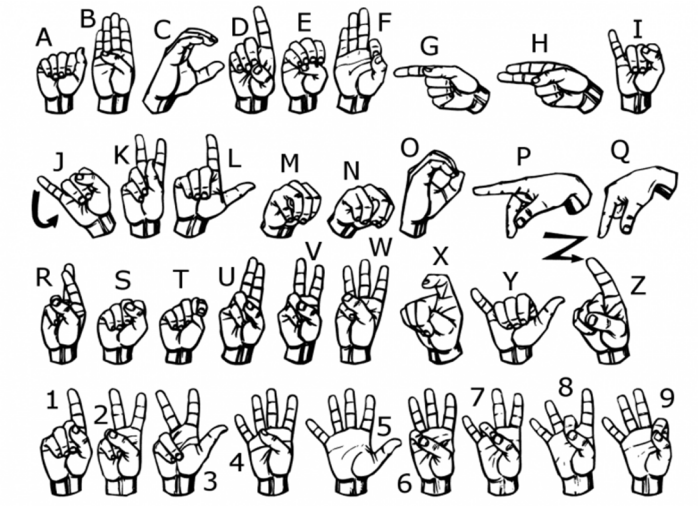


Figure 2. Signs data set

II. RELATED WORK

Sign Language Recognition Based on CNN and Data Augmentation its author G. Li, X. Wang and Y. Liu [1] designed SLR system based on CNN has been proposed for recognizing American sign language (ASL) letters. enhance the model's ability to generalize and mitigate overfitting, the application of data augmentation techniques is applied.

Action Detection for Sign Language using Machine Learning its author A. S Sushmitha Urs, V. B Raj, P. S, P. Kumar K, M. B R and V. Kumar S [2]. This study focuses on American Sign Language (ASL), implementing convolutional neural networks (CNNs) for real-time sign language recognition using monochromatic camera images. The impact of ASL extends beyond communication, promoting inclusivity, empowerment, and societal effects on education and careers.

Sign Language Digits Recognition Technology Based on a Convolutional Neural Network its author O. Voloshynskyi, V. Vysotska, R. Holoshchuk, S. Goloshchuk, S. Chyrun and D. Zahorodnia [3]. This paper employs a convolutional neural network (CNN) to address digit recognition in sign language images, using a dataset of 2062 64x64 pixel images.

Real-time Indian Sign Language Recognition using Skeletal Feature Maps its author V. Kumar, R. Sreemathy, M. Turuk, J. Jagdale and A. Agarwal [4]. This study addresses sign language recognition, a well-explored problem in the context of deep learning. While many models claim high validation accuracy but struggle with real-time performance, the proposed model achieves consistent real-time predictions using Skeletal Feature Maps generated from key points via Media Pipe.

Development of Sign Language Translator for Disable People in Two-ways Communication its author P. Singh, S. V. S. Prasad, R. Singh, K. Dasari and B. L. Prasanna [5]. This paper presents a prototype aimed at aiding the communication of deaf and dumb individuals with the broader community.

Deep Learning-based Methods for Sign Language Recognition: A Comprehensive Study its author Adaloglou, Nikolaos M, et al [6]. Computer vision-based sign language recognition systems are subjected to a comparative experimental assessment. The most recent deep neural network techniques in this field are used to perform a complete evaluation on a variety of publicly available datasets. By mapping non segmented video streams to glosses, the purpose of this study is to learn more about sign language recognition.

The ArSL Database and Pilot Study: Towards Hybrid Multimodal Manual and Non-Manual Arabic Sign Language Recognition its author Luqman, Hamzah, and El-Sayed M. El-Alfy [7]. A new multi-modality ArSL dataset that combines many modalities. It comprises of 6748 video samples recorded using Kinect V2 sensors of fifty signs performed by four signers. In addition, we used state-of-the-art deep learning algorithms to analyse the integration of spatial and temporal characteristics of distinct modalities, both manual and non-manual, for sign language identification.

Deep Learning Multi Stroke Thai Finger Spelling Sign Language Recognition System its author Pariwat, Thongpan, and Pusadee Seresangtakul [8]. On a complicated backdrop, a vision-based approach was used to accomplish semantic segmentation with dilated convolution for hand segmentation, optical flow separation for hand strokes, and learning feature and classification with a CNN. The five CNN structures

Implementing k-Nearest Neighbours with Dynamic Time Warping and Convolutional Neural Network Algorithms in Wearable Electronics for Sign Language Recognition its author Saggio, Giovanni, et al [9]. a wearable electronics-based sign language recognition device with two separate categorization methods A sensory glove and inertial measurement units were used to collect finger, wrist, and arm/forearm motions for the wearable electronics. k-Nearest Neighbours with Dynamic Time Warping (a nonparametric technique) and Convolutional Neural Networks were used as classifiers.

For static signs, a deep learning-based sign language recognition system its author Wadhawan, Ankita, and Parteek Kumar [10]. The paper discusses the use of deep learning-based convolutional neural networks to represent robust static signs in the context of sign language recognition (CNN). A total of 35,000 sign photos of 100 static signs were gathered from various users for this study. On around 50 CNN models, the suggested system's efficiency is assessed.

III. PROPOSED SYSTEM

A Convolutional Neural Network (CNN) is a type of deep neural network used for image recognition and A specific kind of such a deep neural network is the convolutional network, which is commonly referred to as CNN or ConvNet. It's a deep, feed-forward artificial neural network. Remember that feed-forward neural networks are also called multi-layer perceptron's (MLPs), which are the quintessential deep learning models. Image classification, object detection, segmentation, face recognition

Pooling is a technique commonly used to down sample or reduce the spatial dimensions of an image while retaining its essential features. Max pooling involves dividing the input image into non-overlapping rectangular regions (usually 2x2 or 3x3) and selecting the maximum value from each region. The selected maximum values form the output of the pooling layer. Max pooling helps retain the most prominent features in each region and discard less relevant information. Average pooling, on the other hand, involves taking the average value of the pixels within each region. it helps down sample the image, but instead of selecting the maximum value, it uses the average. Average pooling tends to smooth out the features in the down sampled image. Activation Function introduce non-linearity to the network, enabling it to learn the complex relationships in the data. Sigmoid function: we utilize a nonlinear transformation of the linear regression equation to give values between 0 and 1; the sigmoid function $s(x) = 1/(1 + e^{-x})$, where $e \approx 2.71$. Hyperbolic tangent function: The tanh activation function will yield −1 or values very close to −1 when outputs of a neurons are large negatives $f(z) = \tanh(z) = 2\sigma(2z) - 1$.

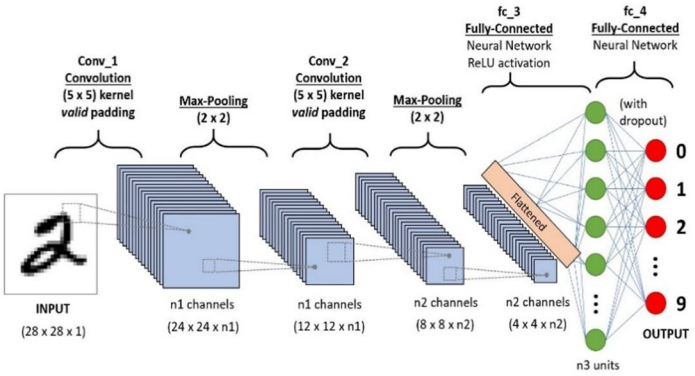


Figure 3. CNN Architecture and layers

A. *Technologies and Tools*

1) Python
2) TensorFlow
3) Keras
4) OpenCV

B. *Setup*

Use command promt to setup environment by using install_packages.txt and install_packages_gpu.txt files.
python -m pip r install_packages.txt
import NumPy as np
import pickle
import cv2, os

```
from glob import glob
from keras import optimizers
from keras.models import Sequential
from keras.layers import Dense
from keras.layers import Dropout
from keras.layers import Flatten
from keras.layers.convolutional import Conv2D
```

C. Training Model

```
def train():
with open("train_images", "rb") as f:
train_images = np.array(pickle.load(f))
with open("train_labels", "rb") as f:
train_labels = np.array(pickle.load(f), dtype=np.int32)
```

D. Algorithm

1) Added gestures and label them using OpenCV which uses webcam feed. by running create_gestures.py and stores them in a database.
2) Add different variations to the captured gestures by flipping all the images by using Rotate_images.py.
3) Run load_images.py to split all the captured gestures into training, validation and test set.
4) To view all the gestures, run display gestures.py.
5) Train the model using Keras by running cnn_model_train.py.
6) Run final.py. This will open up the gesture recognition window which will use your webcam to interpret the trained American Sign Language gestures.

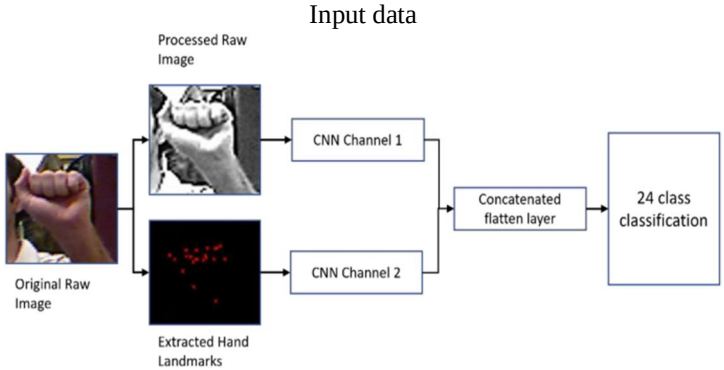
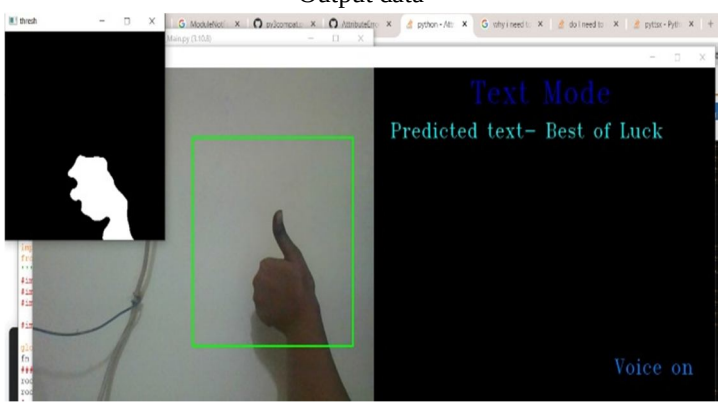


Figure 4. Image processing through CNN layers



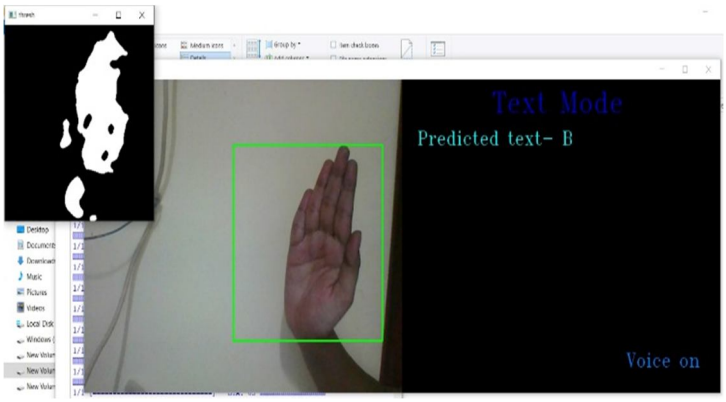


Figure 5. System output with threshold image

E. Result of Sign Detection System

Table.1. Accuracy Table

No.	Sign	Correct	Incorrect	Accuracy (%)
1	A	17	3	85.00
2	B	16	4	80.00
3	C	18	2	90.00
4	D	18	2	90.00
5	E	14	6	70.00
6	F	18	2	90.00
7	G	15	5	75.00
8	H	17	3	85.00
9	I	18	2	90.00
10	J	17	3	85.00
11	K	18	2	90.00
12	L	17	3	85.00
13	M	18	2	90.00
14	N	16	4	80.00
15	O	20	0	100.00
16	P	16	4	80.00
17	Q	15	5	75.00
18	R	17	3	85.00
19	S	14	6	70.00
20	T	17	3	85.00
21	U	18	2	90.00
22	V	18	2	90.00
23	W	15	5	75.00
24	X	16	7	80.00
25	Y	15	5	75.00
26	Z	18	2	90.00
27	Best of Luck	20	0	100.00
28	You	17	3	85.00
29	I/Me	18	2	90.00
30	Like	20	0	100.00
31	Remember	16	4	80.00

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

32	Love	19	1	95.00
33	All the best	16	4	80.00
34	I love you	19	1	95.00
35	0	20	0	100.00
36	1	18	2	90.00
37	2	16	4	80.00
38	3	17	3	85.00
39	4	15	5	75.00
40	5	18	2	90.00
41	6	17	3	85.00
42	7	16	4	80.00
43	8	17	3	85.00
44	9	15	5	75.00
Accuracy Average				85.22

IV. CONCLUSION

Sign language detection project in Python successfully recognizes and interprets sign language gestures through the use of computer vision techniques and machine learning algorithms, we've created a tool that can understand and translate sign language into text or speech.it opens doors for better communication accessibility for the deaf and hard of hearing community, bridging gaps and fostering inclusivity in our society.

Sign language provides children with an alternative way to make themselves understood. This extra tool enables them to express how they feel, their thoughts and wants, so that they can take part in learning and social activities. This not only gives a child a 'voice' but is also important when building relationships. So, the primary goal of gesture recognition research is to create systems, which can identify specific human gestures and use them, for example, to convey information. For that, vision- based

hand gesture interfaces require fast and extremely robust hand detection, and gesture recognition in real time. Sign Language detection system shows what the position of hands in viewfinder of camera module means with good accuracy. It can then be used to help people who are just beginning to learn Sign Language or those who don't know sign language but have a close one who is deaf.

V. FUTURE SCOPE

In future work, proposed system can be developed and implemented using Raspberry Pi. Image Processing part should be improved so that System would be able to communicate in both directions i.e.it should be capable of converting normal language to sign language and vice versa. To develop manpower for using Indian Sign Language (ISL) and teaching and conducting research in ISL, including bilingualism. To promote the use of Indian Sign Language as educational mode for deaf students at primary, secondary and higher education levels. A) Enhanced Accuracy: Machine learning models will become even better at recognizing nuanced sign language gestures, reducing recognition errors and improving overall accuracy.

B) Wider Implementation: Sign language recognition systems will become more prevalent, integrating into everyday devices and platforms, including smartphones, tab.

C) Educational role: Teachers are constantly searching for new ways to engage their students in the learning process. Using sign language within the classroom is one solution to reach all learners. Sign language can enhance the learning process by bringing visual, auditory and kinaesthetic feedback to help reach all students, and smart home assistants.

REFERENCES

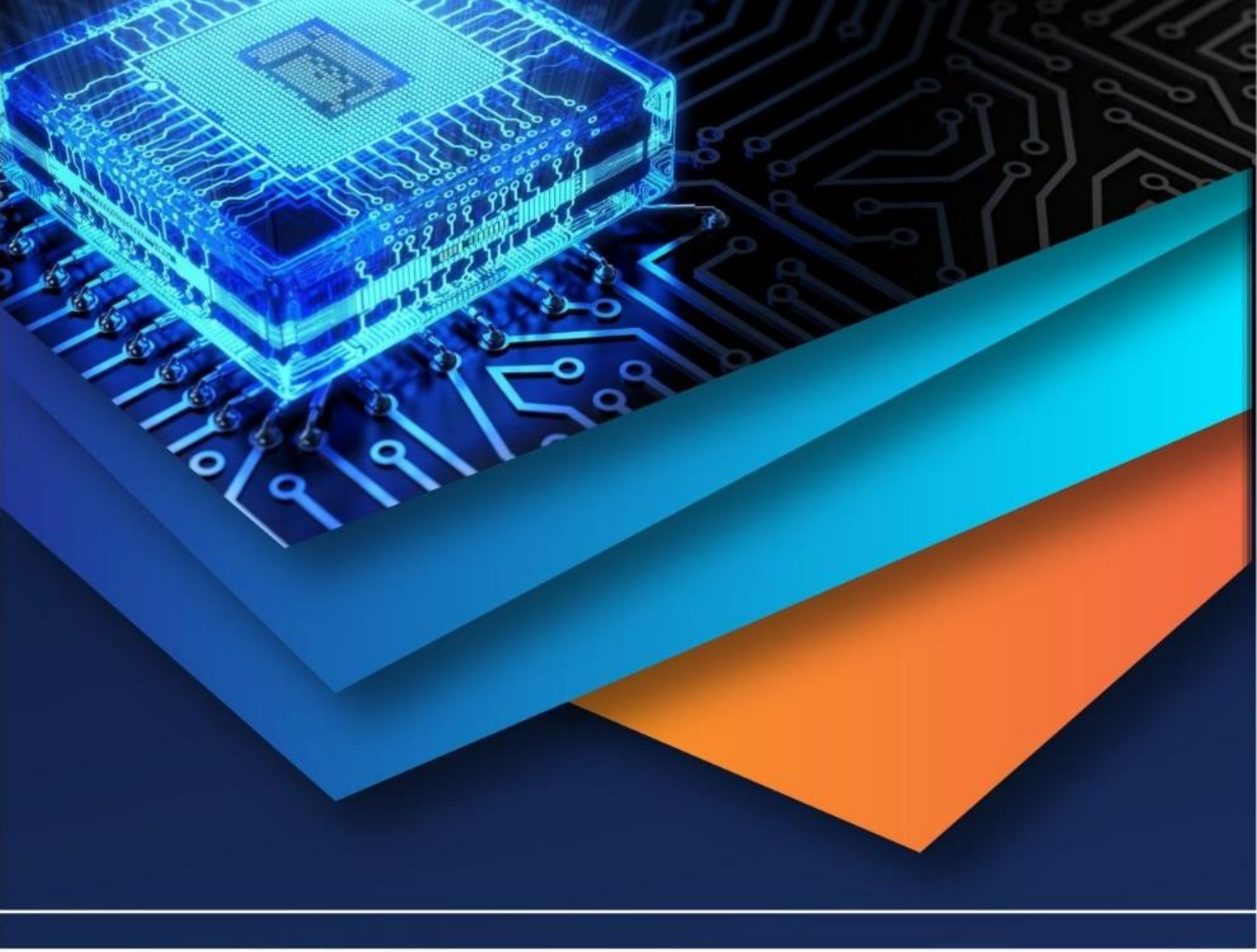
[1] G. Li, X. Wang and Y. Liu, "Sign Language Recognition Based on CNN and Data Augmentation," 2023 IEEE 5th International Conference on Civil Aviation Safety and Information Technology (ICCASIT), Dali, China, 2023.

[2] A. S Sushmitha Urs, V. B Raj, P. S, P. Kumar K, M. B R and V. Kumar S, "Action Detection for Sign Language Using Machine Learning," 2023 International Conference on Network, Multimedia and Information Technology (NMITCON), Bengaluru, India, 2023.

[3] O. Voloshynskyi, V. Vysotska, R. Holoshchuk, S. Goloshchuk, S. Chyrun and D. Zahorodnia, "Sign Language Digits Recognition Technology Based on a Convolutional Neural Network," 2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Dortmund, Germany, 2023.

[4] V. Kumar, R. Sreemathy, M. Turuk, J. Jagdale and A. Agarwal, "Real-time Indian Sign Language Recognition using Skeletal Feature Maps," 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023.
[5] P. Singh, S. V. S. Prasad, R. Singh, K. Dasari and B. L. Prasanna, "Development of Sign Language Translator for Disable People in Two-Ways Communication," 2023 1st International Conference on Circuits, Power and Intelligent Systems (CCPIS), Bhubaneswar, India, 2023.
[6] Deep Learning-based Methods for Sign Language Recognition: A Comprehensive Study its author Adaloglou, Nikolaos M, et al. Computer vision-based sign language recognition systems are subjected to a comparative experimental assessment.
[7] The ArSL Database and Pilot Study: Towards Hybrid Multimodal Manual and Non-Manual Arabic Sign Language Recognition its author Luqman, Hamzah, and El-Sayed M. El-Alfy A new multi-modality ArSL dataset that combines many modalities.
[8] Pariwat, Thongpan, and Pusadee Seresangtakul. "Multi-Stroke Thai Finger-Spelling Sign Language Recognition System with Deep Learning." Symmetry 13.2 (2021): 262.
[9] Saggio, Giovanni, et al. "Sign language recognition using wearable electronics: implementing k-nearest neighbors with dynamic time warping and convolutional neural network algorithms." Sensors 20.14 (2020): 3879.
[10] Wadhawan, Ankita, and Parteek Kumar. "Deep learning-based sign language recognition system for static signs." Neural Computing and Applications 32.12 (2020): 7957-7968.

©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)