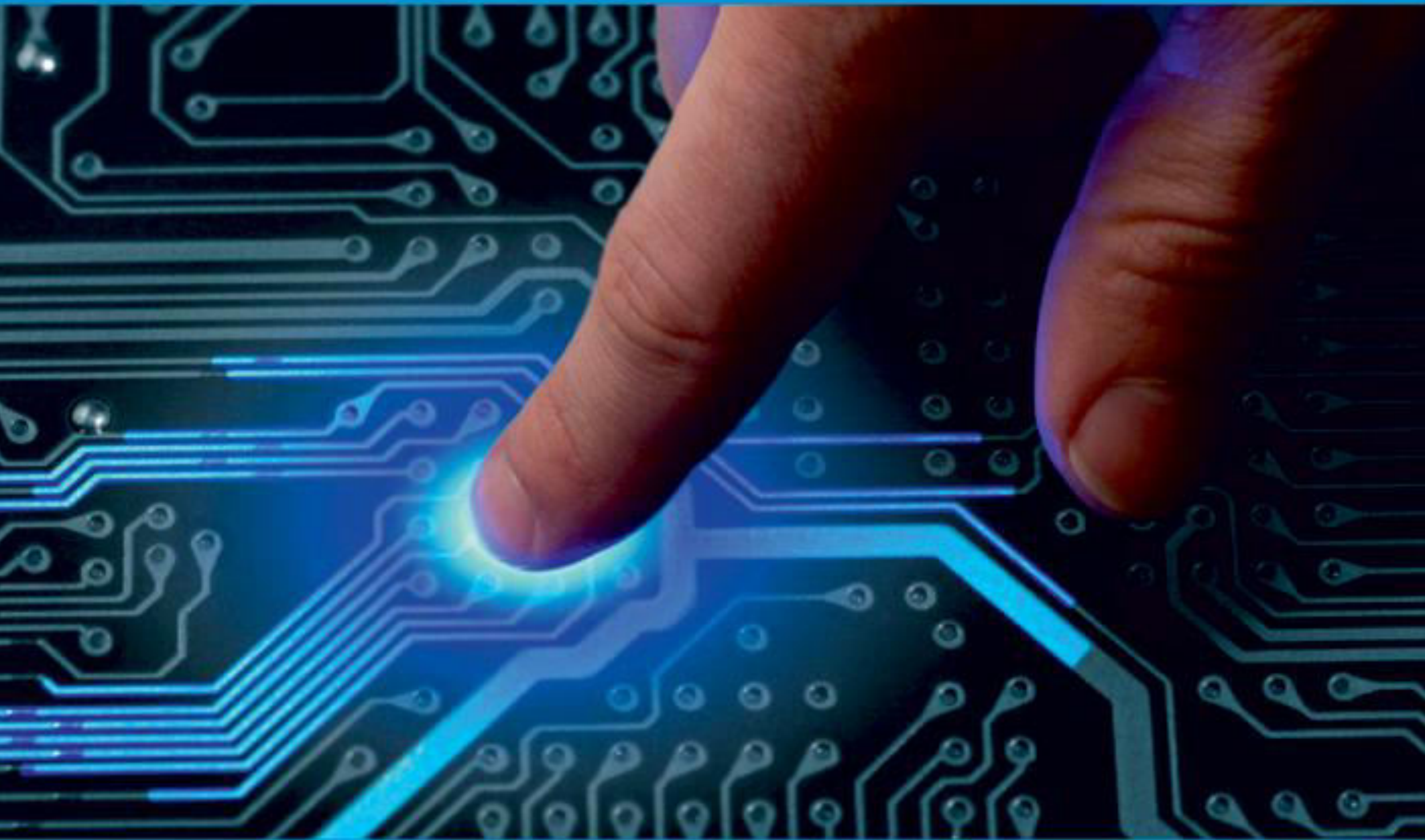
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 5, May 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005252 |

Sign Language Recognition Using Python and OpenCV

Y. Suresh [1], Mandadapu Harika[2], Konduru Anusha[3], Kikkuri Haritha[4],

Mannava Siva Sai Navya Sri[5]

Assistant Professor, Department of IT, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, Andhra Pradesh, India[1]

UG Student, Department of IT, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, Andhra Pradesh, India[2]

UG Student, Department of IT, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, Andhra Pradesh, India[3]

UG Student, Department of IT, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, Andhra Pradesh, India[4]

UG Student, Department of IT, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, Andhra Pradesh, India[5]

ABSTRACT: Sign Language is used by deaf and dumb people for communication. There have been several advancements in technology, and a lot of research happening to help the people who are deaf and dumb. Aiding the cause, with the help of Deep Learning we can make an impact on this cause.

A sign detector is developed in this project to detect signs shown by hand and identify alphabets from A-Z using Convolutional Neural Networks.
So, the people having the disability if deaf or dumb can connect with every person having all the emotions and intentions in which they don't face any complications in interacting with other people.

KEYWORDS: - Hand gesture, Sign language, Communication, OpenCV, ANN, CNN

I. INTRODUCTION

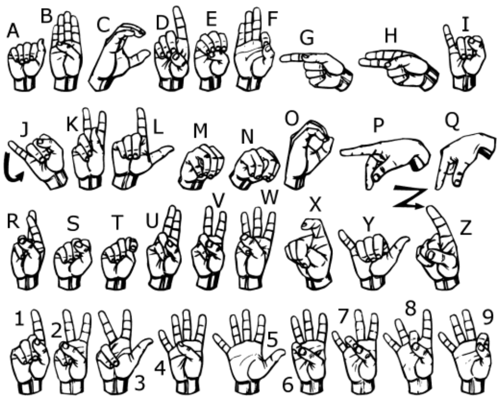
Communication is very crucial to human beings, as it enables us to express ourselves. We communicate through speech, gestures, body language, reading, writing, or through visual aids, speech being one of the most commonly used among them. However, unfortunately, for the speaking and hearing impaired minority, there is a communication gap. Visual aids, or an interpreter, are used for communicating with them. However, these methods are rather cumbersome and expensive, and can't be used in an emergency. Sign Language chiefly uses manual communication to convey meaning. This involves simultaneously combining hand shapes, orientations, and movement of the hands, arms, or body to express the speaker's thoughts.

Sign Language consists of fingerspelling, which spells out words character by character, and word level association which involves hand gestures that convey the word's meaning. Fingerspelling is a vital tool in sign language, as it enables the communication of names, addresses, and other words that do not carry meaning in the word-level association. In spite of this, fingerspelling is not widely used as it is challenging to understand and difficult to use. Moreover, there is no universal sign language and very few people know it, which makes it an inadequate alternative for communication.

A system for sign language recognition that classifies fingerspelling can solve this problem. Various machine learning algorithms are used and their accuracy is recorded and compared in this report.

II. OBJECTIVES

This project aims at identifying alphabets in Sign Language from the corresponding gesture. Gesture recognition and sign language recognition has been a well researched topic for American Sign Language but has been rarely touched for its Indian counterpart. We aim to tackle this problem but instead of using high end technology like gloves or kinect for gesture recognition, we aim at recognition from images(which can be obtained from say webcam) and then use computer vision techniques and machine learning techniques for extracting relevant features and subsequent classification.



The proposed system architecture, which consists of two modules,namely:dataacquisition,pre-processingandfeature extraction and sign language gestureclassification.

III. MOTIVATION

Communication is one of the basic requirement for survival in society. Deaf and dumb people communicate among themselves using sign language but normal people find it difficult to understand their language. Extensive work has been 1 done on American sign language recognition but Indian sign language differs significantly from American sign language.ISL uses two hands for communicating(20 out of 26) whereas ASL uses single hand for communicating. Using both hands often leads to obscurity of features due to overlapping of hands. In addition to this, lack of datasets along with variance in sign language with locality has resulted in restrained efforts in ISL gesture detection. Our project aims at taking the basic step in bridging the communication gap between normal people and deaf and dumb people using Indian sign language. Effective extension of this project to words and common expressions may not only make the deaf and dumb people communicate faster and easier with outer world, but also provide a boost in developing autonomous systems for understanding and aiding them

IV. LITERATURE SURVEY

The researches done in this field are mostly done using a glove based system. In the glove based system, sensors such as potentiometer, accelerometers etc. are attached to each of the finger. Based on their readings the corresponding alphabet is displayed. Christopher Lee and Yangsheng Xu developed a glove-based gesture recognition system that was able to recognize 14 of the letters from the hand alphabet, learn new gestures and able to update the model of each gesture in the system in online mode.

Tanuj Bohra et al. [1] proposed a real-time two-way sign language communication system built using image processing, deep learning and computer vision. Techniques such as hand detection, skin color segmentation, median blur and contour detection are performed on images in the dataset for better results. CNN model trained with a large dataset for 40 classes and was able to predict 17600 test images in 14 seconds with an accuracy of 99%.

Joyeeta Singha and Karen Das [2] proposed a system for indian sign language recognition from a live video. The system comprises of three stages. Preprocessing stage includes skin filtering and histogram matching. Eigen values and eigen vectors are being considered for feature extraction stage and Eigen value weighted euclidean distance

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005252 |

for classification. Dataset consisted 480 images of 24 signs of ISL signed by 20 people. System was tested on 20 videos and achieved an accuracy of 96.25%.

Oscar Kellar et al. [4] introduced a hybrid CNN-HMM for sign language recognition. They conducted experiments on three datasets namely RWTH-PHOENIX-Weather 2012 [19], RWTH-PHOENIX-Weather Multisigner 2014 [20] and SIGNUM single signer [21]. Training and validation set have a ratio of 10 to 1. After the CNN training is finished a softmax layer is added and results are used in HMM as observation probabilities.

G. Anantha Rao et al. [5] proposes an indian sign language gesture recognition using convolutional neural network. This system works on videos captured from a mobile's front camera. Dataset is created manually for 200 ISL signs. CNN training is performed with 3 different datasets. In the first batch, dataset of only one set is given as input. Second batch has 2 sets of training data and third batch respectively has 3 sets of training data. Average recognition rate of this CNN model is 92.88%.

V. EXISTING SYSTEM

In the existing glove-based method the signer has to wear a hardware glove. This glove is attached with several sensors. For every hand movement made the corresponding sensors will work and the hand sign is conveyed to the opposite person. The biggest disadvantage of this method is the person should be carrying this glove with them in order to communicate. The cost of building this glove can be relatively high because of the sensors used and the lifetime cannot be guaranteed.

VI. PROPOSED MODEL

Vision-based method, further classified into static and dynamic recognition. Statics deals with the detection of static gestures(2d-images) while dynamic is a real-time live capture of the gestures. This involves the use of the camera for capturing movements.

In the proposed system, the dumb person should provide a gesture or a sign in front of a camera in the designated area. The system would capture and pre-process the gesture and then send the preprocessed frame to a trained CNN model. The CNN model predicts the gesture and shows the output on the screen in text format.

VII. METHODOLOGY

The system is a vision-based approach. All the signs are represented with bare hands and so it eliminates the problem of using any artificial devices for interaction.

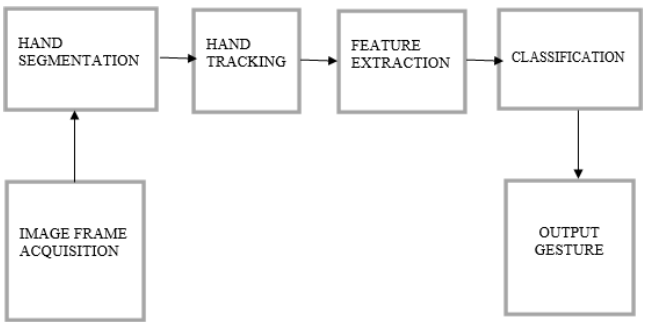


Fig 1: System Architecture

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005252 |

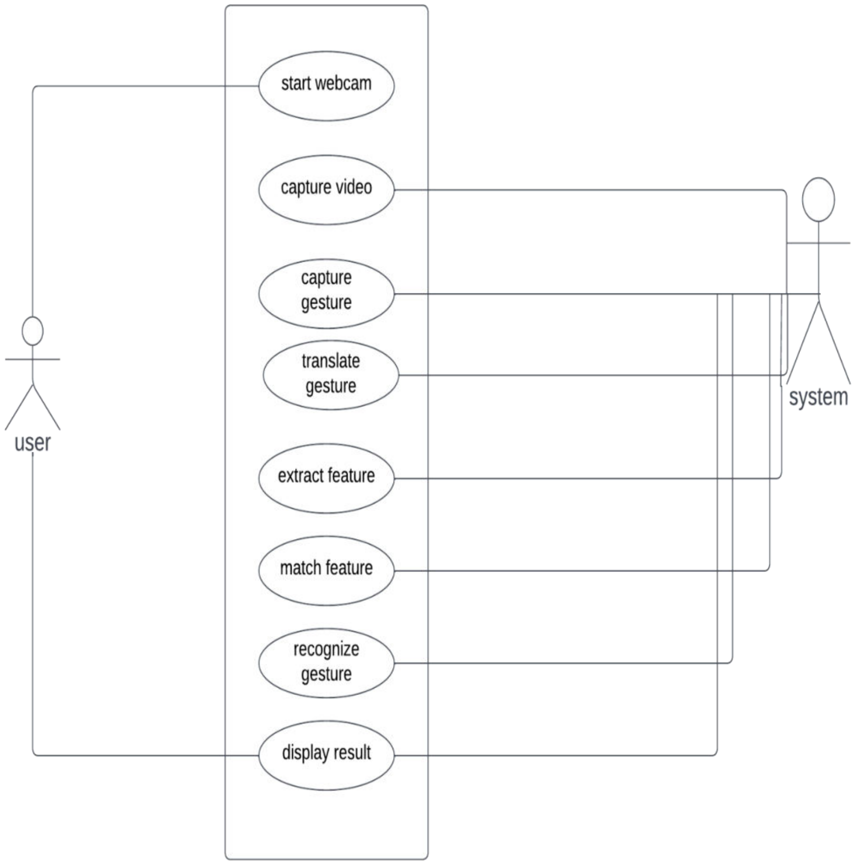


Fig 2 UseCase Diagram

A. DATASET
Dataset used to train the CNN model contains 27455 images in the training set and 7172 images in the test set which belongs to 24 classes. These 24 classes are English alphabets except J and Z. The dataset used is in CSV format with the label column being the output column that is the alphabet and columns pixel1 to pixel784 being the image data. Each row in the dataset is a gray-scale image of size 28 x 28. This 28 x 28 2-d matrix is flattened to get the 784 columns in the dataset.

B. DATA PREPARATION
Pandas framework is used to read data from CSV files. Columns from pixel1 to pixel784 are reshaped to 28x28 2d vector using reshape method of pandas. Then this 2d vector is divided by 255 such that each pixel value ranges from 0-1. The resulting vector is fed into CNN for training.

C. TRAINING ANDTESTING
We convert our input images (RGB) into grayscale and apply gaussian blur to remove unnecessary noise. We apply adaptive threshold to extract our hand from the background and resize our images to 128 x 128. We feed the input images after pre- processing to our model for training and testing after applying all the operations.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |
|| Volume 10, Issue 5, May 2022 ||
| DOI: 10.15680/IJIRCCE.2022.1005252 |

VIII. RESULTS & DISCUSSION

```
[26] plt.plot(history.history['accuracy'])
     plt.plot(history.history['val_accuracy'])
     plt.title("Accuracy")
     plt.xlabel('epoch')
     plt.ylabel('accuracy')
     plt.legend(['train','test'])

     plt.show

<function matplotlib.pyplot.show>
```

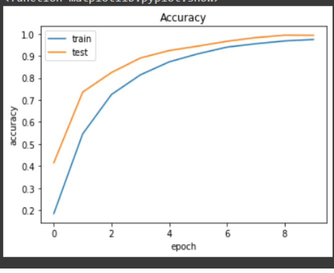


Fig 3 Training Accuracy vs Validation Accuracy

Model: "sequential"

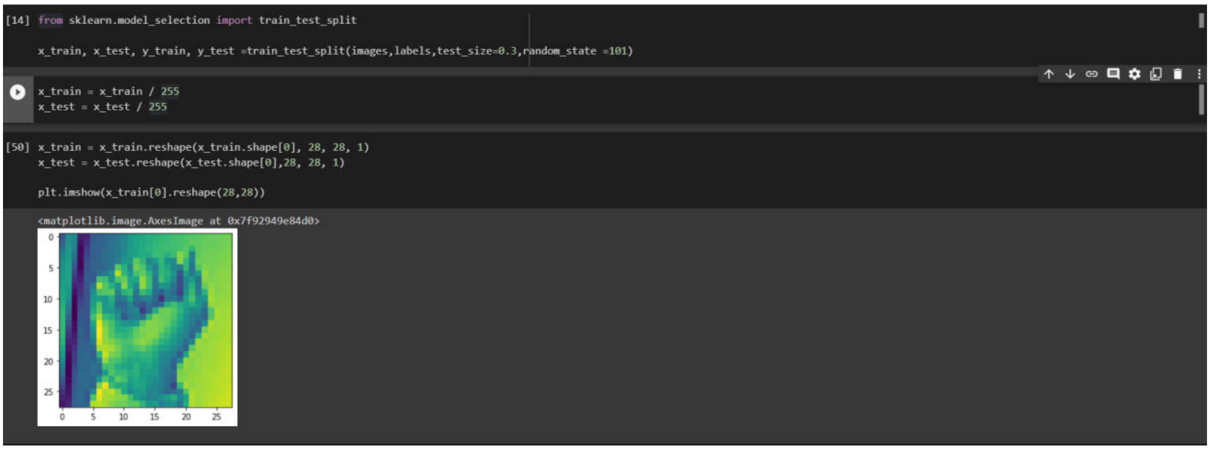
Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 26, 26, 64)	640
max_pooling2d (MaxPooling2D)	(None, 13, 13, 64)	0
conv2d_1 (Conv2D)	(None, 11, 11, 64)	36928
max_pooling2d_1 (MaxPooling2	(None, 5, 5, 64)	0
conv2d_2 (Conv2D)	(None, 3, 3, 64)	36928
max_pooling2d_2 (MaxPooling2	(None, 1, 1, 64)	0
flatten (Flatten)	(None, 64)	0
dense (Dense)	(None, 128)	8320
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 24)	3096

Total params: 85,912
Trainable params: 85,912
Non-trainable params: 0

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005252 |



IX. CONCLUSION

In this paper, a survey on sign language recognition is presented and various techniques have been studied and analysed for the same. In recognition process, segmentation plays a crucial part in which skin region is separated from the background which usually affects the recognition accuracy. Besides segmentation, classification also depends on the feature extraction techniques which performs dimensionality reduction and reduces the computation cost. Study of various classification techniques concludes that deep neural network (CNN, Inception model, LSTM) performs better than traditional classifiers such as KNN and SVM.

REFERENCES

[1]. H. Muthu Mariappan and V. Gomathi, "Real-Time Recognition of Indian Sign Language," 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, India, 2019, pp. 1-6, doi: 10.1109/ICCIDS.2019.8862125.

[2]. Mathavan Suresh Anand, Nagarajan Mohan Kumar, Angappan Kumaresan, “An Efficient Framework for Indian SignLanguage Recognition Using Wavelet Transform” Circuits and Systems, Volume 7, pp 1874- 1883, 2016.

[3]. Kumar Mahesh, “Conversion of Sign Language into Text,” International Journal of Applied Engineering Research ISSN 0973- 4562 Volume 13, Number 9 (2018) pp. 7154-7161.

[4]. M. Xie and X. Ma, "End-to-End Residual Neural Network with Data Augmentation for Sign Language Recognition," 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chengdu, China, 2019, pp. 1629-1633, doi: 10.1109/IAEAC47372.2019.8998073.