International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

# Expressive Sign Language Recognition using CNN

**K.Ravi Raju[1], G.Madhavi[2] , B.Sravani[3], G.Sneha[4], G.Yasasiwani[5]**
*[1] Assistant Professor, Department of Computer Science and Engineering ,*
*Raghu Engineering College, Vishakhapatnam.*

*[2,3,4,5]B.Tech Students*
*[2,3,4,5]Department of Computer Science and Engineering , Raghu Institute of Technology ,Vishakhapatnam*

***Abstract: A major shortcoming in our society is a social barrier between the differently abled members of the society and the abled folks. One of the most important aspects of human beings, being regarded as social animals, is in fact communication. Communication is also a major obstacle faced by thehearing and vocal disabilities people. This inability to communicate leads to frequent problems and hinders the daily activities of a person with hearing and vocal disabilities.***
***The underlying reason for this disparity is that abled folks don't learn and aren't taught Sign Language which is the main means of communication for a person with hearing and vocal disabilities. Thus, abled folks are incapable of having a normally fluent conversation with these different sections of the society. Consequently, in a verbal exchange among hearing and speech impaired individuals and an able person the convenience of communique and consequently the consolation degree is hampered. So, in our project, we have proposed a cost-efficient solution to overcome this communication barrier. This solution can be easily used by everyone and can also be, with some modifications, made to work on most platforms which have a camera module. Our approach uses the integrated camera module to capture real time hand gestures based on hand key points or landmarks and the algorithm using machine learning techniques, displays the alphabet that the gesture is representing.***

***KEY WORDS**: Convolution Neural Network, Transfer Learning, Deep Learning, Artificial Neural Network, Support Vector Machines*

## I. INTRODUCTION

In the past few years, huge advancements have been made in the fields of science and technology. Not only this, technology has got much cheaper and its availability has widened as itis now available to the common man. So, it is vital to no longer overlook the duty of our generation to make use of this accessibility to technology to contribute to the progress and improvement of society at large.Human beings have, since the beginning of time, been described as a social animal. As a social being, one of the principal aspects of our life is communication. Social interaction or simply communication has always been regarded as one of the major aspects of living a happy life. For an individual to live a normal lifestyle, communication is necessary and is required for almost allof our daily tasks. But there is a not so blessed segment of society which faces hearing and vocal disabilities. A hearing-impaired individual is one who either can't hear at all or is able to hear sounds which are above a certain frequency, or what we'd generally call as 'can only hear when spoken to loudly'. An individual with the inability to speak due to any reason whatsoever is considered as a mute or silent person.

In an enormous research conducted in diverse domain names, it turned into determination that impairments such as hearing-impairment, vocal-impairment or the ineptitude to express oneself causes loss of opportunities for such people when compared to able people. Not only does it lead to this, but also hinders day-to-day activity of an individual such as normal conversations.

According to MoSPI, Govt. of India [1], in 2002 about 30.62 lakh of the then population were suffering from hearing disorder and 21.55 lakh of the then population were suffering from speech disorder.Another 2001 Census [2] states that around 21 million Indian citizens (which constituted 12.6 million males and 9.3 million females approximately), that is, about 2.1 per cent of the then population of India, were facing certain disabilities. People with speech disability accounted for the 7.5 per cent while those with hearing disability accounted for 5.8 per cent of these 21 million people in total.

These statistics also show evidence of the problems and discrimination faced by these people. Additionally, they also provide us with a wealth of facts about specific kinds of disabilities, the number of humans tormented by these disabilities and the barriers they face in their life. One of the foremost boundaries a disabled Individual faces in his existence is incapable of talking with an everyday man or woman.So with our knowledge of technology, we hope to help such people through our project so that they are able to communicate normally with others.

## II. LITERATURE SURVEY

**1.Sign Language Recognition System to aid Deaf-dumb People Using PCA**

Author: Shreyashi Narayan Sawant
Publication: International Journal of Computer Science & Engineering Technology, 2014 Summary:
26 gestures from Indian Sign Language, using MATLAB with WebCam live capture.260 images in total, 10 of each sign. Captured with white background so as to avoid illumination effects.Image segmentation using Otsu's method (a method for quantization), noise removal & contour smoothing.Centroid is calculated (using image moment) to separate the hand part from the arm.
Skin detection using HSV color model.Feature extraction using PCA (eigenvector matrix).Finally, subject gesture is normalized with respect to the average gesture and then projected onto gesture space using the eigenvector matrix. Finally, Euclidean distance is computed between this projection and all known projections. The minimum value of these comparisons is selected for recognition during the training phase.

**2.Hand Gesture Recognition for Deaf People Interfacing**
Author: Isaac Garcia Incertis, Jaime Gomez Garcia-Bermejo, Eduardo Zalama Casanova
Publication: International Journal of Computer Science & Engineering Technology, 2014 Summary:The work is carried on the static gesture case corresponding to the alphabet letters in Spanish Sign Language(LSE).
On captured images, hand regions and corresponding contours are extracted through color segmentation.Contours are Sampled at every arc distance.The resulting points are compared to those of a target gesture in the dictionary.The comparison is performed on the basis of four Distance Criteria- L0 norm, L1 norm , L2 Norm , L inf norm.A positive identification is assumed for the closest model.

**3.An American Sign Language Detection System using HSV Color Model and Edge Detection**
Author: A. Sharmila Konwar, B. Sagarika Borah, C. Dr.T.Tuithung
Publication: International Conference on Communication and Signal Processing, April 3-5, 2014, IndiaThe aim is to build a user-friendly human-computer interface in which the computer can understand the human language. .The system is described into two phases- training phase and testing phase. In the first phase , a database is created by capturing the images by web camera or any camera followed by preprocessing, feature extraction and training.In the second phase , image acquisition, preprocessing, feature extraction and classification which was based on the testing phase is doneCanny edge detection algorithm is used to detect hand gestures.PCA and ANN were used for the feature extraction and recognition part respectively.
The project was able to detect five alphabets successfully.
Due to the geometric variation and uneven background and lighting conditions, some images were not detected successfully.

**4. Real Time Hand Gesture Recognition Using Different Algorithms Based on American Sign Language**
Author: Md. Mohiminul Islam,Sarah Siddiqua,and Jawata Afnan3
Publication: IEEE Summary:
Different algorithms are used for the feature detection. Some of the algorithms are K convex hull for fingertip detection, pixel segmentation, eccentricity, elongatedness of object etc.Apart from the K-convex hull algorithm, many other algorithms were also used to get better accuracy.
Images are taken using a mobile phone, over that dataset ANN was used.The ANN is trained on 1850 sample images. The model was trained on the real time environment.The proposed model was able to detect ASL alphabets and numbers with the accuracy of 94.32%.Further, the model is improved for the movement detection of the hand for word recognition**.**

**OPEN CV &K-Nearest Neighbour**

Any machine learning based application can be summed up to have at least three phases - data collection and preprocessing phase, training phase and visualization.Our program also follows these steps in order. At first, the data is collected and a base dataset is prepared. This dataset is then divided into training data and testing data which in our case is a multi-label classification data as we have to predict 26 gestures. To generate our dataset, we've collected hand keypoints from images for each gesture using the laptop's web camera. Features are then selected and extracted from the training data. The next step is to decide which machine learning models to use. Since ours is a multi-class classification problem, the models used were K-Nearest Neighbours, Support Vector Machine and Decision Tree. These models are then trained on the training set. Then they are made to make predictions on the test set based on which their performance is evaluated and changes are made to the parameters so as to squeeze out the best results from these models.

**A.OPENCV**

OpenCV is short for Open Source Computer Vision Library which is a library for achieving actual-time programs. OpenCV is written in C and C++ and is cross-platform, that is, it works on all machines regardless of the operating system that is installed on that machine. And it is available as a library for languages such as Java, Python, C++, etc. As it is an open source project, it is freely available at http://sourceforge.net/projects/opencvlibrary/.

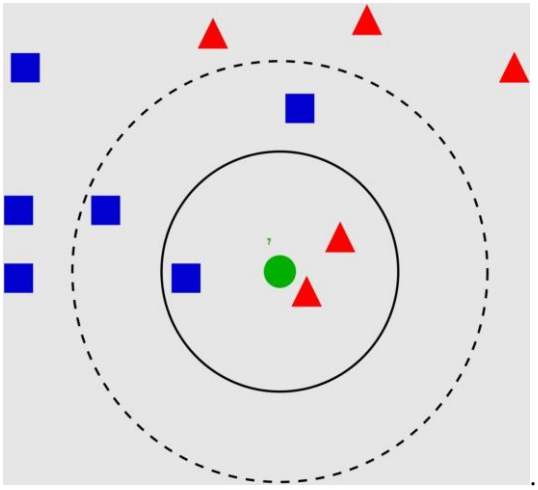
The creation of OpenCV is credited to Grad Bradsky of Intel. It was implemented in 1999, with only one mission in mind, that is, to encourage research in the field of computer vision and also to make computer vision available freely even for commercial applications.Being an open source project, OpenCV also has its documentation available on the web residing at http://opencv.willowgarage.com/documentation/index.htm

A digital image is an array of discrete values or a matrix of light intensities taken or capture by a device like camera and are organized into a two-dimensional matrix of pixels, in such a way where each of the pixels is represented by a number, generally ranging from 0 - 255 (255 due to it being 8-bit).

## B. k-Nearest Neighbours

k-Nearest Neighbours or simply k-NN, is a machine learning model that can be used for both classification and regression problems. The k-NN algorithm is based on the similarities between the dataset and the sample to be predicted and it puts the sample into the category that is the

closest to the categories present in the dataset. It stores the entire dataset and classifies the sample data based on similarity.



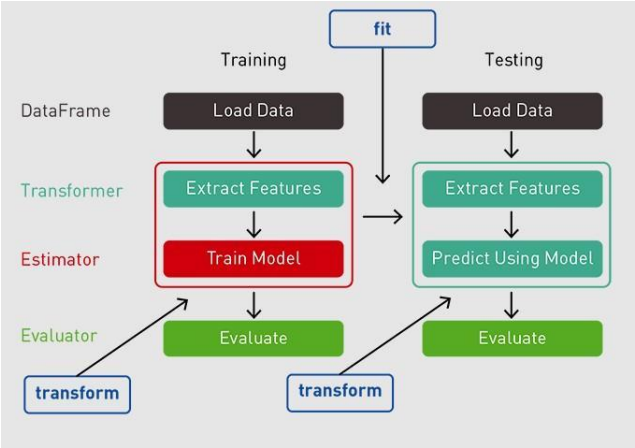
.

**III. METHODLOGY**

Any machine learning based application can be summed up to have at least three phases - data collection and preprocessing phase, training phase and visualization.

Our program also follows these steps in order. At first, the data is collected and a base dataset is prepared. This dataset is then divided into training data and testing data which in our case is a multi-label classification data as we have to predict 26 gestures. To generate our dataset, we've collected hand keypoints from images for each gesture using the laptop's web camera. Features are then selected and extracted from the training data. The next step is to decide which machine learning models to use. Since ours is a multi-class classification problem, the models used were K-Nearest Neighbours, Support Vector Machine and Decision Tree. These models are then trained on the training set. Then they are made to make predictions on the test set based on which

their performance is evaluated and changes are made to the parameters so as to squeeze out the best results from these models.



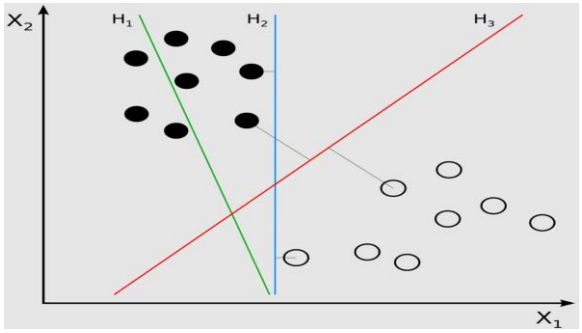

.

## Support Vector Machines

Support Vector Machines or simply SVM, is placed under the category of supervised machine learning algorithm. It is generally used for classification problems. In this model, each data item is plotted as a point in a n-D (here, n denotes the number of features in the training dataset) space and then it performs classification by identifying hyper-plane(s) that separates the two classes with the maximal distances.

## MediaPipe Hands

MediaPipe Hands employs machine learning to deduce 21 3-dimensional landmarks of the hand using just a single frame. Thus, it is regarded as a hi-fi and fairly accurate hand and finger detection and tracking solution as compared to current state-of-the-art models which generally rely on high performing machines. MediaPipe is available on various platforms even on the web and smartphones. MediaPipe Hands is also capable of inferring landmarks of both hands simultaneously.




**OPENCV**

OpenCV is short for Open Source Computer Vision Library which is a library for achieving actual-time programs. OpenCV is written in C and C++ and is cross-platform, that is, it works on all machines regardless of the operating system that is installed on that machine. And it is available as a library for languages such as Java, Python, C++, etc. As it is an open source project, it is freely available at http://sourceforge.net/projects/opencvlibrary/.

The creation of OpenCV is credited to Grad Bradsky of Intel. It was implemented in 1999, with only one mission in mind, that is, to encourage research in the field of computer vision and also to make computer vision available freely even for commercial applications.Being an open source project, OpenCV also has its documentation available on the web residing at http://opencv.willowgarage.com/documentation/index.htm

A digital image is an array of discrete values or a matrix of light intensities taken or capture by a device like camera and are organized into a two-dimensional matrix of pixels, in such a way where each of the pixels is represented by a number, generally ranging from 0 - 255 (255 due to it being 8-bit). .

**B. Technologies Used**

**1.Numpy:**

NumPy is a versatile Python package essential for scientific computing, offering a high-performance multidimensional array object and tools for efficient array manipulation. It serves as a fundamental tool for mathematical operations, data analysis, and integration with databases, making it a cornerstone in various fields such as finance, statistics, and linear algebra.

**2.Pandas:**

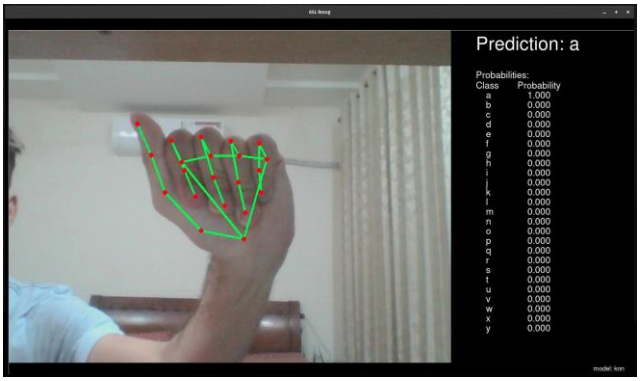
Pandas is an open-source Python library for data manipulation and analysis, offering high-performance tools for data loading, preparation, manipulation, modelling, and analysis. It is widely used in various fields, academia and industry, for data-driven tasks such as finance, economics, statistics, and analytics. Pandas provides powerful data structures like Data Frames, enabling efficient data handling, cleaning, and merging.

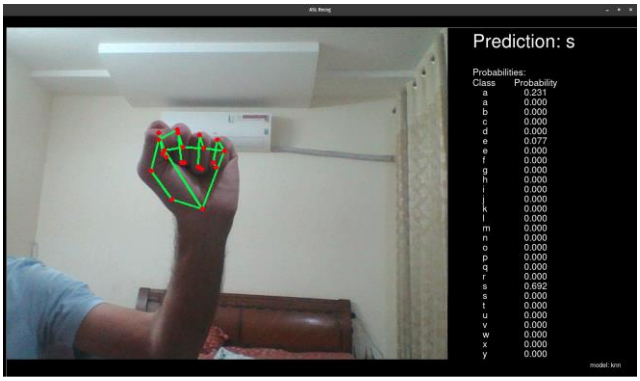
**3.Skikit-Learn:**

Scikit-learn is a versatile Python library for machine learning, offering a comprehensive suite of supervised and unsupervised learning algorithms with a consistent interface. Licensed under a permissive BSD license, it is widely used in academic and commercial settings. Scikit-learn provides tools for classification, regression, clustering, and model selection, making it a popular choice for data analysis and machine learning tasks due to its seamless integration with other Python libraries.

**4.scikit-image (skimage)** :It is a Python library designed for image processing and computer vision tasks. It offers a comprehensive collection of algorithms for tasks such as image enhancement, restoration, segmentation, and feature extraction. Built on top of NumPy, SciPy, and Matplotlib, scikit-image provides a user-friendly interface for tasks like image I/O, geometric transformations, filtering, morphological operations, color image processing, and feature detection. It is widely used in scientific research, medical imaging, remote sensing, and industrial inspection due to its versatility and ease of use.

## IV. RESULT



## VI. CONCLUSION

### 1.Sign Language Communication

Our project can be used to help in communication with sign language and remove the barrier between people with speech disabilities and others.

### 2.Gesture Control

It can be used to control various other products and software applications using such gestures.

An example will be controlling your music, skipping tracks, changing volume or even playing games using gesture control.

## V.REFERENCES

1.Census India,Disabled Population. https://censusindia.gov.in/census_and_you/disabled_population.aspx

2.OpenCV Documentation, https://docs.opencv.org/master/

3.Media Pipe Hands, https://google.github.io/mediapipe/solutions/hands.html

4.Model Card – Media Pipe Hands, https://drive.google.com/file/d/1yiPfkhb4hSbXJZaSq9vDmhz24XVZmxpL/preview

5.Altman, Naomi S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression

6.Cortes, Corinna; Vapnik, Vladimir N. (1995). Support-vector networks

7.Shalev-Shwartz, Shai; Ben-David, Shai (2014). Decision Trees

8. Sawant, S. (2014). Sign Language Recognition System to aid Deaf-dumb People Using PCA.

9.Patil, Purushottam & Patil, Neel & Tewar, Chandru & Bhoi, Rohit & Yadav, Deepak. (2019). Deaf Communicator.

10.Machine Learning Pipeline. https://algorithmia.com/blog/ml-pipeline