International Journal of Education and Science Research Review
Volume-11, Issue-2 March-April-2024 E-ISSN 2348-6457
(Peer-Reviewed, Open Access, Fully Refereed International Journal) P-ISSN 2349-1817
www.ijesrr.org Impact Factor - 7.02 Email- editor@ijesrr.org

Sign Language Recognition Using Machine Learning

[1]Ms. Kohale Sakshi Rajendra ,[2]Ms. Ingole Dipali Anandrao,

[3]Ms. Gavali Arati Mahaling, [4]Mr. Dhole Rohan Arun, [5]Prof. Miss. Patil D.S.

Dept. of Electronics and Telecommunication Engineering, FTC college of engineering & Research, Sangola.

E-mail: sakshirkohale2002@gmail.com, dipaliingole2002@gmail.com, aartigavali2121@gmail.com, dholerohan74@gmail.com

Abstract: - One of the ways to communicate with deaf and dumb individuals is through sign language. So, in order to speak with deaf and dumb people, one should learn sign language; yet, because not everyone can learn it, communication becomes nearly impossible. The goal of this study is to use machine learning to break through these communication hurdles. The majority of existing technologies and sensors, which are out of reach for most people. We utilize Open CV to take images and the CNN technique to train the machine, with the output being text. Many previous studies have offered methods for partial sign language identification, however this study intended for the full acceptance of Indian Sign Language comprises of 26 letters. As a result, the goal of this research was to extract features from finger and hand motions in order to distinguish between static and dynamic gestures.

Hand Gesture recognition system provides us an innovative, natural, user friendly way of communication with the computer which is more familiar to the human beings. By considering in mind the similarities of human hand shape with four fingers and one thumb, the software aims to present a real time system for recognition of hand gesture on basis of detection of some shape-based features.

Keywords-Sign Language, CNN Technique, Human hand gestures, Gesture Recognition, AI machine learning,

I.INTRODUCTION: -

Deaf (hard hearing) and dumb people use Sign Language (SL) as their primary means to express their ideas and thoughts with their own community and with other people with hand gestures. It has its own vocabulary, meaning, and syntax which is different from the spoken language or written language. Humans have a variety of methods for communicating with each other. This includes actions like bodily gestures, face expressions, spoken words, etc. However, people who are hard on hearing are limited to communicate with hand motions. People with hearing disabilities and/or speech disabilities use a standard sign language which cannot be understood by people who do not know it. Also, learning sign language is hindered by their disability.

Spoken language is a language produced by articulate sounds mapped against specific words and grammatical combinations to convey meaningful messages. Sign language uses visual hand and body gestures to convey meaningful messages. There are somewhere between 138 and 300 different types of Sign Language used around globally today. In India, there are only about 250 certified sign language interpreters for a deaf population of around 7 million.

Sign Language Recognition is an attempt to recognize these hand gestures and convert them into corresponding text.

II. LITERATURE SURVEY

This works focuses on static fingerspelling in Indian Sign Language A method for implementing a sign language to text/voice conversion system without using handheld gloves and sensors, by capturing the gesture continuously and converting them to voice[1]. in this method, only a few images were captured for recognition. The design of a communication aid for the physically challenged [3].

Design of a communication aid for physically challenged:

The system was developed under the Visual Studio environment. It consists of mainly two phases via training phase and the testing phase. In the training phase, we used feed-forward neural networks [2].

1.1 Dataset

We have used multiple datasets and trained multiple models to achieve good accuracy.

1.2 Dataset: We have used multiple datasets and trained multiple models to achieve good accuracy [1].

1.3 ISL Alphabet The data is a collection of images of the alphabet from the Indian Sign Language, separated into 29 folders that represent the various classes. The training dataset consists of 87000 images which are 200x200 pixels. There are 29 classes of which are English alphabets A-Z [2]

1.4 Sign Language Gesture Images Dataset The dataset consists of 37 different hand sign gestures which include A-Z alphabet gestures. Each gesture has 10 to 12 images which are 50x50 pixels, so altogether there are 37 gestures which means there 55,500 images for all gestures. Convolutional Neural Network (CNN) is well suited for this dataset for model training purposes and gesture prediction. 3.2 Data Pre-processing An image is nothing more than a 2-dimensional array of numbers or pixels which are ranging from 0 to 255.Typically, 0 means black, and 255 means white. Image is defined by mathematical function f(x,y) where 'x' represents horizontal and 'y' represents vertical in a coordinate plane. The value of f(x, y) at any point is giving the pixel value at that point of an image[3].

III. BLOCK DIAGRAM:

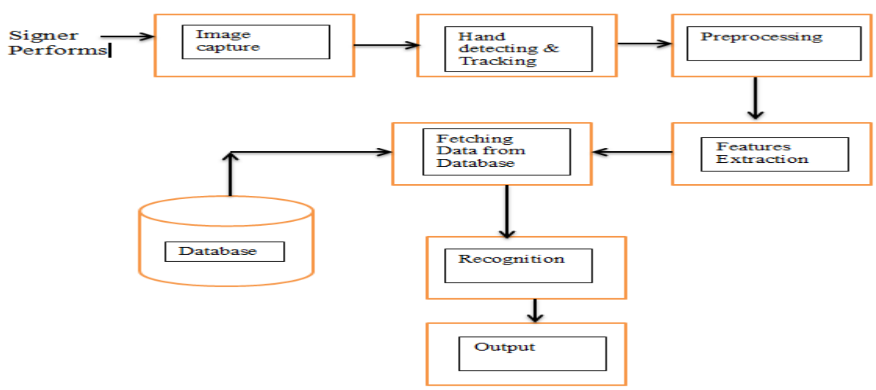

Figure 1. Block Diagram of Sign language Recognition using ML

3.1.1 Working of Block Diagram:

3.1 Image Acquisition: It is the action of extracting an image from a source, typically a hardware-based source, for process of image processing. Web Camera is the hardware-based source in our project. It is the first step in the workflow sequence because no processing can be done without an image. The picture that is obtained has not been processed in any way.

3.2 Segmentation: The method of separating objects or signs from the context of a captured image is known as segmentation. Configure 2. Proposed Methodology. Text subtracting, skin-color detection, and edge detection are all used in the segmentation process. The motion and location of the hand must be detected and segmented in order to recognise gestures.

3.3 Features Extraction: Predefined features such as form, contour, geometrical feature (position, angle, distance, etc.), colour feature, histogram, and others are extracted from the pre-processed images and used later for sign classification or recognition. Feature extraction is a step in the dimensionality reduction process that divides and organises a large collection of raw data. Reduced to smaller, easier-to-manage classes As a result, processing would be simpler. The fact that these massive data sets have a large number of variables is the most important feature. To process these variables, a large amount of computational power is needed.

3.4 Pre-processing: Each picture frame is pre-processed to eliminate noise using a variety of filters including erosion, dilation, and Gaussian smoothing, among others. The size of an image is reduced when a colour image is transformed to grayscale. A common method for reducing the amount of data to be processed is to convert an image to grey scale.The phases of pre-processing are as follows:

3.5 Recognition: We'll use classifiers in this case. Classifiers are the methods or algorithms that are used to interpret the signals. Popular classifiers that identify or understand sign language include the Hidden Markov Model (HMM), Nearest Neighbour classifiers, Support Vector Machine (SVM), Artificial Neural Network (ANN), and Principle Component Analysis (PCA), among others. However, in this project, the classifier will be CNN. Because of its high precision, CNNs are used for image classification and recognition.

3.6 Text output: Understanding human behaviour and identifying various postures and body movements, as well as translating them into text.

IV PROPOSED ALGORITHM

Creating the sign language recognition dataset: Any frame that detects a hand within the ROI (region of interest) generated can be transferred to a directory that contains a pair of directories, train and take a look at, every containing 10 folders containing images captured mistreatment the produce gesture knowledge.py perform. Now, to create the dataset, we have a tendency to use Open CV to urge the live cam feed and create a ROI that is solely the portion of the frame wherever we have a tendency to want to find the hand for the gestures. For differentiating between the backgrounds we have a tendency to calculate the accumulated weighted average for the background then cipher this from the frames that contain some object ahead of the background which will be distinguished as foreground. This can be accomplished by computing the accumulated weight for specific frames and the context's accumulated average. After we've the accumulated average for the background, we tend to subtract it from each frame that we read after sixty frames to seek out any object that covers the background.

3.5 Recognition: We'll use classifiers in this case. Classifiers are the methods or algorithms that are used to interpret the signals. Popular classifiers that identify or understand sign language include the Hidden Markov Model (HMM), Nearest Neighbour classifiers, Support Vector Machine (SVM), Artificial Neural Network (ANN), and Principle Component Analysis (PCA), among others. However, in this project, the classifier will be CNN. Because of its high precision, CNNs are used for image classification and recognition. The CNN uses a hierarchical model that builds a network, similar to a funnel, and then outputs a fully-connected layer in which all neurons are connected to each other and the output is processed.

3.6 Text output: Understanding human behaviour and identifying various postures and body movements, as well as translating them into text.

SDK Requirements:

1. **Keras:**

Keras a high-level, deep learning API developed by Google for implementing neural networks [2]. It is written in Python and is used to make the implementation of neural networks easy. It also supports multiple backend neural network computation.

Keras is relatively easy to learn and work with because it provides a python frontend with a high level of abstraction while having the option of multiple back-ends for computation purposes. This makes Keras slower than other deep learning frameworks, but extremely beginner-friendly.

Keras allows you to switch between different back ends. The frameworks supported by Keras are

Tensor flow, Theano, PlaidML, MXNet, CNTK (Microsoft Cognitive Toolkit)

Out of these five frameworks, Tensor Flow has adopted Keras as its official high-level API. Keras is embedded in Tensor Flow and can be used to perform deep learning fast as it provides inbuilt modules for all neural network computations. At the same time, computation involving tensors, computation graphs, sessions, etc can be custom made using the Tensor flow Core API, which gives you total flexibility and control over your application and lets you implement your ideas in a relatively short time.

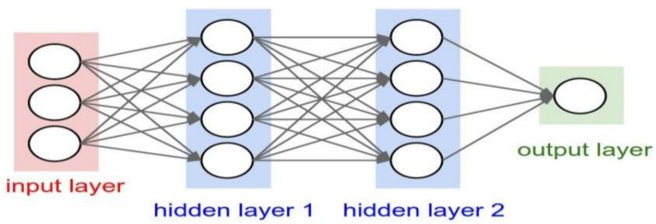

Figure 2. Process of Keras SDK

Open CV:

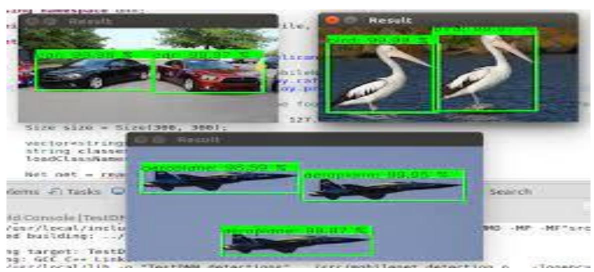

Figure 3. Eg. Of Open CV

Open CV is the huge open-source library for the computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today's systems. By using it, one can process images and videos to identify objects, faces, or even handwriting of a human [2]. When it integrated with various libraries, such as NumPy, python is capable of processing the Open CV array structure for analysis. To Identify image pattern and its various features we use vector space and perform mathematical operations on these features [3].

2. **Tensor flow:**

Tensor Flow is an open-source library developed by Google primarily for deep learning applications. It also supports traditional machine learning. Tensor Flow was originally developed for large numerical computations without keeping deep learning in mind. However, it proved to be very useful for deep learning development as well, and therefore Google open-sourced it.

Tensor Flow accepts data in the form of multi-dimensional arrays of higher dimensions called tensors. Multi-dimensional arrays are very handy in handling large amounts of data.

The next part of the What is Tensor Flow tutorial focuses on why should we use Tensor Flow.

3. **CNN:**

Yann LeCun, director of Facebook's AI Research Group [3], is the pioneer of convolutional neural networks. He built the first convolutional neural network called LeNet in 1988. LeNet was used for character recognition tasks like reading zip codes and digits.

Have you ever wondered how facial recognition works on social media, or how object detection helps in building self-driving cars, or how disease detection is done using visual imagery in healthcare? It's all possible thanks to convolutional neural networks (CNN). Here's an example of convolutional neural networks that illustrates how they work [2]:

Imagine there's an image of a bird, and you want to identify whether it's really a bird or some other object. The first thing you do is feed the pixels of the image in the form of arrays to the input layer of the neural network (multi-layer networks used to classify things). The hidden layers carry out feature extraction by performing different calculations and manipulations. There are multiple hidden layers like the convolution layer, the ReLU layer, and pooling layer, that perform feature extraction from the image. Finally, there's a fully connected layer that identifies the object in the image.

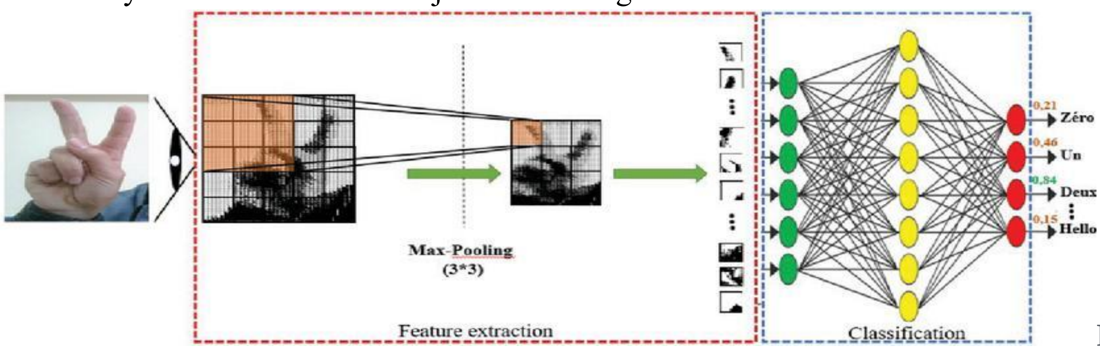

Figure 4. Process of Hand Gesture in CNN

V. RESULTS:

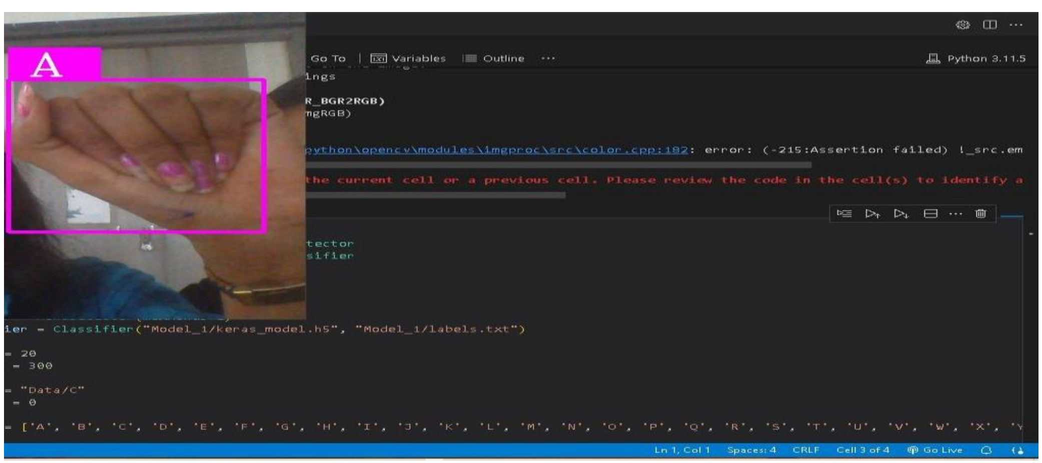

Figure. Output of SLR Symbol 'A'

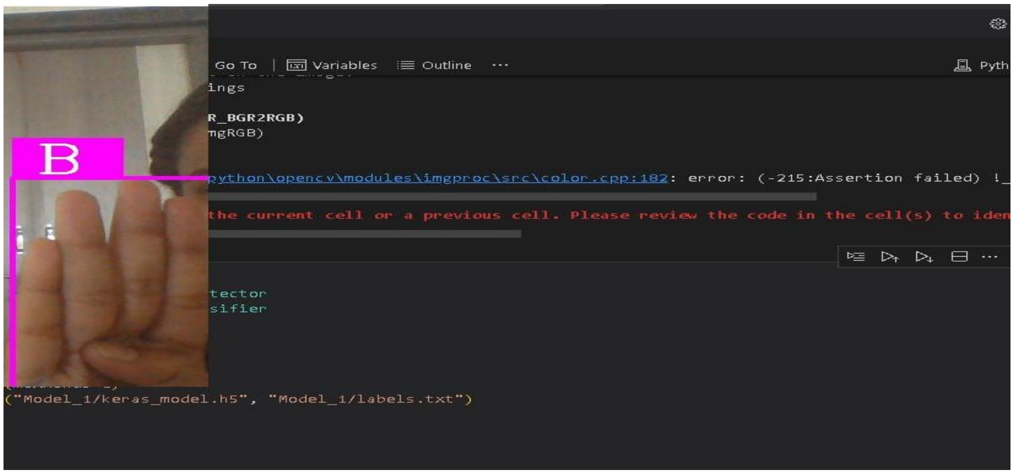

Figure. Output of SLR Symbol ‘B’

VI. CONCLUSION:

In conclusion, we were successfully able 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 like this system can detect A-Z alphabets hand gestures but doesn’t cover body gestures and other dynamic gestures. We are sure and it can be improved and optimized in the future

VIII. FUTURE PROSPECTS:

The proposed sign language recognition system used to recognize sign language letters can be further extended to recognize gestures facial expressions. Instead of displaying letter labels it will be more appropriate to display sentences as more appropriate translation of language. This also increases readability. The scope of different sign languages can be increased. More training data can be added to detect the letter with more accuracy. This project can further be extended to convert the signs to speech. The sign language start in android Mobile.

In future we try to convert text into Audio or Sound which can easily help to Dumb & Deaf persons for communicating to other peoples.

IX. REFERENCES:

[1] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 4510-4520, doi: 10.1109/CVPR.2018.00474. [2] L. K. Hansen and P. Salamon, "Neural network ensembles," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 10, pp. 993-1001, Oct. 1990, doi: 10.1109/34.58871.

[2]Kang, Byeongkeun, Subarna Tripathi, and Truong Q. Nguyen. ”Real- time sign language fingerspelling recognition using convolutional neural networks from depth map.” arXiv preprint arXiv: 1509.03001 (2015).

[3] Suganya, R., and T. Meeradevi. ”Design of a communication aid for phys- ically challenged.” In Electronics and Communication Systems (ICECS), 2015 2nd International Conference on, pp. 818-822. IEEE, 2015.