Real Time Sign Language Detection Using OpenCV

Pavan Kumar Meka and Yesu Raju Parusu and Vaddi Radhe Syam

Department of Information Technology, Velagapudi Ramakrishna Siddhartha Engineering College, Andhra Pradesh, India

Pavanmeka74@gmail.com
yesurajuparusu@gmail.com
syam.radhe@gmail.com

Abstract. Sign language is communication commonly used by people who are deaf and dumb. They use hand gestures to exchange information between their community and with the rest of the world. Real time sign language detection is the solution that provides the communication between normal people and those who are deaf and dumb. It deals with image acquisition and continues until text generation. Previously provided solutions are all sensor based and they don't give good accuracy. The old models customized with limited gestures. In order to overcome this problem, a model has been developed to detect the 40 different classes of signs. Which includes alphabets, numbers and some special symbols. This model can effectively recognize gestures dynamically. The dataset is created in well planned manner to include 10 numbers, 24 alphabets and 6 special symbols. The sign language recognition steps are in this survey are Data collection, image preprocessing and segmentation, training model and last classification and recognition of result. The captured images are preprocessed to remove the extra background in image. A convolutional neural network is used to train the model and perform multi-class classification on image data. The results are displayed in the result window with text recognition. Some future improvements for research in this area also suggested.

Keywords: sign language, classification, convolutional neural networks, hand gestures, communication, real-time

1.Introduction

Communication is very important for sharing information and to express oneself. Deaf and dumb people use sign language to express their ideas and thoughts with their community and with the rest of the world. Sign language is actually a hand gesture that convey the meaning message. There are around 400 different types of sign languages are used globally. And in India there are around 250 certified sign language interpreters for deaf people. This is very difficult to teach the sign language for all the deaf and dumb people as there are limited number of interpreters available today.

To overcome this problem, a model is to be developed to detect the signs automatically. Previously developed models are all sensor based and they

need gloves and high-resolution cameras to capture and identify the hand gesture. They are all limited to certain hand gestures only. Some models are able to detect only statics images, when it comes to dynamic images they do not work properly. These are all drawbacks of the previously developed models.

For images convolutional neural networks is the most popular neural networks than the other one. The convolutional neural network (CNN) that we have used here 3-layer network. The network layers are convolutional layer, pooling layer and fully connected layer. Datasets that we have created includes 40 different classes which are 24 alphabets, 9 numbers and 6 special symbols. With the help of deep learning model, we can easily identify the gestures. So, this system will help deaf and dumb people to communicate easily with rest of the world.

2. Literature Study

Here is the study of various previously available models, methodologies and techniques to sign language which were referenced to create our proposed model.

Hand Gesture Recognition Using PCA (2020) [1]:

Authors presented the static hand gesture recognition system using digital image processing. For hand gesture feature vector SIFT algorithm is used. Proposed a system to aid communication of deaf and dumb people communication using Indian sign language

An Automated System for Indian Sign Language Recognition (2018) [2]:

Developed a method for automatic recognition of signs on the basis of shape-based features is presented. Features of segmented hand region are calculated using Hu's invariant moments that are fed to Artificial Neural Network for classification.

Design of ANFIS system for recognition of single hand and two hand signs (2019) [3]:

Author presented application that helps the deaf and dumb person to communicate with the rest of the world using sign language. The key feature in this system is the real time gesture to text conversion

Finding relevant image content for mobile sign language recognition: Signal Processing, Pattern Recognition and Application [4]:

This work deals with deep learning algorithms like CNN, RNN which work better on image data. But the accuracy for this model is not up to the mark. It works well for the static images.

Vision based hand gesture recognition. (2021) [5]:

He proposed a real time vision-based system for hand gesture recognition for human computer interaction in many applications. The system can recognize 35 different hand gestures given by Indian and American Sign Language or ISL and ASL at faster rate with virtuous accuracy.

Hand Gesture Recognition using Image processing and feature extraction techniques [6]:

This paper is mostly focuses on use of traditional preprocessing techniques to effectively detect the correct hand gesture. And this model is works well for static images. Whereas in the case of dynamic images model do not work properly.

3. Proposed Work

The proposed model uses the OpenCV to read and capture the hand gestures using webcam. After that we will pre-process image to remove the extra background from the image so that model will focus on main part of the gesture. A model is trained using Convolutional Neural Network (CNN) to classify the different classes of hand gestures. Convolutional Neural Networks works well for the images rather than any other neural networks. It classifies the hand gestures with more accurate and effectively whenever gestures are shown to webcam.

3.1 Data set

The proposed model was trained on 8000 colour images of different classes of hand gestures. Which includes 26 alphabets, 9 numbers and 6 specials symbols or English Phrases. For this model we have created our own data set by capturing different classes of hand gestures in different angles and light adjustments.

Gesture type	No. of data samples
Alphabets (A-Z)	Each gesture contains 200 images 26*200 =5200
Numbers (1-9)	Each gesture contains 200 images 9*200 =1800
Okay	200
Call me	200
Rock	200
Smile	200
Thumbs down	200
Thumbs up	200

Fig. 1. Dataset Structure

3.2 Pre-processing

The captured images were pre-processed to remove the additional background in the hand gestures by using image filtering and Segmentation. So, this will modify or enhance image properties and extract the valuable information from the pictures such as corners, edges and etc… A filter is nothing but kernel or some particular pattern, which is applied to each pixel and its neighbours within an image.

A total of 200 images were pro-processed to enhance the image properties.

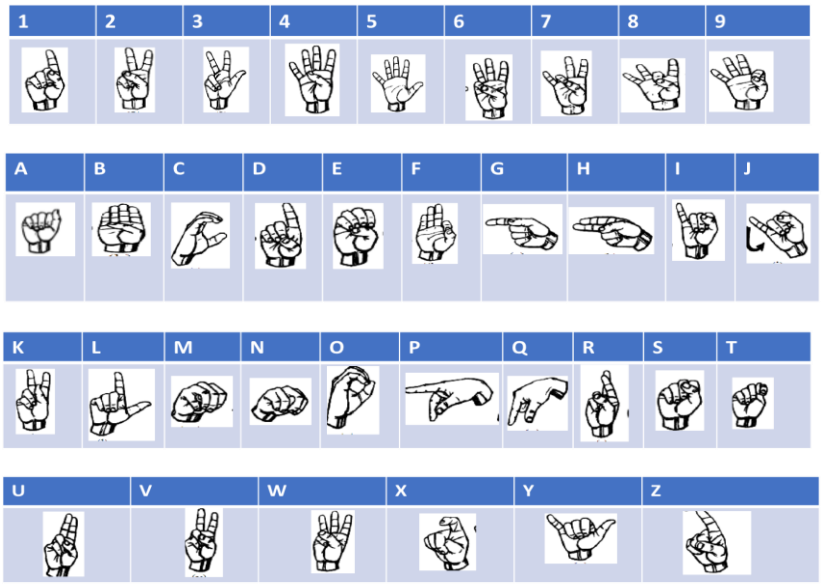

Fig. 2. Created dataset

3.3 Design methodology

The design methodology of the proposed model has been described below:

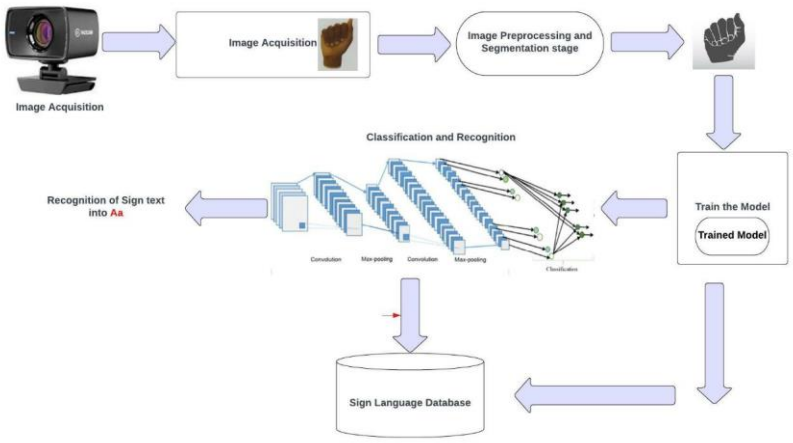

Fig. 3. Architecture Diagram

The first step is we have captured hand gestures of different classes using different angles. After that each image is pre-processed to enhance the image properties. And all the images are stored in database with labelling. Then, these images are used to train the model. A CNN model was created by including three main layers named convolutional layer, pooling layer and fully connected layers.

This deep learning model was completely trained and tested using our own dataset. Once the we got the required accuracy, the model has been used to make predictions of the hand gestures.

Algorithm:

Step 1: Create a separate file for dataset in project folder.
Step 2: Create required dataset by capturing images of hand gestures using webcam.
Step 3: Add labels to each image and store them in a database (separate file)
Step 4: Perform image pre-processing to each image that is captured.
Step 5: Build a CNN (convolutional Neural Networks) model using our own dataset.
Step 6: Train and test the model using customised dataset.
Output image using convolutional layer =**N-F+1**
Output image using pooling layer = **N+2P-F+1, S=2, P=2**
Where n= size of input, F= size of kernel/filter, P=padding, S=stride
Step 7: Use this model to detect the sign language
Step 8: Use the webcam to test the trained sign language recognition model

The trained model gives the accuracy of overall 90.56%.

4. Results and Observations

Fig. 4. Outputs produced by model

The model was executed for 20 spans until required accuracy was obtained. The Interpretation Standards like precision, recall and F1-score of the model are displayed below.

Interpretation Standards	Proposed model
Precision	0.89
F1-score	0.85
Recall	0.91
Support	91.62
Testing Accuracy	90.56
Weighted average	89.43

Fig. 5. Interpretation standards of the model

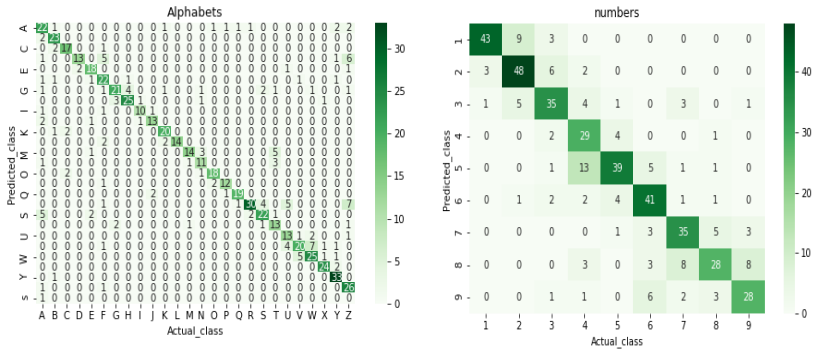

Fig. 6. Confusion Matrix

The model performance is low at starting up accurateness. But the model executes exceptionally well as the training progresses. The validation set loss is seen to fluctuate heavily during the entire training period.

The testing accuracy of 90.56%, precision of 0.89, recall of 0.91 and an f1 score of 0.85.

5. Advantages

i. It helps the children who have autism spectrum disorder (ASD).
ii. It gives the literacy for Dump and deaf people.
iii. Elderly people benefited greatly from the use of sign language detection model.
iv. Sign language model gives efficient and accurate way to covert sign language into text in real time.

6. Conclusion

The model is to plays an important role in the society to help the people who are suffering with autism spectrum disorder (ASD). The aim of the model to recognize various sign language signs and transform them into text. It facilitates communication between the people who are suffering with dump and deaf impairments. It gives high accuracy and response time when applied in practical scenarios. In future we can develop apps on these models such it will easily availably phones to help the people. It can also be deployed as an API using cloud such that it can be integrated with other applications easily. The Model provides communication and can play a key role in disabled community.

7. Improvements

1. we have developed a model that runs on a PyCharm IDE. So, they need a laptop whenever they want to detect the sign.

2. Instead of using laptops we can make this project as mobile application so that people can easily understand sign language using mobile.
3. Mobile camera will be used as input for the model and model will recognition and put the result in the mobile screen

8. References

1. Survey on sign language recognition in the context of vision- based and deep learning by S. Subburaj, S. Murugavalli (2022).
2. Sign Language Recognition Application Systems for Deaf-Mute People: A Review Based on Input-Process-Output by Suhajito, Ricky Anderson, Fanny Wiryana, Meita Chandra Ariesta (2019).
3. Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions- Xiaoyang Tan and Bill Triggs (2020).
4. Indian Sign Language Character Recognition by Sanil Jain and K.V. Sameer Raj.
5. Shanmukha Swamy, M.P. Chethan, Mahantesh Gatwadi Indian sign language interpreter with android implementation Int J Comput Appl (2015), pp. 975-8887.
6. Sign Language Recognition Application Systems for Deaf-Mute People: A Review on Input-Process-Output