INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 8, Issue 1 - V8I1-1257)
Available online at: https://www.ijariit.com

Indian sign language interpreter

Shivam Sonawane
sonawaneshivam01@gmail.com
Vishwakarma Institute of Technology, Pune, Maharashtra

Nisha Sodha
nishaben.sodha@vit.edu
Vishwakarma Institute of Technology, Pune, Maharashtra

Shubham Mangrule
shubham.mangrule20@vit.edu
Vishwakarma Institute of Technology, Pune, Maharashtra

Srushti Shingade
srushti.shingade20@vit.edu
Vishwakarma Institute of Technology, Pune, Maharashtra

Aniket Joshi
aniket.joshi20@vit.edu
Vishwakarma Institute of Technology, Pune, Maharashtra

Nupur Shinde
nupur.shinde20@vit.edu
Vishwakarma Institute of Technology, Pune, Maharashtra

ABSTRACT

Hand gestures which are a form of non-verbal communication are a strong way to communicate with a hearing-impaired community. It helps to establish human-computer communication. The purpose of this project is to help the dumb and deaf people for ease in communication as this project will help us to understand their sign language to some extent as both hands are used for performing any gesture and this will eventually help them to communicate with others. Image classification is done using neural networks. The task is accomplished using Machine learning and Image classification. InceptionV2 model with feature vector is used to train the dataset of around 10000 images in total and 300 images of each sign which gives the optimum accuracy.

Keywords *— ASL, Tensorflow, OpenCV, Image processing, Deep learning*

1. INTRODUCTION

The Indian Sign language is widely used by deaf and dumb people as a medium for communication. And the deaf and dumb people use these for communicating with each other. Sometimes the deaf need to take the help of a sign language interpreter. to express their doubts and thoughts to normal people.

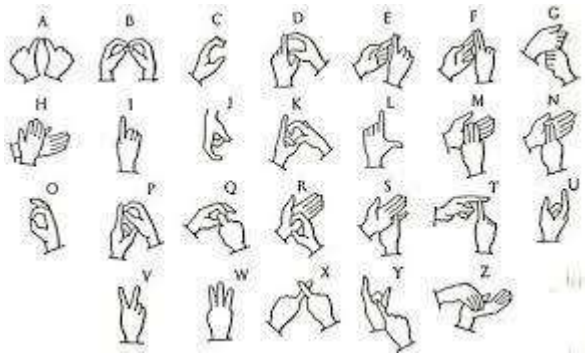


fig.1 Signs in Indian Sign Language[8]

Sign language is nothing but the recognition of the hand gestures composed by the different shapes and movements of hands. Indian Sign Language (ISL), is the main sign language used in India, by anywhere between 1 million to 2.7 million people. A sign language consists of a well-structured code of signs, and gestures, each of which has a particular meaning assigned to it. The existing methodology has 3 layer CNN network. But the model we have used is InceptionV2 with a feature vector. We have also preprocessed the input data using data augmentation. We have got good results with accuracy of more than 99%.

2. LITERATURE REVIEW

1] Kumund Tripathi, Neha Baranwal, and G.C.Nandi carried out a project on Continuous Indian Sign Language Gesture Recognition and Sentence Formation. Gestures are performed by the deaf and dumb communities for sign language all over the world but in their regional form like ASL, ISL. So for the Project, 10 sentence database is created which is static as well as dynamic in nature.

Extracting the start frame and end frame of each gesture is the problem. This problem is solved by the gradient-based Key Frame Extraction Method. In preprocessing Silhouette image of every gesture is created. Skin color segmentation and HSV (Hue, Saturation, Value) is applied for extracting skin region and eliminating non-skin region. Later Median filter is applied for preserving outer boundaries. The keyframe extraction is performed for extracting meaningful frames. Middle frames are the most informative frames so only those are considered. Then Orientation Histogram is applied as a feature extraction technique for extracting the most appropriate features. Orientation Histogram with 36 bins gives higher accuracy than 18 bins.

The principal component analysis is used for finding patterns in input data. For gesture recognition, different distance matrices are tested and used of which Euclidean distance and

Correlational have higher recognition rates than others. Outcomes are tested the usage of everyday Webcam and get appropriate results. Experimental effects shows nice accuracy and consequences. [1]

2] The following report is a literature review of the paper of the project "Indian Sign Language Recognition" carried out by students of IIT Kanpur.

This project aims at identifying alphabets in Indian Sign Language from the corresponding gestures with the use of computer vision and machine learning algorithms. Major challenges while completing the project were the lack of a standard dataset and occluded features of Indian Sign Language like using both the hands for sign language and multiple signs for the same character. The dataset was collected from a deaf school by clicking images of sign language from 8 deaf students. In the implementation, the first stage is called image segmentation, in which the skin part is segmented from the image. For implementing this, first, they used the SVM learning algorithm to segment out non-skin pixels and then used the HSV model to separate color and intensity components which makes it robust to lighting. Finally, they transformed the image from RGB space to YIQ and YUV space. The second stage is called feature extraction, which is used to extract relevant features from skin segmented. For extracting feature vectors, they used approaches like Bag of Visual Words, Histogram of Oriented Gradient with dimension reduction and without dimensionality reduction. The third step is to apply the published features as inputs to the various supervised learning models for training and finally to use the training models to classification.

The results of consecutive merging to separate one-hand image and two-handed image and use SVM with a linear line in individual clusters lead to poor accuracy. While, methods such as converting svm kernel to rbf and using random forests for training lead to relatively good accuracy.

Photos taken in poor light provide sound effects while separating images. A better set would help to obtain more accurate features and would result in higher accuracy.[2]

3] In this scientific world a lot of research is being done on automatic sign recognition techniques and it's a very effective way of transferring information to the Deaf and Hard of Hearing (HOH) peoples. So technically this can be done with the help of some sensors and machine learning. Here we propose a system for Indian Sign Language recognition, which uses Microsoft Kinect sensor and Machine learning for effectively recognizing some signs used in Indian Sign Language. The Kinect Sensor detects the 20 joints of human bones for detecting the signs. We use 11 out of 20 joints and extract 34 novel features per frame. The recognition depends upon distance, shape and angle of the 11 out of 20 joints and bones. We get such features because of the presence of a multi class support vector machine which has an accuracy of almost 86.16% to 100%. Till date the system is able to recognise 37 signs. The data used in the proposed system is generated by the Deaf and Hard of Hearing (HoH) persons in our lab.[3]

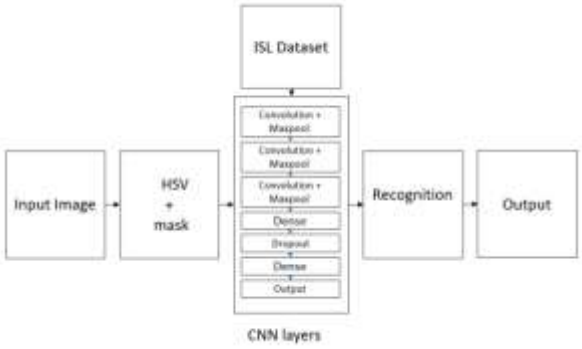
4] This research paper is about the Study on Graphic Design based on Deep Learning and Tensorflow. In this research paper, we get to learn about Deep Learning using the TensorFlow framework or also we can say image classification using a Deep Neural Network (DNN). Deep Neural Network (DNN) basically produces a high percentage of accuracy because of which it is the best option for the training process and also image classification is discussed in terms of percentage. One of the frequent and best systems that have been applied towards Image classification is Machine Learning. Basically there are four phases in this whole process and each of them is discussed thoroughly. Each phase has Python as its main programming language and uses TensorFlow as the open source software. Then, according to the process they collect some of the pictures or the inputs and by applying Deep Neural Network (DNN), all the images will be classified according to their groups. So the first process is : 1)Training Images , second is : 2) Implementing Deep Neural Network (DNN) , third is : 3) MobileNet Performance , and the last i.e. fourth process is : 4) mapping the Flowchart of the classification systems.

For now only photos of flowers were recognised but for later work or the future work this work can be extended to recognize animals and objects. [4]

3. METHODOLOGY/EXPERIMENTAL

Existing Methodology:

As shown in figure, HSV method for feature extraction is used and 3-layer CNN network is used for classification. CNN is lightweight and is suitable for Raspberry Pi as it requires lower computational power. The observed accuracy is 96%.



Proposed Methodology

The existing methodology has 3 layer CNN network. But the model we have used is InceptionV2 with feature vector. We have also preprocessed the input data using data augmentation.

4. STEPS FOLLOWED

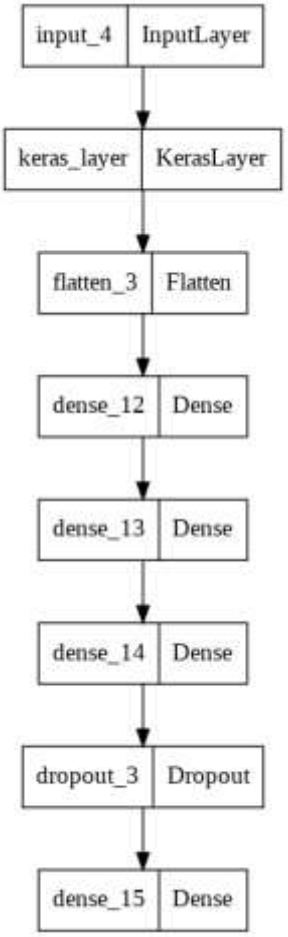
1. Imported necessary libraries like os, cv2, numpy, pandas, tensorflow, keras , sklearn etc.
2. Imported dataset into the working environment from github
3. Created subsets of training and validation with validity split of 0.25.
4. Initialized the training and testing directories.
5. Built training and testing data with data augmentation.
6. Created model checkpoint callbacks and earlystopping callback
7. Built feature vectorizer using Tensorflow Hub
8. Built model using functional API.
9. Created prediction function.
10. Used Gradio for providing easy UI.

About InceptionV2 and Feature Vector

In machine learning, feature vector is vector containing multiple elements about an object. Puting feature vectors for objects together can make up a feature space. The features may represent , as a whole , one mere pixel or an entire image. The granularity depends on what someone is trying to learn about object or represent about object. We have used InceptionV2

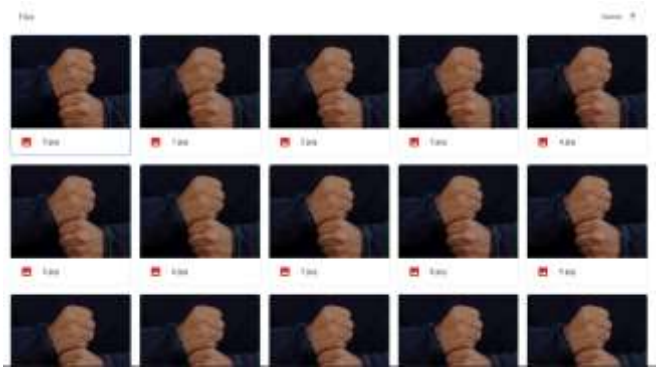
feature vector from tensorflowhub in our project which provides optimum accuracy.

Workflow of our model:
Our model contains various layers. We have used InceptionV2 feature vector transferlearning model at first. Then we have added flatten and dense layer. Also added dropout layer to avoid overfitting of model. At at last once again a dense layer as output layer.



workflow of model

Dataset
We have used dataset of around 10000 images. Dataset of each sign contains 300 images in training and 100 images in testing folder which is split of 0.25.



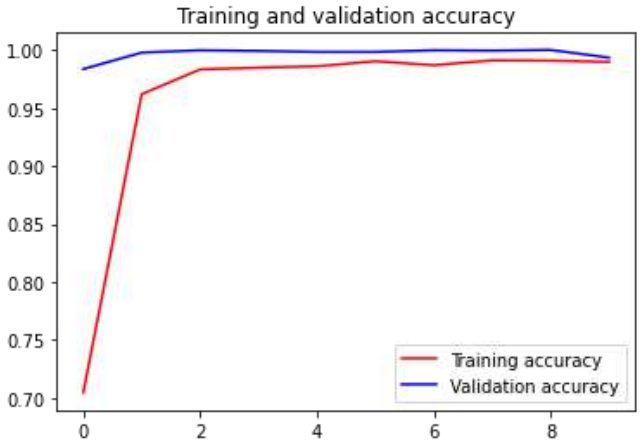
5. RESULTS AND DISCUSSIONS

This "Indian Sign Language Interpreter" is developed using Convolution Neural Network. It has achieved an accuracy of 99.68% and achieved validation accuracy of 99.8%. It successfully recognizes the alphabets and numbers of Indian Sign Language. 10 epochs are used to train the data set. Data Augmentation is done to increase the diversity of existing data. Loss after training is 0.03 and validation loss is 0.0079 which is really low and shows good results.

Results for prediction

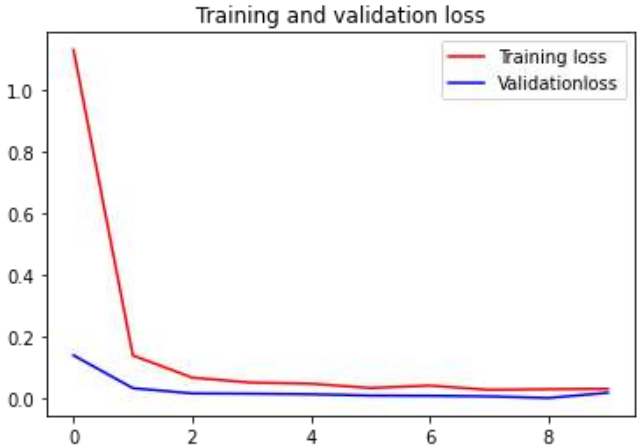
Graph for training and validation accuracy:

Following graph depicts training accuracy and validation accuracy during the learning process. We can see after each epoch training and validation both accuracies are increasing. and settling at almost 99% accuracy.



Graph for training and validation loss:

Following graph depicts training loss and validation loss during the learning process. Both losses are really low which shows the efficiency of model. Both the curves have a small gap between them which depicts that the model is approximately a good fit model.



6. LIMITATIONS

- The model gives predictions for the static images only so it is not possible for this model to be used in continuous type of sign language.

- Currently model is not capable to predict real time gestures using webcam.
- Text to voice feature is also not there in model.

7. FUTURE SCOPE

Various hand gestures are detected and used as inputs. Hand gestures that represent numbers can be converted into commands to perform related tasks in real time. The Facial Gesture Recognition Method can be used on vehicles to alert drivers who are about to fall asleep. We can include full body motion recognition, or something small like a change in facial expression. Using hand gestures many users can handle applications from distance without even touching it

8. CONCLUSION

With the use of sign language, this system will provide voice to the voiceless. This method will help impaired people communicate more easily. This project demonstrates the most efficient method for transliterating 26 static ASL alphabet . The project's major goal was to create a system that could translate static sign language into its word counterpart, which would contain letters.

9. ACKNOWLEDGMENT

We would like to express my gratitude towards our project guide Prof. Nisha R Sodha ma'am for there kind cooperation and encouragement which helped us in completing this project. We would also like to express our special gratitude to Vishwakarma Institute of Technology for allowing us to do this project.

10. REFERENCES

[1] cyberleninka.org/article/n/1321510
[2] www.irjet.net/archives/V7/i3/IRJET-V7I3418.pdf
[3] cse.iitk.ac.in/users/cs365/2015/_submissions/vinsam/report.pdf
[4] https://www.ripublication.com/ijaer18/ijaerv13n9_90.pdf
[5] https://link.springer.com/content/pdf/10.1007%2F978-3-642-02707-9_3.pdf
[6] https://www.researchgate.net/profile/Shivashankara-Ss/publication/326972551_American_Sign_Language_Recognition_System_An_Optimal_Approach/links/5b6e952492851ca65055261b/American-Sign-Language-Recognition-System-An-Optimal-Approach.pdf?origin=publication_detail
[7] https://www.kaggle.com/vaishnaviasonawane/indian-sign-language-recognition-using-inceptionv3/data
[8] https://i.pinimg.com/564x/e9/d1/c7/e9d1c752dd7f49ed4d2c54b7cfa81020.jpg

