Comparative Study On Sign Language

C. Sriinitha reddy. Stanley College of engineering and technology for women, Hyderabad, India

M. Sai anmol. Stanley College of engineering and technology for women, Hyderabad, India

E. Sai Vaishnavi. Stanley College of engineering and technology for women, Hyderabad, India

Abstract World Health Organization's (WHO) survey states that above 6% of the world's population is suffering from hearing impairments. In March 2018, the number of people with this disability is 466 million, and it is expected to be 900 million by 2050. • Text messaging, writing, using visual media, and fingerspelling are a few methods used to establish communication between normal and hearing and speech impaired people. However, they prefer sign language only because, they can express their emotions and feelings through signs only. • Text messaging, writing, using visual media, and fingerspelling are a few methods used to establish communication between normal and hearing and speech impaired people. However, they prefer sign language only because they can express their emotions and feelings through signs only.

1. INTRODUCTION

1.1 Introduction

What is Sign Language?

• Sign language typically refers to any mode of communication via movements of the body, especially associated hands and arms, used when verbal communication is impossible or not feasible. This practice is undoubtedly older than speech.

• Why do we need Sign Language?

• Sign language is the medium that introduces us to the world of people, who have an inability in hearing or verbal ability. A collection of hand gestures accompanied by facial expressions, constitute sign language. It enables the deaf and the dumb to interact with the people around them and vice versa. It helps them to understand the world around them through visual descriptions and thereby, contribute to society.

• There is no Sign Language that is considered universal. Sign language has its different types depending upon regions and their culture like any other spoken language. So, we can't say that there is a specific language that is considered a universal sign language in the world.

1.2 Problem Statement

The aim of this work is to create a system or model that can recognize alphabets in Indian Sign Language (ISL) and check the evaluationaccuracy.

• In this, we train the data to two different image processing algorithms namely: CNN and SVM.

• It also aims to understand which algorithm works best for the classification of images for the given dataset. Objective

.

1.3 OBJECTIVE

• The aim of this work is to create a system or model that can recognize alphabets in Indian Sign Language (ISL) and check the evaluation accuracy. • In this, we train the data to two different image processing algorithms namely: CNN and SVM.

• It also aims to understand which algorithm works best for the classification of images for the given dataset.

2. Literature Survey

Sign Language Recognition System For Deaf And Dumb People.

In this paper Sakshi Goyal, Ishita Sharma and Shanu Sharma have used SIFT Algorithm is for comparison of captured images with the images stored for each symbol in the directory, intended for detection of key points, orientation for obtaining feasible and expected results. The SIFT algorithm is typically a fourstep process, namely Image Acquisition, Feature Extraction, Orientation Detection, and Gesture Recognition. The process starts with Image Acquisition as the name suggests is of acquiring the image during runtime through an integrated webcam. Then the feature extraction stage is intended to provide a "feature" description of the object. In the orientation detection step, we will take the input of hand movement in any form or any orientation of the captured image, the gesture will be detected through the described section of feature extraction. In the final step Gesture Recognition stage, when the process is complete, the finalized gesture will be converted into its recognized character or alphabet. Cons: The feature extraction stage consists of a multiple-level filtering mechanism, thus requiring only those images which satisfy a certain set of constraints.

Automatic Indian Sign Language Recognition System.

This paper presents a methodology that recognizes the Indian Sign Language (ISL) and translates it into a normal text. the proposed system consists of three phases a training phase, a testing phase, and a recognition phase. In the training phase, each class is trained with a multi-class support vector machine (MSVM). Hu invariant moment and structural shape descriptors are combined to make a combinational feature vector that is to be extracted from the input image in the testing phase after applying to pre-process. In the recognition phase, different classes are used for testing an input gesture. The outcome with the most probable group is identified to recognize the gesture. Finally, after the recognition of the input image, their meaning is displayed on the screen. The effectiveness of the proposed method is validated on a dataset having 720 images. Experimental results demonstrate that the proposed system can successfully recognize hand gestures with a 96% recognition rate. Cons: It is quite challenging to define the constraints of the MSVM in the installed system in order to carry out this process.

Indian Sign Language Recognition Using SVM In this paper

A vision-based technique has been used in this paper for hand sign recognition, and a support vector machine (SVM) was used to classify as a sign of ISL. This process consists of the following steps: Capturing video, Skin filtering, Preprocessing, Feature calculation, SVM classification, sign recognition. The captured video was converted to HSV color space for pre-processing and then segmentation was done based on skin pixels. Also, Depth information was used in parallel to get more accurate results. Hu- Moments and motion trajectory were extracted from the image frames and the classification of gestures was done by Support Vector Machine. In this way, This system targets to facilitates disabled people who are not able to hear and there are not many people who can understand their sign language. The Indian Sign Language Recognition system was implemented and 4 gestures were classified using SVM. Results were 97.5% accurate. Cons : Long training time for large datasets. Difficult to understand and interpret the final model, variable weights, and individual impact

3. OVERVIEW OF THESYSTEM

3.1 Existing System

• In the past few years, neural abstractive text summarization with sequence-tosequence (seq2seq) models have gained a lot of popularity. Many interesting techniques have been proposed to improve seq2seq models, making them capable of handling different challenges, such as saliency, fluency and human readability, and generate high-quality summaries. Generally speaking, most of these techniques differ in one of these three categories: network structure, parameter inference, and decoding/generation. There are also other concerns, such as efficiency and parallelism for training a model. In this paper, Yaser Keneshloo provided a comprehensive literature survey on different seq2seq models for abstractive text summarization from the viewpoint of network structures, training strategies, and summary generation algorithms. Several models were first proposed for language.

3.2 Proposed System

CNN (CONVOLUTION NEURAL NETWORK) In this, we propose to train the ISL data to the CNN (Convolution Neural Network) and SVM (Support Vector Machine) and do a comparative study and conclude VGG16: VGG16 is a Convolution neural network (CNN). The most unique thing about VGG16 is that instead of having a large number of hyper-parameters they focused on having convolution layers of 3x3 filter with a stride 1 and always uses same padding and max pool layer of 2x2 filter of stride 2. This network is a pretty large network and it has about 138 million (approximately) parameters.

SVM (SUPPORT VECTOR MACHINE)

SVM is a feasible algorithm for performing image classification. It is primarily a learning algorithm that is mainly used to classify data into different sets. SVM trains on a collection of label data. The basic advantage of SVM is that it can be used for both classification and object detection problems. Stanley college of Engineering & Technology for Women, B.E. (Information Technology) 13 SVM makes use of a decision boundary which acts as a distinct parameter between any two datasets enough to classify them. SVM is also used in Object Detection and image classification problems. SVMs hold the capability of efficiently performing image classification processes using the kernel trick, which is defined as implicitly mapping their inputs into high-dimensional feature spaces.

3.3 Proposed System Design

In this project work, I used five modules and each module has own functions, such as:

1. The Dataset
2. Data Cleaning
3. Utility Functions
4. The Model
5. Training the Model
6. Inference

3.3.1 Dataset

Kaggle is the world's largest Data Science community consisting of powerful tools and resources. We train the model with a pre-processed dataset found on the Kaggle website. • Name of the Dataset: Indian Sign Language (ISL) Dataset • The Dataset consists of the number and alphabetical representations in Indian Sign Language. The images are pre-processed and fed to the model as part of the training dataset. Each Sign in the dataset has 1200 pictures taken in various environments and processed.

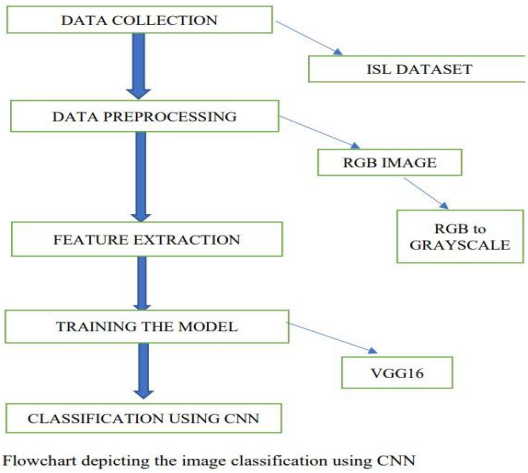
3.3.2 Preprocessing

Data Pre-processing is the removal of noise (unwanted information) from the image. The initial image which is captured is unprocessed and has a lot of unwanted data. The very first step in image pre-processing is the blurring or smoothing of the image. This helps in reducing sharp edges which might affect the functionality of the model.

• The image is generally captured as an RGB image and this can be converted to other color models. The image is then converted to gray-scale and by applying the global threshold technique on the histogram of the gray color image, we can convert the image into binary form.

• Global thresholding: The global thresholding method take advantage of the image histogram. The image histogram is a type of statistical graph with a grayscale value on the x-axis and the number of pixels for each grayscale on the y-axis.

3 Architecture



Flowchart depicting the image classification using CNN

Fig 1: Architecture

Algorithm:

CNN (CONVOLUTION NEURAL NETWORK)

In this, we propose to train the ISL data to the CNN (Convolution Neural Network) and SVM (Support Vector Machine) and do a comparative study and conclude VGG16: VGG16 is a Convolution neural network (CNN). The most unique thing about VGG16 is that instead of having a large number of hyper-parameters they focused on having convolution layers of 3x3 filter with a stride 1 and always uses same padding and max pool layer of 2x2 filter of stride 2. This network is a pretty large network and it has about 138 million (approximately) parameters

SVM (SUPPORT VECTOR MACHINE)

SVM is a feasible algorithm for performing image classification. It is primarily a learning algorithm that is mainly used to classify data into different sets. SVM trains on a collection of label data. The basic advantage of SVM is that it can be used for both classification and object detection problems. SVM makes use of a decision boundary which acts as a distinct parameter between any two datasets enough to classify them. SVM is also used in Object Detection and image classification problems. SVMs hold the capability of efficiently performing image classification processes using the kernel trick, which is defined as implicitly mapping their inputs into high-dimensional feature spaces.

4 RESULTS

Accuracy and loss using CNN

The ISL dataset has been fed to VGG16 (one of the standard CNN models used for image classification). The accuracy achieved using this model is approximately 98%

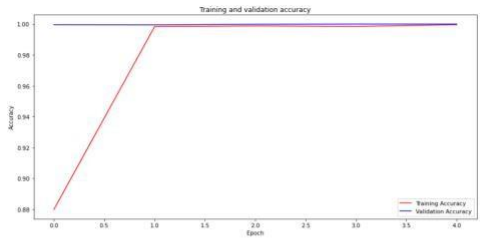


Fig 5.1 Training and validation accuracy (CNN)

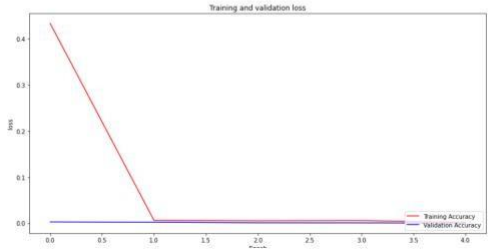


Fig 5.2 Error rate (CNN)

.

Accuracy and loss using SVM The ISL dataset has been fed To SVM model. The accuracy achieved is approximately 99.1%. As we analyze the graph, it is concluded that as the number of epochs is increasing, the accuracy of the system is improving as well.

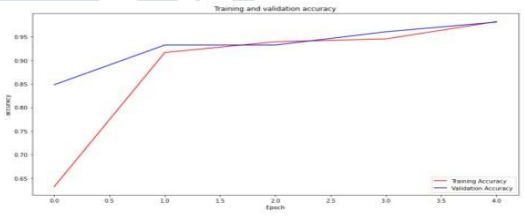
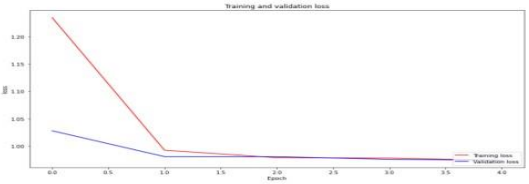


Fig 5.3 Training and validation accuracy (SVM)



Comparison: CNN vs SVM:

5.1 Epochs = '5'

The training accuracy of CNN and SVM when epochs ="5" is observed to be 98.2% and 96.3% respectively.

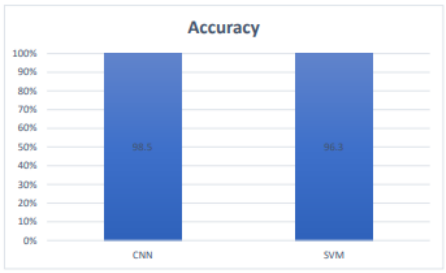


Fig 5.5 CNN vs SVM chart when epochs= 5

5.2 Epochs = "10"

The training accuracy of CNN and SVM when epochs ="10" is observed to be 98.4% and 96.3% respectively.

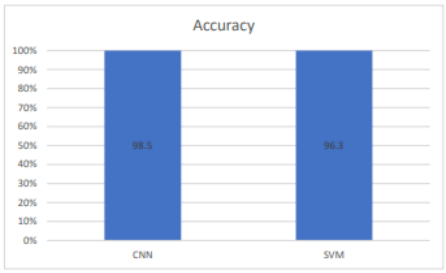


Fig 5.5 CNN vs SVM chart when epochs= 10

6. CONCLUSION

This system helps to build the communication gap between the specially-abled and normal people and since this project is based on ISL, which is one of the lesser-known sign languages, this work can act as one of the references for future work that will happen in this field. Our proposed system is expected to effectively classify the alphabets and numbers in the Indian Sign Language using two different algorithms namely: CNN and SVM. This work aims to do a comparative study and conclude the Machine Learning model that is best suitable for the considered dataset. It is observed that the accuracy of CNN is approximately around 99% whereas the accuracy of SVM is around 96%. This accuracy was tested by varying the epochs and it was concluded that there is no major change in the accuracy achieved by varying the epochs in both the models. The accuracy of the system in real-time is yet to be tested, we choose to implement CNN and SVM and did a comparative study to choose an optimal model for the dataset.

FUTURE SCOPE:

This work has been limited alphabets and numbers of the ISL (Indian sign language). Since these are static images, there need not be much pre-processing done (because for dynamic videos we have to convert videos to frames). This can be extended to dynamic datasets of the ISL (i.e videos). Due to lack of availability of the datasets, there has been no much work done on the same. PM Modi has recently launched an ISL dictionary, which consists of videos of words such as police, pay, party, hello, etc. Since, a single video is not enough for a single class, once the dataset set grows large, the same can be extended to videos as well. The same videos can eb pre-processed and fed to both CNN and SVM and choose the algorithm most appropriate for this dataset

7. References

1.Nello Cristianini, Colin Campbell and John Shawe-Taylor, "A Multiplicative Updating Algorithm for Training Support Vector Machines".

2. Brandon Garcia, Sigberto Alarcon Viesca "Real time American Sign Language Recognition with Convolutional Neural Networks"

3.Mark Devaney and Ashwin Ram, "Efficient Feature Selection in Conceptual Clustering",

4. Byung-woo min, Ho-sub yoon, Jung soh, Takeshi ohashi and Toshiaki jima," Visual Recognition of Static/Dynamic Gesture:

5. Raheja JL, Mishra A, Chaudary A. Indian Sign Language Recognition Using SVM

6.Huang J, Zhou W, Li H, Li W. Sign language recognition using 3D convolutional neural networks.

7.. Pigou L, Dieleman S, Kindermans PJ, Schrauwen B. Sign Language Recognition using Convolutional Neural Networks.

8.] PADMAVATHI . S, SAIPREETHY.M.S, V. Indian sign language character recognition using neural networks.

9. Adithya V, P. R Vinod and Usha Gopalakrishnan. “Artificial Neural Network based method for Indian Sign Language Recognition”.

10. Sharma, Madhuri, Ranjna Pal and Ashok Kumar Sahoo, “Indian Sign Language Recognition using Neural Networks and KNN Classifiers”.

11. Rashid, O., Al-Hamadi, A., and Michaelis, B, "A Framework for the Integration of Gesture and Posture Recognition Using HMM and SVM."