

SIGN LANGUAGE DETECTION USING MACHINE LEARNING ALGORITHMS

Pelluru Venkata Sai Satya Tarun Department of Electronics and Computer Engineering Vignan's institute of information technology Visakhapatnam, India pvstarun@gmail.com

Janaparedi Roopa Department of Electronics and Computer Engineering Vignan's institute of information technology Visakhapatnam, India janapareddiroopa143@gmail.com

Gorle Divya Department of Electronics and Computer Engineering Vignan's institute of information technology Visakhapatnam, India meetme.divya.gorli@gmail.com

Bellala Geethanjali Department of Electronics and Computer Engineering Vignan's institute of information technology Visakhapatnam, India bellalageetha9@gmail.com

Abstract

Communication is a very important tool or necessity. We use it to express our feelings, our pain, information etc. But there are few people who couldn't communicate and they need something to rely on so that they could communicate. Sign language can be useful but it is very difficult to understand and convert to our normal language. By using image processing techniques, we can accurately identify the signs. There are many signs based upon the language through which they want to communicate. We are using American Sign Language (ASL) to communicate. There are two modes to implement the detection of sign language, one is hardware mode and another is the software mode. In Hardware mode we have to build an arm to deliver sign language which is expensive and difficult to make, instead we can switch to machine learning algorithms to do the same thing even better. The main problem with the machine learning is to find the optimal algorithm for sign language detection. So primarily we have chosen SVM (Support Vector Machine), CNN (Convolutional Neural Network), Naive Bayes, KNN (K Nearest Neighbors) algorithms to test the accuracy. But we had to collect images for each alphabet so, we collected photographs of different signs. We used OpenCV to convert those images to gray scale and then we trained our models with the training data set. Then we tested it with different algorithms. We found Convolutional Neural Network produces accurate results compared with KNN, Naive Bayes, SVM. Thus, we concluded that CNN is the optimal algorithm for sign language detection.

Keywords: Sign language, Machine learning, CNN, Naive Bayes, SVM, KNN, ASL.

Introduction

Communication is a very important medium and became a very vital part in each and every life on this planet earth. This is considered as a fortune for human beings because of the intelligence human beings possess. Human's need always led to new discoveries and invention. But what if the communication is a barrier? People who are deaf and dumb cannot communicate properly as they lack the organs that work perfectly. That's the reason why we invented sign language. Movement of hand gestures which make a unique word or a saying which helps the deaf and dumb people to communicate with others [10]. Sign language is a basic hand Gestures gestures instead of sound [11]. But there used to be a translator always who understands the sign language and translates it to normal human language. But that's a very difficult process, because we can't have human translator every time with us and learning ourselves sign language is very difficult. And this tradition has been from ages. But in this hurry world we have to find an optimal solution in order to continue the work flow of this human race. Computer vision is the technology emerging and people are switching to facial recognition and this could solve the problem of language translation. There are basically two ways to solve such a problem. One is to make a hardware that could literally convert the sign language to text and text to sign language. But this is very expensive and complex to make. There is another way, and that is as mentioned above using computer vision. Computer vision has the capability to not only mimic the human actions, but also learn and extract the unknown actions which make it precise and give a scope to the future to study about it. So, machine learning plays a vital part in this computer vision. So, by extracting data through images and classifying them with different classifiers we have in machine learning.

Literature Survey .

C.R.Dyer, Guidong guo stated Because there is a lack of contact with deaf people in our society, the system's deployment here is based on image representation [1]. Narendra Ahuja and Sinisa Todorovic This study proposes that a human-interaction system based on sensing human hand movements be designed and implemented. J Sreedevi, Dr M rama bai ,mohammed Maqsood ahmed The quickest approach to solve the problem of gesture recognition is to use CNN[3]. Yanting pei, yapping huang, qi zou, xingyuan zhang, song wangy discovered that when a picture is deteriorated, image classification performance suffers dramatically, especially when the training images are unable to accurately reflect the degradation levels of the test images [4]. Hoo-chang shin and holger R. Rohith Data set characteristics, deep convolutional neural network (CNN) architecture, and transfer learning[5]. Bing Tu, Jinping Wang, dong Kang, Guoyun Zhang, Xianfeng Ou , and Longyuan Guo On four genuine hyperspectral datasets, experimental results show that the suggested method may achieve substantially greater classification accuracies, especially when the number of training samples is limited .Because pixels of multiple classes coexist in the same super pixel region, the pixels in these super pixels are unable to be classified efficiently. Second, the computational cost imposed by the KNN operation is a significant issue [6]. yanhui guo, han cao1, siming han, yunchuan sun , yu bai Methods for classification that combine joint representation KNN with a guided filter. To extract spatial information, a front guided filter approach is used [7]. Li Ma, Student Member, Melba M. Crawford, Fellow, and Jinwen Tian ULML and SLML as DR pre-processing methods for the KNN classifier were examined [8]. Emma Izquierdo-Verdiguier and Luis Gómez-Chova The support vector machine (SVM) classifier has been effectively applied to the classification of remote sensing images. This method has been frequently used in the classification of contextual images [9]. walaa aly , saleh aly and sultan almotairi Though signs and gestures are categorized as non-verbal communication, they are regarded the most natural way to express messages among individuals through body motions [10]. Md Azher Uddin , Shayhan Ameen Chowdhury consider The most fundamental mode of communication is sign language. This entails employing hand gestures and nonverbal communication rather than sound or spoken words. Deaf and dumb persons communicate via hand gestures, which are a graphic method of communication including a combination of hand posture [11]. Mathur and G. M. Foody. The possibility of one-shot multiclass SVM for the categorization of remotely sensed data was discussed by a member. Not only did the one-shot multiclass SVM classification have the maximum accuracy, but it also used the fewest support vectors [12]. Yun-Fu Liu, Jing-Ming Guo, and Jiann-Der lee did a classification using naïve bayes and LMS [13].

Research Methodology

The problem is to find an algorithm for sign language detection which gives you an optimal result with more accuracy. But there are many ways to detect sign language. Finding the best suitable algorithm is the trickiest part. There are many people who are suffering with lack of communication and I guess communication is the biggest tool to enrich the humanity. There are many ways to solve such problems and I guess there are only two modes of approach. Those are software and hardware modes. We can collect the motions and create actuator that would actually create motions by sensing and create sign language [2]. Hardware has a lot of friction because of a few factors. They are costly and very hard to make, and not so portable. Through image processing techniques it is easy to detect the sign language and this could be the best way to create a medium for communication and help them to communicate with others [1]. For image recognition we need images as the data set. Basically, the data set comprises training data and also testing data. The image data should be the hand symbols of alphabets. Each and every alphabet has a sign language. So, we are going to collect the data manually by taking photographs of each sign and making those symbols by hand. We had to take more than 300 photos of each symbol. So, we collected the data and stored them in separate folders in our local system. We have not relied on any website to extract the images. The images were completely photographed by ourselves. We have parallelly developed the model for sign language but for the machine to understand the images, we had to convert the images to other types of images like Gray, canny, HSV. Any algorithm understands data in its binary format. Our algorithm cannot recognize those high-quality images. The algorithm can understand the images in

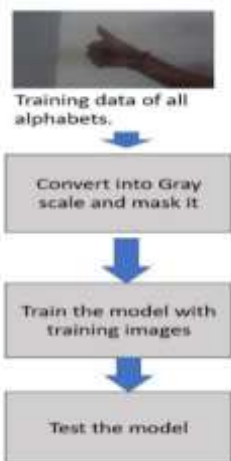
its truest format. So, we are going to convert those images into Gray images so that it would be easy for any algorithm to understand the images. After all those are algorithms, we need to make them understand the characteristics of an image. The reason why we are converting them into Gray is to make the data set as lite as possible, because responsiveness is much more important than anything in the world. If we use color images it would take way more time than anyone could possibly think of because color image has three channels called RGB. The values of RGB are in the form of matrix and there are three with range of 0 to 256 values (depending upon the intensity of the color). We have used HSV (Hue saturation value) to separate the luminance. We have also used canny for edge detection and masking process to detect the target object.

We have collected the data and we have also converted them into its meaning full form so there is one thing left to do and that's the main part. We have to select the algorithm. In order to convert the images to Gray for HSV or canny, we are using python module called open cv for computer vision.

We have used Naïve bayes, Support Vector Machines, KNN and Convolutional Neural Network (CNN) classifiers for sign language detection in this paper. First, we are going to use Convolutional Neural Network which is a machine learning algorithm used for image classification. CNN is the most expeditious way for image classification [3]. We are going to use python and its libraires and modules to implement this algorithm. For using CNN, we are going to use TensorFlow frame work. We are going to use RELU activation function for this CNN algorithm. We are going to pass the converted images to train the algorithm. Convolutional Neural Network is a machine learning algorithm which is a neural network (combination of neurons form a network) which is a classifier used in classification of data in supervised learning, also used in Natural language processing. Performance of the algorithm also degrades with the image degradation [4]. We have also considered checking out images used in classification in medical field but that was as same as we do in this paper [5]. We have trained all the alphabets symbols to the CNN model and tested it with the testing data. Naïve bayes is a machine learning algorithm used in classification problems. We are going to feed the training data into the model which is a function. Naive bayes algorithm can also be combined with LMS but Naïve bayes alone is sufficient [13]. And later the accuracy results are collected. KNN (K-nearest -neighbor) is a machine learning algorithm also used in classification problem. But the problem with the KNN is it may cause computational burden [6]. Trained images were given to the KNN function and accuracy is calculated. We could have used the guided filter but that isn't that necessary [7]. A new SLML based KNN classifier could be used for classifying the hyperspectral image but we have used masked images [8]. Support Vector Machine (SVM) is a classifier used in machine learning algorithm. SVM's are used in contextual image classification and are very efficient [9]. Multi class SVM are major disadvantages due to its parameter optimization. But to test its accuracy, we have passed the training data to the SVM function and extracted the accuracy results [12]. Here we have used python IDLE for executing this algorithm. In order to visualize the results with graphs, we used matplotlib library.

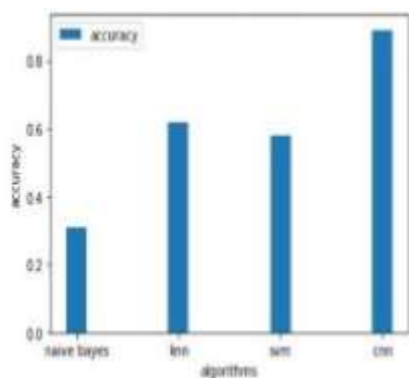
FLOWCHART:

The above flow chart explains how the overall algorithm of identifying an image works. First, we had taken hundreds of photos of different sign with hand. The different signs are just English alphabets. We took the photos and tried to convert them into gray images. Before going for testing, the images are converted into gray level and masking is applied. And then algorithm train with the images. After completion of training the classifier is ready to test. We test the model using testing data and check for the results.



Results and Discussion

The results clearly state that Convolutional Neural Network algorithm stands the best algorithm for sign language detection. CNN doesn't lose any quality or information even though we reduce parameters. K- nearest neighbors stand the second with 63% accuracy. Support vector machine stands 3rd with 56% accuracy. Naïve bayes stands the last with least accuracy percentage of 33%. So Convolutional Neural Network is proven to be the best algorithm with highest accuracy.



classifiers	accuracy
KNN	63%
CNN	92%
Naïve bayes	33%
SVM	56%

VI. CONCLUSION:

Convolutional Neural Network stands as the best for sign language when compared to other algorithms like K Nearest Neighbors, Support Vector Machine and Naïve bayes classifiers. Despite of losing parameters, CNN managed to maintain the quality of the image and gave us the highest accuracy. So, we conclude that Convolutional Neural Network is the best suited algorithm for sign language detection.

References

[1] C.R.Dyer, Guidong guo "Learning from examples in the small sample case:face expression recognition" IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) (Volume: 35, Issue: 3, June 2005).

- [2] Narendra Ahuja and Sinisa Todorovic "Unsupervised Category Modeling, Recognition, and Segmentation in Images" IEEE transactions on pattern analysis and machine intelligence, vol. 30, no 12, december 2008
- [3] J SreeDevi, Dr M Rama Bai, Mohammed Maqsood Ahmed " An Approach To Translate Hand Gestures Into Telugu Text" INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 9, ISSUE 04, APRIL 2020.
- [4] Yanting Pei, Yaping Huang, Qi Zou, Xingyuan Zhang, Song Wang "Effects of Image Degradation and Degradation Removal to CNN-based Image Classification" journal of latex class files, vol. 14, no. 8, august 2015 .
- [5] Hoo-Chang Shin, Member, IEEE, Holger R. Roth, Mingchen Gao, Le Lu, Senior Member, IEEE, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel Mollura, and Ronald M. Summers "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning" iee transactions on medical imaging, vol. 35, no. 5, may 2016.
- [6] Bing Tu , Member, IEEE, Jinping Wang, Student Member, IEEE, Xudong Kang , Senior Member, IEEE, Guoyun Zhang , Member, IEEE, Xianfeng Ou , Member, IEEE, and Longyuan Guo, Member, IEEE " KNN-Based Representation of Superpixels for Hyperspectral Image Classification" iee journal of selected topics in applied earth observations and remote sensing.
- [7] yanhui guo 1,2, han cao1, siming han1, yunchuan sun 3, (Senior Member, IEEE), AND YU BAI 4, (Member, IEEE)" Spectral_Spatial Hyperspectral Image Classification With K-Nearest Neighbor and Guided Filter" Received January 30, 2018, accepted March 14, 2018, date of publication March 28, 2018, date of current version April 23, 2018. Digital Object Identifier 10.1109/ACCESS.2018.2820043.
- [8] Li Ma, Student Member, IEEE, Melba M. Crawford, Fellow, IEEE, and Jinwen Tian "Local Manifold Learning Based k-Nearest-Neighbor for Hyperspectral Image Classification" ransactions on geoscience and remote sensing, vol. 48, no. 11, november 2010.
- [9] Emma Izquierdo-Verdiguier, Student Member, IEEE, Valero Laparra, Luis Gómez-Chova, Member, IEEE, and Gustavo Camps-Valls, Senior Member, IEEE "Encoding Invariances in Remote Sensing Image Classification With SVM" iee geoscience and remote sensing letters, vol. 10, no. 5, september 2013.
- [10] Md Azher Uddin , Shayhan Ameen Chowdhury " Hand Sign Language Recognition for Bangla Alphabet using Support Vector Machine" 978-1-5090-61228/16/\$31.00 ©2016 IEEE.
- [11] walaa aly 1,2, saleh aly 1,2, and sultan almotairi 3 " User-Independent American Sign Language Alphabet Recognition Based on Depth Image and PCANet Features" Received July 25, 2019, accepted August 28, 2019, date of publication September 2, 2019, date of current version September 12, 2019.
- [12] A. Mathur and G. M. Foody, Member, IEEE "Multiclass and Binary SVM Classification: Implications for Training and Classification Users" iee geoscience and remote sensing letters, vol. 5, no. 2, april 2008.
- [13] Yun-Fu Liu, Student Member, IEEE, Jing-Ming Guo, Senior Member, IEEE, and Jiann-Der Lee, Member, IEEE "Halftone Image Classification Using LMS Algorithm and Naive Bayes" iee transactions on image processing, vol. 20, no. 10, october 2011.