DOI: http://dx.doi.org/10.26483/ijarcs.v13i3.6828

ISSN No. 0976-5697

Volume 13, No. 3, May-June 2022

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

MACHINE LEARNING ALGORITHMS FOR ASL IMAGE RECOGNITION WITH LENET5 FEATURE EXTRACTION

Ratan Kumar Sajja
Dept. of Computer Science and Engineering
Anil Neerukonda Institute of Technology and Sciences
Visakhapatnam, India

Vandana Killari
Dept. of Computer Science and Engineering
Anil Neerukonda Institute of Technology and Sciences
Visakhapatnam, India

Satya Avinash Nimmakayala
Dept. of Computer Science and Engineering
Anil Neerukonda Institute of Technology and Sciences
Visakhapatnam, India

Varun Ippili
Dept. of Computer Science and Engineering
Anil Neerukonda Institute of Technology and Sciences
Visakhapatnam, India

Abstract: American Sign Language is used by mute and deaf people so that they can interact withthe people around them. It is used by approximately 2,50,000-5,00,000 Americans (and some Canadians) of all ages.Over the period of time, many have proposed different methods for recognition of ASL. Sign language Recognition is a complex technical problem due to the difficulty of visual analysis of hand motions and the highly structured nature of sign language. Hence the accuracy is not achieved. To enhance this accuracy, the proposed system compares different machine learning classification algorithms using Lenet5 architecture for feature extraction. Lenet5 is one of the architectures of Convolution Neural Network (CNN). The proposed system uses machine learning algorithms like Neural Network, Decision Tree Classifier, K-Nearest Neighbour and Support Vector Machine. There is a raise in accuracy using Neural Network algorithm when compared to other machine learning algorithms. The proposed solution was tested on data samples from ASL data sets and achieved an overall accuracy of 99.99% using Neural Network.

Keywords:Human Computer Interaction (HCI)**,**Convolution Neural Network (CNN),American Sign Language (ASL), Hand Gesture Recognition (HGR), Sign Language Recognition (SLR).

I. INTRODUCTION

Generally deaf and mute people communicate less with the normal people in the world. So, itwould be difficult to understand what they are telling through signs. But with the help of HCmechanism we can remove the communication problem with individuals through devices. SomeHCI user interfaces are Non-touch, Speech recognition, Gesture recognition are mostly used but, it is expensive. So, everyone cannot afford it. Some standard features should be there in order toachieve the optimal design are simplicity, scalability, flexibility, and precision. Among different HCInow a days speech and hand gestures usability has increased a lot. Hand gestures are widelyused in augmented realities (SLR), video games, sign recognitions, cars. Based on SLR this papersuggests feature extraction through CNN. The HGR has many uses in various areas, such as video gaming, sign language recognition and augmented reality (SLR). Between these, an SLR is the best widely utilized method where voice communication is difficult. On this perspective, this paper suggests an effective method for the extraction of features via CNN. The purpose of this paper is to offer the extraction of features by CNN and classify using different machine learning classification algorithms. The main goal is to provide an intelligent application with better accuracy among different classification algorithms.

In Section 2 of this paper, we have discussed about Literature Review in which it describes about previous existing models for ASL image classification. In the section 3, we will talk about the proposed model about feature extraction and different machine learning algorithms. In the section 4, we have discussed about the results and estimations for each algorithm. In the section 5, we have discussed about conclusion.

II. LITERATURE REVIEW

Hand gesture recognition system received great attention in the recent few years because of the ability to interact with machine efficiently through human computer interaction. Many kinds of research have also been done on hand gesture interpretation using variety of methods. Jayshree R.Pansare et al.(2016) have implemented vision-Based approach using Edge orientation Histogram. They use canny Edge Detection to significantly reduce the amount of data and to filter out the useless information in an image and should be able to mark as many real edges as possible. After finding the edges of the hand image using canny edge detection technique, the Edge Oriented Histogram features are extracted from it. The results of the proposed system in respect to hand sign recognition is 88.26% efficiency with recognition time of 0.5 second [1].PonlawatChophuk et al.(2018) introduced American sign Language recognition using Leap Motion Sensor. They used an external device(Leap Motion Sensor) Fist alphabets were detected by Leap Motion Sensor which can detect finger bone and provides palm and tip position of fingers. They used Shoelace Algorithm and Decision Tree algorithm is a mathematical algorithm to determine the area of a simple polygon whose vertices are described by their Cartesian coordinates in the plane. This System detects only Fist Alphabets. The sign recognition score of the system is 96.1% [2].Deepali Naglot etal.(2017) suggested a model Sign Language Recognition using Leap Motion Controller(LMC).LMC is a 3D non-contact motion sensor which can tracks and detects hands, fingers, bones and finger-like objects. They have used Multilayer Perceptron neural network(MLP) algorithm. Multi-Layer Perceptron(MLP) is self-learning, supervised algorithm used for pattern recognition and classification. The Suggested model identifies 26 of Alphabets.The accuracy of the model is 96.15% [3].Rim Barioul et al.worked on wrist force myography to identify sign language symbols. Myography is a technique used to find the intensity of the muscle while the muscle is in contraction. By using this technique with the help of four sensors they have identified only 9 symbols with an accuracy of 89.65% by using the Select Vector Machine(SVM) and Least Discriminant Analysis(LDA) algorithms. Main advantage of this work is that the identification rate is so fast [4].

III. PROPOSED MODEL

The application is configured to obtain images. The dataset is collected where the images are in the form of grayscale. These images are used to train the model. These images are used for feature extraction using CNN (Convolution Neural Network. The CNN model will create a numpy array of features. The numpy array is converted into .csv file. The suggested model will use that .csv file for training and classification. This will be done using different machine learning algorithms.

The basic principle of our model is proposed to classify the ASL (American Sign Language) alphabet, based on the human hand gesture. The operating method for the suggested model is seen in Figure 1.

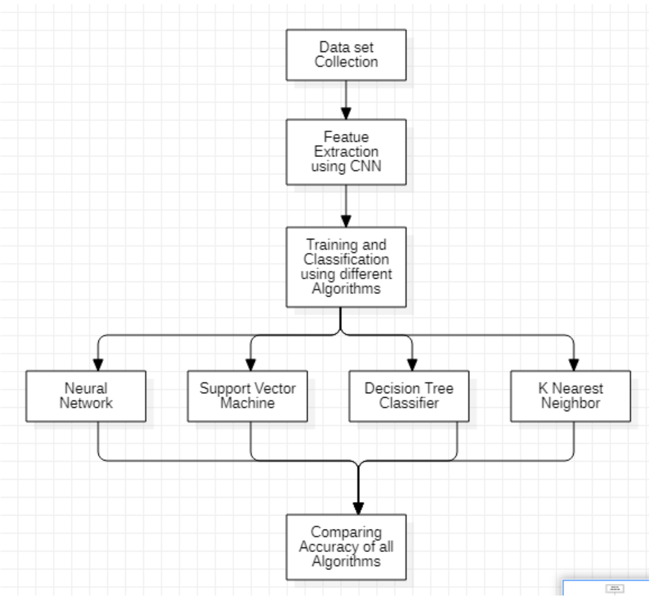

Figure 1. Operational process of the suggested model

A. Data Collection:

The dataset is collected from the Kaggle website. There are 27,455 samples in the dataset. The image size is 28x28x1. There is a total of 784 pixels for each row. There are nearly 1,000 samples for each sign. There are images of all ASL alphabets except for J and Z which are dynamic letters. Any representation of the sign language can be used, the interpretation proposed is for the American Sign Language Alphabet, which is shown in Figure 2.

https://www.kaggle.com/datasets/datamunge/sign-language-mnist

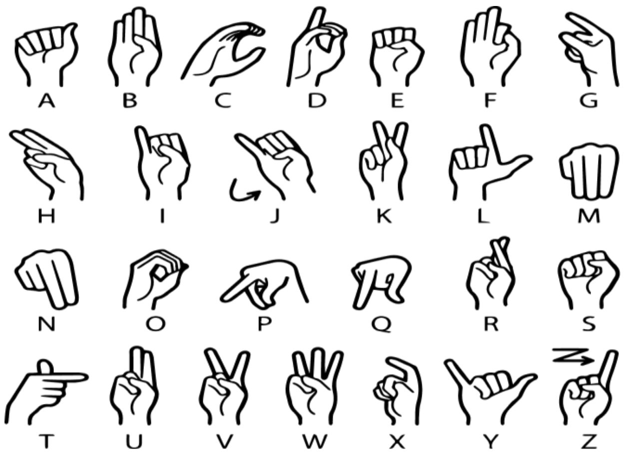

Figure 2. Alphabet ASL Images

B. Feature Extraction:

Many algorithms of machine learning are used to extract features. One of the strongest deep learning strategies is Convolution Neural Network (CNN). A broad of different images varies that CNN is used. CNN will extract features for the classification model across wide variety of pictures. There are so many architectures for the CNN model. The proposed model uses the Lenet5 architecture is seen in Figure 3 and Figure 4.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 24, 24, 6)         156
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 12, 12, 6)         0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 8, 8, 5)           755
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 4, 4, 5)           0
_________________________________________________________________
flatten (Flatten)            (None, 80)                0
=================================================================
Total params: 911
Trainable params: 911
Non-trainable params: 0
```

Figure 3. Lenet5 model description

C. Training and Classification

All machine learning classification algorithms takes 80 features of input from deep learning algorithm Lenet5.

Neural Network

The proposed network is one hidden layer with 120 neurons. The features will be fed to the hidden layer which is a fully connected layer. The SoftMax activation function is used to differentiate each alphabetical sign as a part of this suggested model. This function provides the distribution of the likelihood between mutually exclusive output classes. In SoftMax, it will give probability for each alphabet. According to those probabilities, the final decision is made for this procedure by choosing the maximum output probability [5].

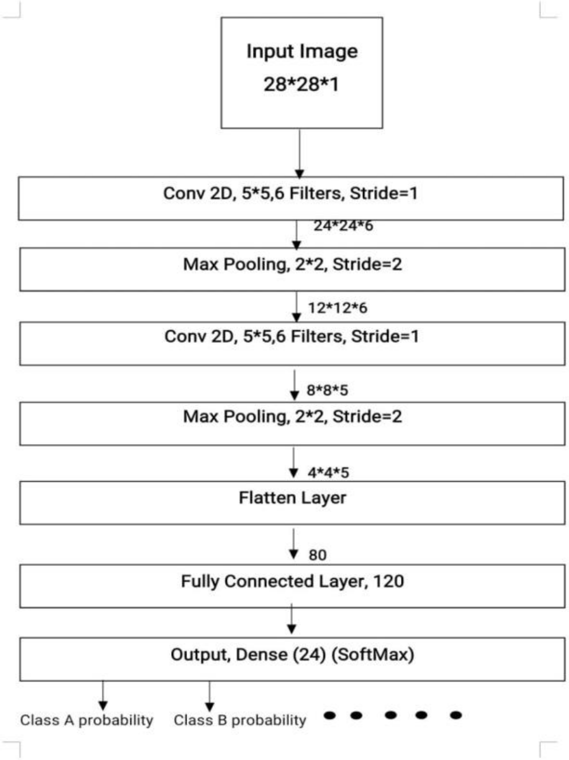

Figure 4. Lenet5 Architecture

Decision Tree Classifier

Out of these 80 features, the feature which contributes towards classification will become the root node and the next most contributed feature will be the child node of the root. Likewise, the entire tree is constructed. Hence the leave node defines the class to which it belongs to.

K Nearest Neighbor (KNN)

Out of these 80 features, the features which are relevant to the class it belongs to by calculating the distance of all the classes. The minimum distance will signifies that the feature is belongs to that class.

Support Vector Machine (SVM)

The SVM constructs a graph and a line is drawn (classification line). The portion that is on the left side of the line belongs to one class and the portion on the right side of the line belongs to another class.

IV. Result and Estimation

Neural Network

Python with the Keras and TensorFlow backend libraries was used to apply the suggested deep learning algorithm. The dataset is split into two groups. The first branched collection comprises 70 percent for training images and the second one includes the remainder of the 30 percent for testing images. The model accuracy about 99.99% which is seen in Figure 5 and Figure 6.

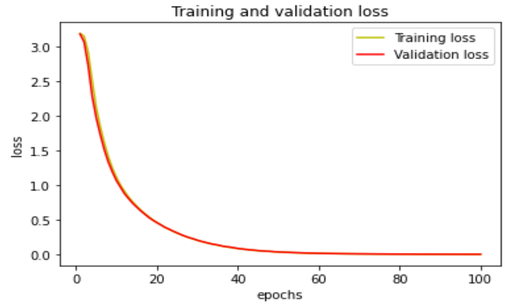

Figure 5. Graph of loss against number of epochs

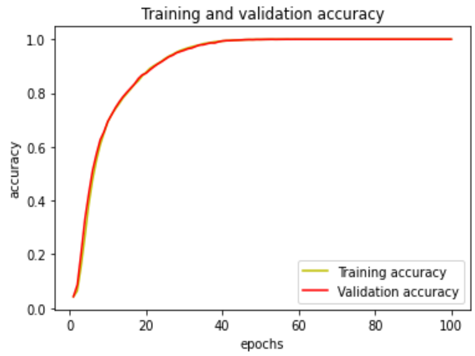

Figure 6. Graph of accuracy against number of epochs

With the more detailed features of training and testing, each of the signs individually performed by the SoftMax has been identified.

Decision Tree Classifier:

Python with scikit-learn backend libraries was used to apply the suggested Decision tree algorithm. The dataset is split into two groups. The first branched collection comprises 85 percent for training images and the second one includes the remainder of the 15 percent for testing images. After performing training with this classifier we have obtained an accuracy of 67%.

K Nearest Neighbour (KNN)

Python with scikit-learn backend libraries was used to apply the suggested K-Nearest Neighbor. The dataset is split into two groups. The first branched collection comprises 80 percent for training images and the second one includes the remainder of the 20 percent for testing images. After performing training with this algorithm for k = 1 we have obtained an accuracy of 97% which is shown in Figure 7.

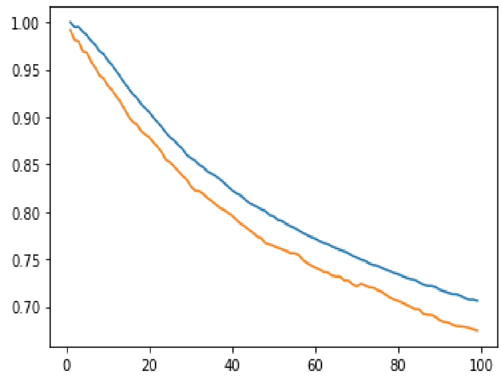

Figure 7. Graph of training and testing accuracies against the k values for proposed model

Support Vector Machine (SVM)

Python with scikit-learn backend libraries was used to apply the suggested Support Vector Machine. The dataset is split into two groups. The first branched collection comprises 85 percent for training images and the second one includes the remainder of the 15 percent for testing images. After performing training with this algorithm we have obtained an accuracy of 80%.

V. CONCLUSION

Hand gesture recognition is a big problem in real-life implementations concerning the precision and reliability correlated with. This paper introduces the static hand gesture recognition in ASL. As shown in table 1, Classification accuracy was attained by99.99% when using neural network compared to other different classification algorithms like Decision tree classifier, K-Nearest Neighbor and Support vector machine.

Table 1: Comparing the state of the art with different classification algorithms

S. No.	Algorithm	Training accuracy	Test accuracy
1	Neural Network	0.99	0.99
2	Decision tree classifier	0.86	0.66
3	K-Nearest Neighbor	0.99	0.97
4	Support Vector Machine	0.80	0.80

VI. REFERENCES

[1] Jayashree R. Pansare and Maya Ingle, “Vision-based approach for American SignLanguagerecognition using Edge Orientation Histogram”, no. September, 2016, doi:10.1109/ICIVC.2016.7571278

[2] Ponlawat Chophuk, Kanjana Pattanaworapan and Kosin Chamnongthai, “First american signlanguage recognition using leap motion sensor”, no. January, 2018, doi:10.1109/IWAIT.2018.8369790

[3] Deepali Naglot and Milind Kulkarni, “Real time sign language recognition using leap motion controller”, no. August, 2017, doi:10.1109/INVENTIVE.2016.7830097

[4] Rim Barioul, Sameh Fakhfakh Ghribi, Houda Ben Jmaa Derbel and Olfa Kanoun, “Four SensorsBracelet for American Sign Language Recognition based on Wrist Force Mygography”, no. July,2020, doi:10.1109/CIVEMSA48639.2020.9132742

[5] Azher Atallah Fahad, Hassan Jaleel Hassan and Salma Hameedi Abdullah, “Deep learning-based Deaf and Mute Gesture Translation System”, no. May,2020, doi:10.21275/SR20503031800