TAMIL SIGN LANGUAGE RECOGNITION IN REAL TIME

Manisha. A[1], Neela Hairsh. M[2], Thirekha.B.T[3], Kavya varshini.AG[4], Pradeepa.S[5]

[1] *Post Graduate Student, Dept. of Biomedical Engineering, Rajalakshmi Engineering College, Chennai, India*
[2] *Assistant Professor, Dept. of Biomedical Engineering, Rajalakshmi Engineering College, Chennai, India*
[3,4,5] *Dept. of Biomedical Engineering, Rajalakshmi Engineering College, Chennai, India*
Email: [1]manisha.a.2019.meme@rajalakshmi.edu.in, [2]neela.m@rajalakshmi.edu.in, [3]thirehka.bt.2017.bme@rajalakshmi.edu.in, [4]kavyavarshini.ag.2017.bme@rajalakshmi.edu.in, [5]thirehka.bt.2017.bme@rajalakshmi.edu.in

Abstract: Communication is interchanging of ideas information or message from one person to the other person, sign language is made with the hand and other movement, including facial expressions and posture of the body used by the people of unable to speak and hearing, there are different type of sign language. Tamil sign language is the regional sign language, the aim of the work is provide the real time recognition of Tamil sign Language (TSL) in to Tamil letter, here we introduced the convolutional neural network (CNN) as a classifier used as training the Tamil sign language and predict the Tamil sign language, the process are divided in two section one is to train by using the keras model and the second section is the skin segmentation of the hand gestures in the region of the interest, there are two phase in training the model, having set of the hand gestures images in the training set and testing set by using the image train the model, the model occurred the 100% accuracy in the white background and good lighting condition, 97.5 % in the low lighting condition.

Keywords: **Tamil sign Language, CNN, and Keras**

1. Introduction

Sign language is used by the deaf and dumb people to convey their thought and emotion, the sign language different from one country to other country, one region to the other region in sometimes it different from the one state to the other .The sign language recognition can be identified by the two method one is placing the sensor in hand, by the motion variance the sign can be detected, other one is camera based the sign is captured by the camera, the prediction of the sign has to be identified, the hand gesture vary from the one handed, double handed, static and dynamic.

There are different type of sign language used in the different country they are American Sign Language (ASL), Chinese Sign language (CSL), Indian Sign Language (ISL) and Japanese Sign Language (JSL).

Deep learning, there are several types of models such as the Artificial Neural Networks (ANN), Auto encoders, Recurrent Neural Networks (RNN) and Reinforcement Learning. But there has been one particular model that has

contributed a lot in the field of computer vision and image analysis which is the Convolutional Neural Networks (CNN).

CNNs are a class of Deep Neural Networks is used to identify and predict the individual features from images and it is broadly used for analyzing the images. Image and video recognition application are classification of image, computer vision and natural language processing.

2. Literature Survey

The intention of the paper is the real time model which convert the Indian sign language in to text. There are two phase in this paper in the first phase the execution of the convolutional neural Network using the keras, in the phase two other method is used to detect the region of interest by using the skin segmentation, the segmented area is feed in to the model to identify the sign, the accuracy obtained in this paper is 99.56% in the same subject, in the low light condition the accuracy is 97.26% (1).

The research paper introduced the automated recognition of Indian sign language numeric sign, the database of numeric sign 5000 signs, the image per numerical sign is 500 images, there are two technique used to extract the feature from the sign image is direct pixel value and hierarchical centroid techniques, after the feature extraction to classify the image by using neural network and knn classification technique, the accuracy of the experiment is 97.10% accuracy, the dark background is selected for the background of the hand gestures (2).

The images are taken from the kinetic sensor the first step is pre-processing image and the second step is the segmentation the classifier classify the signals, the third step is feature extraction to extract the feature of the particular hand gesture, the final step is recognize the hand gesture, the image was resized in to200 × 200 pixels by using this image the database was created, the accuracy of one word is different from other word, comparing the accuracy the highest accuracy rate is 92.24% (3).

This paper discussed the various technique of classifier used to identify the sign language recognition technique, here focusing the deep learning Based classifier is used to recognise the hand gesture based on the deep learning technique in Arabic sign language recognition, in the deep learning technique of some classifier are discussed here such

as convolutional neural network, recurrent neural network, multilayer perceptron, linear discriminant analysis, hidden Markov model, artificial neural network, support vector machine, k nearest neighbour are discussed here with the accuracy (4)

The principle focal point of this framework is to make a vision based system, a Convolutional Neural Network (CNN) model to recognize the six distinctive gesture based communications from the image taken here two CNN models was created have divergent type of optimizers, the Stochastic Gradient Descent (SGD) and Adam, the convolution neural network have different layers such as Convolutional layer, Pooling layer, Dropout layer and Fully Connected layer, the resized input image pixel is 64×64 and the first layer is activated by the rectified linear unit, hand gesture of different individual are taken, the Adam optimiser having more accuracy then the SGD (5)

This paper presents a learning based way to deal with Bangla Sign Language (BdSL) recognize utilizing the convolutional neural network, here we using leep motion controller by using lmc hand position, movement can be extracted, from the continuous frame we extract the hand gesture by using Hidden Markov Model (HMM), by using convolutional neural network we train the data, the error rate is 3% (6).

In this paper proposed the work of static finger spelling in Japanese sign language, identify by the single grey image, finger spelled sign language can be identify by the convolutional neural network, here 5000 image in the dataset and 1000 images are captured by using rotating and shifting phases . the image are captured by RGB camera, the convolutional neural network having six layers, the input image size is 32×32 pixel gray image, here testing samples are 80% and testing samples are 20%, the accuracy rate of the system is 93% (7).

This paper developed the American sign language finger spelling translator by using Google net translator trained the ILSVR2012 dataset, train the model between a-y again initialised the classification layer with Xavier, drastically increase the batch size is from 4to 20, the accuracy was different from 1_int and 2_int, there 22 deep layers the robust model for letters a-e and a modest letter for a-k. The largest difference of two layer accuracy is 7%, the accuracy acquired for the five letters 98% and ten letters is 74% (8).

This paper work on the convolutional neural network and recurrent neural network and by suing CNN to identify the spatial features and by using RNN to train the temporal feature, the database here used is American Sign Language (ASL), to extract the temporal feature from the feature to predict the hand gestures, each signs of 2400 image in the dataset, the training set have 1800 image and the testing set have 600 images, the classification done by the two different layer, the prediction is done by the softmax layer and the output is done by the global pool layer both the model is trained by the cross – entropy cost function ADAM, the batch size of the model is 64 and the total epochs used here is 10, the drawback of the model is varying the cloth the poorly performing model (9)

This paper proposed the methodology of identify the Tamil sign language recognition by using the algorithm of scale – Invariant feature transform and the edge is detected by the canny edge detector, the accuracy of the model is 97.31%, here using the right hand the hand gesture are differ by varying the finger up and down and the 31 alphabets of hand gesture can be obtained, the accuracy of the SIFT is higher than the existing methodology, the Tamil sign language is categorised by the regional category (10).

This paper proposed the recognition of Chinese sign language by using convolutional neural network, the database extract from the 40 daily vocabularies, the accuracy rate of the recognition 99%, the movement of the Chinese sign language video are over 0.1 s length, the sample rate of the video is 1000 rate per second, the full set is divided in two section in this first section 50% of training set and 50 % testing set and in the second section 80% of training set and 20% testing data and the validation rate is 5 %, there are two learning algorithm here used adagrad and ADADELTA, the accuracy of adagrad is 90%, the convolutional neural network achieve the accuracy of 99% (11).

This paper proposed collecting the data of 10 selected words of Korean sign language, the input image with 9 frames and developing the convolutional neural network based model, the accuracy of the Korean sign language for the 10 words is 84.5%, the accuracy of the testing data by the training model is 80%, the accuracy level is differ from one word to the other, the highest accuracy of the word is happy and hello, the convolutional neural network distinguish the bright colour and the light colour, the overall accuracy is 80% (12) .

In this paper the hand gesture date is acquired by using the flex sensor and accelerometer, the sensor is fixed in the fingers and the wrist to analyse the change of dynamic hand gestures, acquiring the voltage difference from the sensor is processed by the microcontroller and it send the signal to the module and the voice is stored and the speaker produce the voice, the eight word of the Indian sign language are processed (13)

This paper propose the study of the different type of the technique used for the different type of sign language recognition and what are the challenges come across for the recognition of the sign language (14).

3. Proposed System

Tamil sign language is the regional language, the proposed system is to develop the system of real time application of the Tamil sign language recognition using the

convolutional neural network, and the flow of the system as follows

- Capturing the frame
- Find the hand position
- Hand gesture extracted by using HSV
- Classifying and identification

Block Diagram/Process Diagram
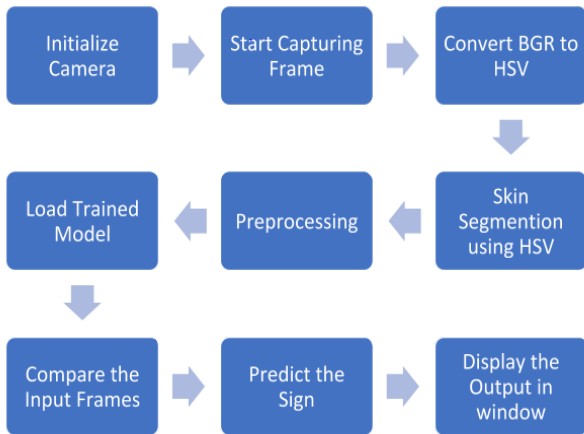
Fig.1.1Block Diagram

The problem of hand gesture identification is finding the position of hand in the frame is difficult, the region of interest in the form of the rectangular box, inside the rectangular box positioned the hand to identify the hand gesture is easy.

In the Region of interest, the skin segmentation is done by using HSV (Hue saturation value) variation of the value, in this system proposed the HSV value is given in two different form one is lower HSV value and other one is upper HSV value, the feature is extracted and fed-up in to the model for the training and the prediction.
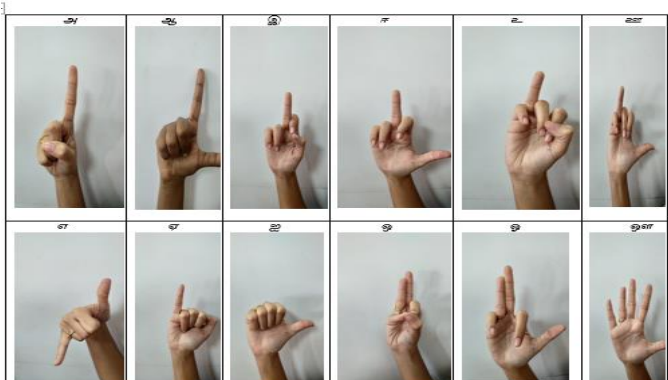
Fig 1.2 Different Tamil Letters Signs

This layer is the first layer of CNN that is used to extract the various features from the input images. In this layer, between the input image and a filter implementing the

convolutional mathematical operation to convert the image in particular size

Two main process of CNN

A convolution tool that separates and identifies the various features of the image for analysis in a process called as Feature Extraction

A fully connected layer that utilizes the output from the convolution process and predicts the class of the image based on the features extracted in previous stages.

The basic Architecture of the CNN having different layer and mentioned below

- Convolution layer
- Pooling layer
- Fully connected layer
- Dropout
- Activation Function

The region of the interest is feed in to the first input layer, the size of the input layer is 64×64, the convolutional layer is applied in the form of 3×3 kernel with the 32 filter, applying the pooling layer in the convoluted image, the size of the pooling layer is 2×2 and the pooling layer reduce the size of the image is 32×32 and after three layer of the convolution and pooling the image size is reduced and the activation function of the convolutional layer and pooling layer is RELU (Rectified Linear Unit). In simple terms, here two images are consider as matrices are multiplied to give an output that is used to features from the image. The filter size used here in the convolutional neural network is 3×3. The Pooling Layer usually serves as a bridge between the Convolutional Layer and the FC Layer, The primary aim of this layer is to decrease the size, In Max Pooling, the largest element is taken from feature map, here we using the max pooling the pooling size is 2×2 it decrease the exact size of the image in to the half stride 2. In the pooling layer we using the activation function called RELU.

The third step is flatten the convolution layer and the pooling layer, the processed the image is linear array, it act as a input of the next layer, the fully connected layer consists of the weights and biases along with the neurons and is used to connect the neurons between two different layers. Here we using the activation function of the fully connected layer is Softmax, here we using the 500 epochs in each epochs the accuracy level is increase, In the compiling step of convolutional neural network consisting of optimiser and loss function the optimiser here used is SGD and the loss function is cross entropy, in the output the error is calculate, based on the error the weight is adjusted, the trained model predict the image in the real time.

4. Experimental Result

The system has trained using the 350 images in the training set and 250 image in the testing set , the image are captured through the camera QHM4952M, the image are trained in the system, the system take 35 minutes to train the model using 350 images of the single sign, the learning rate is 0.01, the system is occurred 97% accuracy in 30 epochs, in python implementing the deep learning, The method here used is Tenser flow and keras, tensor flow is flexible to quickly implemented the model and deploy the model .

The model is tested with the 250 images per the individual sign and it display the good result while testing the image, the output contain two frames one is recognised sign and mask hand gesture, the first one is hand gesture identification in the normal camera and the second one is HSV processed image and the track bar panel are also available here for adjust the thresholding value, the model is not work in the low light condition and without white background the prediction accuracy level is less, in other Tamil sign language recognition there is no real time application are done, this paper overcome the problem of the real time prediction

The dataset of the images are taken from the different person, here we processed the 12 different sign of Tamil sign language the 12 sign are Uyir Ezhuthu, here we processed the image in CPU we compare to GPU it a bit slower.
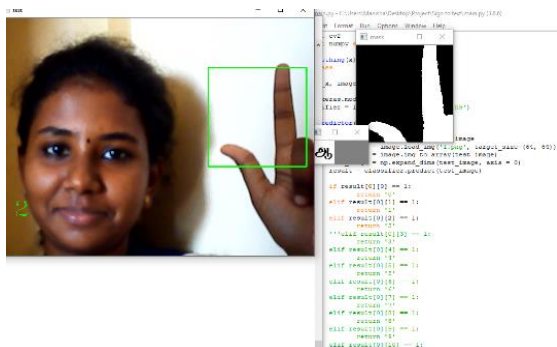

Fig 1.3 output image

5. Conclusion

Tamil sign language recognition in the real time system has developed for letter of Uyir Ezhuthu, Total 4550 image in training set and 3250 images in the testing set of 12 letters each letter have the sign set of 350 image in training set and 250 image in the testing set the model is is processed by using convolutional accuracy us reduced in low light condition. In the future we include the Tamil words as symbol of Tamil sign language and the background and low light condition problem has been solved.

LETTERS	ACCURACY	LOSS
அ	1.00	0.001
ஆ	1.00	0.039
இ	1.00	0.018
ஈ	1.00	0.018
உ	0.99	0.025
ஊ	0.99	0.028
எ	1.00	0.039
ஏ	1.00	0.016
ஐ	1.00	0.010
ஒ	1.00	0.030
ஓ	0.99	0.002
ஔ	0.97	0.025

Table 1. Accuracy and Loss of Signs

References

[1] Sajanraj T D, Beena M V, Indian Sign Language Numeral Recognition Using Region of Interest Convolutional Neural Network Proceedings of the 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT 2018)

[2] Madhuri Sharma, Ranjna Pal and Ashok Kumar Sahoo, Indian Sign Language Recognition Using Neural Networks and Knn Classifiers ,ARPN Journal of Engineering and Applied Sciences, 2014

[3] Nantinee Soodtoetong, Eakbodin Gedkhaw, The Efficiency of Sign Language Recognition using 3D Convolutional Neural Networks,International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology,2018

[4] Mohammed Mustafa, A study on Arabic sign language recognition for differently abled using advanced machine learning classifiers, Journal of Ambient Intelligence and Humanized Computing 2019

[5] Surejya Suresh, Mithun Haridas T.P. Supriya M.H,Sign Language Recognition System Using Deep Neural Network, International Conference on Advanced Computing & Communication Systems (ICACCS),2019

[6] Farhad Yasir1, P.W.C. Prasad1, Abeer Alsadoon1, A. Elchouemi2, Sasikumaran Sreedharan3 1Charles Sturt University Study Centre, Sydney Bangla Sign Language Recognition using Convolutional Neural Network, International Conference on Intelligent Computing, Instrumentation and Control Technologies ,2017

[7] Hana Hosoe, and Shinji Sako, Bogdan Kwolek, Recognition of JSL Finger Spelling Using Convolutional

[8] Neural Networks, Fifteenth IAPR International Conference on Machine Vision Applications (MVA), 2017

[9] Brandon Garcia, Sigberto Alarcon Viesca Real-time American Sign Language Recognition with Convolutional Neural Networks,

[10]   Kshitij Bantupalli, Ying Xie, American Sign Language Recognition using Deep Learning and Computer Vision, IEEE International Conference on Big Data,2018

[11]   Hancy Jose and Anitha Julian Tamil Sign Language Translator—An Assistive System for Hearing- and Speech-Impaired People .

[12]   Su Yang, Qing Zhu, Video-Based Chinese Sign Language Recognition Using Convolutional Neural Network, 9th IEEE International Conference on Communication Software and Networks, 2017

[13]   Korean Sign Language Recognition Based on Image and Convolution Neural Network, Hyojoo Shin, Woo Je Kim, Kyoung-ae Jang

[14]   Mrs.NeelaHarish, Dr.S.Poonguzhali, Design and development of hand gesture recognition system for speech impaired people2015 International Conference on Industrial Instrumentation and Control (ICIC) College of Engineering Pune, India, 2015.

[15]   Manisha. A, Neela Harish,Study of Sign language Recognition Techniques, journal of huazhong university of science and technology,2021