IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

TRANSLATION OF DRAVIDIAN SIGN LANGUAGE (TAMIL) TO EQUIVALENT TEXT USING CNN

[1]Suvetha S,[2]Arockia Sachin A,[3]Aiswarya G

[1,2,3] BE Student

[1,2,3] Department of Computer Science and Engineering

[1,2,3] Dr.Mahalingam College of Engineering and Technology, Pollachi, India

Abstract: Communication between the Speech-impaired and normal people has always been difficult. It is a challenge to develop an appropriate technology to help these people communicate with others. Recognizing the sign language of deaf and dumb people is the only way that offers speech-impaired a proper platform to easily communicate with the normal people. According to Census 2011, out of 6.79 crore people 21,932 people are speech-impaired in Tamil Nadu who speak Tamil Language. It is a tedious process to make all the normal people learn the sign language, but whereas it is an optimal process for a machine to do this job. There are numerous applications developed for recognizing American Sign Language (ASL), but few efficient applications for recognizing Dravidian Sign Language. So, the goal is to develop a gesture recognition system for Dravidian Sign Language (Tamil) that efficiently translates the sign language into an equivalent Tamil text. This project bridges the communication gap between the speech-impaired and normal people with the help of Human-Computer Interaction (HCI) by using the CNN (Convolutional Neural Networks) technique.

Index Terms **- Convolutional Neural Networks (CNN), Tamil Sign Language (TSL), Region of Interest (ROI).**

I. INTRODUCTION

Hand gesture is one of the methods used in sign language for non-verbal communication. It is most commonly used by deaf & dumb people who have hearing or speech problems to communicate among themselves or with normal people. Various sign language systems have been developed by many around the world but they are neither flexible nor cost-effective for the end users. Hence a software is introduced which presents a system prototype that is able to automatically recognize sign language to help deaf and dumb people to communicate more effectively with normal people. Pattern recognition and Gesture recognition are the developing fields of research. Being a significant part in nonverbal communication, hand gestures are playing key role in our daily life. Hand Gesture recognition system provides us an innovative, natural, user friendly way of communication with the computer which is more familiar to the human beings. By considering in mind the similarities of human hand shape with four fingers and one thumb, the software aims to present a real time system for recognition of hand gesture on basis of detection of some shape based features like orientation, Centre of mass, centroid, fingers status, and thumb in positions of raised or folded fingers of hand.

Tamil Sign Language (TSL) is possibly the prevalent Sign Language in Tamil Nadu used by most of the mute people. It has its own phonetics, grammar and syntax which differ from other sign languages. Research related to TSL is now only being standardized. Considering the challenges in TSL recognition, a new method for recognition of static gestures of TSL has been proposed in this work. Tamil Sign Language (TSL) Interpretation system is away to help the mute people to interact with normal people with the help of computer. If a system developed is strong enough for processing the static gestures then it would be the finest system to process the frames obtained while processing the continuous gestures. It is shown that this system performs well with 95% accuracy. Depth images are useful for making highly precise predictions on what is being communicated by the user. Early works used different kind of image convolution to form feature vectors based on a single RGB image of a hand. The authors use wavelet families, computed on edge images, as features to train a Neural Network for 24 sign classification. Haarlet-like features computed on grey-scale images and on silhouettes were used for classification of 10 hand shapes. Principle Component Analysis (PCA) was applied directly on images to derive a sub space of hand poses, which is then used to classify the hand poses. A modification of HOG descriptors is employed to recognize static signs of the British Sign Language. SIFT-feature based description was used to recognize signs of ASL. All these methods depend heavily on the lighting conditions, subject's appearance and Background. This project uses Convolutional Neural Networks for classifying gestures using depth data and its features for more accuracy.

II. LITERATURE REVIEW

C.Anushuya, S.Md MansoorRoomi and M. Senthilarasi [1] have used Sparse Representation Classifier to provide 71% accuracy on a maximum scale and 44% accuracy on average but only in uniform background.SRC has been proved to have better performance than the nearest neighbor and nearest subspace. In SRC, the dictionary is constructed from the training samples of 31 hand signs.

Lakshman Karthik Ramkumar, Sudharasana Premchand and Gokul Karthi Vijayakumar [2] used Simple Recurrent Networks (SRNs) which are known to be a useful tool in cognitive modeling of sequence learning. One advantage of SRNs over feed-forward networks is their ability to implicitly learn the temporal characteristics of a given sequence.Ankit Chaudhary and J.L.Raheja [3] improves the accuracy of the system by using a light invariant gesture recognition system using the ANN technique. Keskin have developed an automatic vision-based tutor for sign language. Some common features extracted include hand silhouettes, contours, key points distributed along the hand i.e. fingertips and joints. Grzeszcuk used stereo based 3D modeling for hand gesture recognition with six types of gestures. He used arm orientation for hand position detection and used color based image segmentation.

Jayan Mistry and Benjamin Inden [4] used Intel RealSense camera instead of a web-cam that improves accuracy and speed but does not support skeleton tracking.Intel RealSense Technology is a suite of depth and tracking technologies designed to give machines and devices depth perceptions capabilities that will enable them to "see" and understand the world. There are many uses for these computer vision capabilities including autonomous drones, robots, AR/VR, smart home devices amongst many others broad market products.

Aashni Haria, ArchanaShri Subramanian, Nivedhitha Asokkumar, Shristi Poddar and Jyothi S Nayak [5] recognized both static and dynamic gestures and converted into respective actions like opening VLC player and PowerPoint for the particular gestures.Noise in images can be defined as a random variation of brightness or color information that is usually produced during the image acquisition process from the webcam. This noise is an undesirable aspect of the image and needs to be removed. In order to do this, Gaussian filter is applied. Gaussian filtering is performed by the convolution of Gaussian kernel with each point in the input array.

Joyeeta Singha and Karen Das [6] detected the gestures using Eigen value Weighted Euclidean Distance based Classification Technique. Eigen values and Eigen vectors are a part of linear transformations. Eigen vectors are the directions along which the linear transformation, acts by stretching, compressing or flipping and Eigen values gives the factor by which the compression or stretching occurs. Using Eigen value weighted Euclidean distance as a classification technique, their project has achieved 97% recognition rate.

S.Sudha, S.Jothilakshm and, R.Rajasoundaramani [7] implemented various classification techniques like K-Nearest Neighbour (KNN), Proximal Support Vector machine (PSVM) and Naïve Bayesian. In comparison to all three, Naïve Bayesian produces the maximum efficiency of 93%.Pratibha Pandey, Vinay Jain and Bhilai C.G [8] identified hand gestures using the vision based approach which includes 2D and 3D model, instrumental glove and based on Colored marker.Vision based approach, Glove based approach and Color Marker based approach are some of the common techniques used for this purpose.Recognition Time, Computational complexity, robustness are some of the constraint posed by the system.

III. PROPOSED METHODOLOGY

The communication between the user and the system occurs as follows:

- The user makes the gestures in-front of the camera which captures the Depth Image and separates the gestures from the entire scene.
- The important features such as position of hand and fingers are extracted to make decision on what is being gestured.
- The Gesture Recognition is performed using Convolutional Neural Networks by classifying the gestures.
- The sentences are interpreted from the gesture and finally sent to the user interface which displays the letters/words.

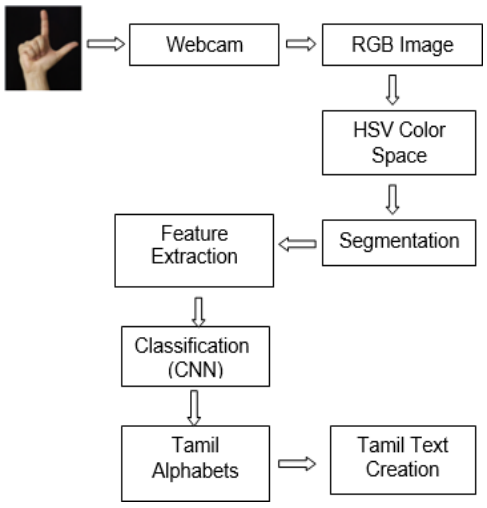


Figure 1 Block Diagram of Proposed System

3.1Preprocessing of Input Data: Conversion of RGB Image to HSV Color Space

The input RGB image is first converted to the HSV image. The motive of performing this step is, RGB image is very sensitive to change in illumination condition. The HSV color space separates three components: Hue which means the set of pure colors within a color space, Saturation describing the grade of purity of a color image and Value giving relative lightness or darkness of a color.

3.2 Feature Extraction by Finding Region of Interest (ROI)

After the desired portion of the image is being cropped, feature extraction phase is carried out. Here, Background Subtraction is used to find out the ROI.

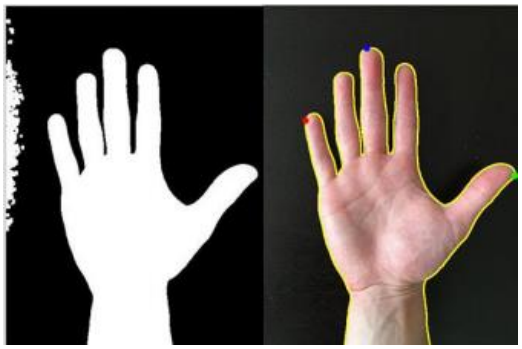


Figure 2 Feature Extraction

3.3 Classification using Convolutional Neural Networks (CNN)

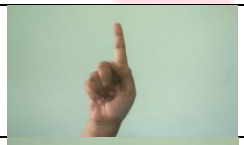
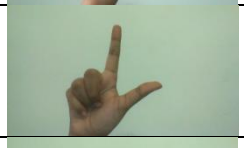
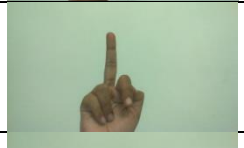
Convolutional Neural Networks (also referred to as CNN or ConvNet) are a class of deep Neural Networks that have seen widespread adoption in a number of computer vision and visual imagery applications. The overall architecture of a CNN consists of an input layer, hidden layer(s), and an output layer. CNN proves to be a faster and effective classifier in terms of speed and accuracy. Almost every gesture with support values greater than 370 is predicted correctly. CNN can be used to predict the gestures. The CNN predicts the gestures accurately under better lighting conditions.

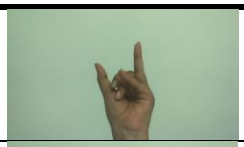
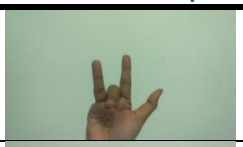
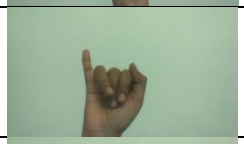
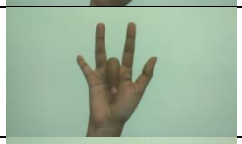
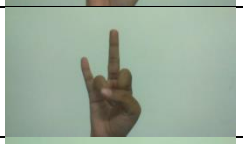
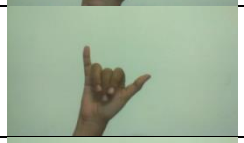
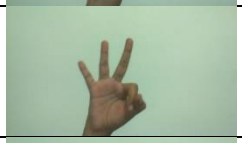
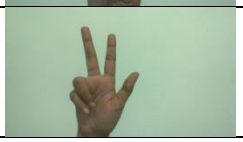
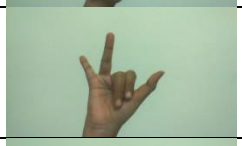
IV. RESULTS AND DISCUSSION

4.1 Dataset

The CNN is tested with the dataset that has been generated. The TSL has total of 31 different signs with 12 vowels, 18 consonants and 1 Aytham. The dataset comprising 37,200 samples for 31 gestures is used for classification and training. The dataset is generated by our team. The captured image is converted to a monochrome by using threshold function for easy recognition of gesture.

Table 1 Dataset

Letter	Gesture	Letter	Gesture	Letter	Gesture
அ		ஔ		ம்	
ஆ		ஃ		ய்	
இ		க்		ர்	
ஈ		ங்		ல்	
உ		ச்		வ்	

ஊ		ஞ		ழ்	
எ		ட்		ள்	
ஏ		ண்		ற்	
ஐ		த்		ன்	
ஒ		ந்			
ஓ		ப்			

4.2 Evaluation Metrics

In the field of machine learning and specifically the problem of statistical classification, a confusion matrix, also known as an error matrix is used to evaluate the system. A confusion matrix is a table that is often used to describe the performance of a classification model on a set of test data for which the true values are known. It allows the visualization of the performance of an algorithm. It allows easy identification of confusion between classes e.g. one class is commonly mislabelled as the other. Most performance measures are computed from the confusion matrix.

4.3 Experiments & Results

In our experiment, a confusion matrix has been used to analyze the accuracy of each gesture and some of the dataset sample is used for this purpose. A confusion matrix is a summary of prediction results on a classification problem. The number of correct and incorrect predictions are summarized with count values and broken down by each class. This is the key to the confusion matrix. The confusion matrix shows the ways in which your classification model is confused when it makes predictions. It gives us insight not only into the errors being made by a classifier but more importantly the types of errors made.

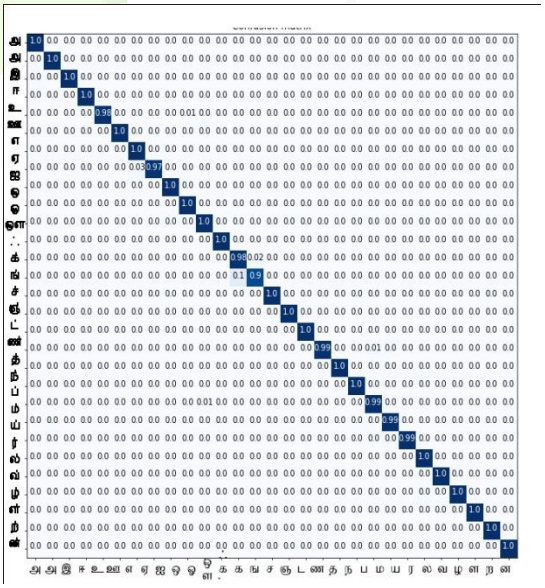


Figure 3 Confusion Matrix

IV. CONCLUSION

The test accuracies for the model obtained is 95%, but the real time accuracy is lesser because the data set needs more manipulation and augmentation. The gestures of vowels letters are predicted accurately whereas in consonants, the gestures of the letters "ட்", "த்" and "ர்", "வ்" are slightly similar, so the accuracy of these letters are difficult to predict. CNN is used because it proves to be a faster and effective classifier in terms of speed and accuracy.

REFERENCES

[1] C.Anushuya, S.Md MansoorRoomi, M. Senthilarasi, "A Signature Based Dravidian Sign Language Recognition By Sparse Representation", International Journal on Natural Language (IJNLC), Vol.5, No.1, February 2016.

[2] C.Anushuya, S.Md MansoorRoomi, M. Senthilarasi, "A Signature Based Dravidian Sign Language Recognition By Sparse Representation", International Journal on Natural Language (IJNLC), Vol.5, No.1, February 2016.

[3] Ankit Chaudhary,J.L.Raheja,"Light Invariant Real-Time Robust Hand Gesture Recognition", Science Direct, 2018.

[4] Jayan Mistry, Benjamin Inden,"An Approach To Sign Language Translation Using The Intel RealSence Camera", International Journal Computer Science and Electronic Engineering (IEEE), Vol.5, No.1, February 2018.

[5] Aashni Haria, Archana Shri Subramanian, Nivedhitha Asokkumar, Shristi Poddar, Jyothi S Nayak "Hand Gesture Recognition For Human Computer Interaction," International Conference on Advances in Computing and Communications, ICACC-2017,22-24 August 2017.

[6] Joyeeta Singha, Karen Das, "Indian Sign Language Recognition using Eigenvalue Weighted Euclidean Distance based Classification Technique" International Journal of Advanced Computer Science and Applications (IJACSA),Vol.4,No.2, 2013.

[7] S.Sudha, S.Jothilakshmi, R.Rajasoundaramani "Tamil Sign Language to Speech Translation" International Journal of Science and Research (IJSR), 2012.

[8] PratibhaPandey, Vinay Jain, Bhilai C.G "Hand Gesture Recognition for Sign Language Recognition", International Journal of Science, Engineering and Technology Research (IJSETR), Vol.4,No.3, March 2015.

[9] Ashok K Sahoo, Gouri Sankar Mishra and Kiran Kumar Ravulakollu, "Sign Language Recognition: State Of The Art ", ARPN Journal of Engineering and Applied Sciences, Vol. 9, No. 2, February 2014.

[10] S. Sudha, S. Jothilakshmi, "Dynamic Tamil Sign Language Recognition System", International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE), Vol. 2, Issue 8, August 2016.

[11] Vivek Bheda, N. Dianna Radpour, "Using Deep Convolutional Network for Gesture Recognition in American Sign Language",2017.

[12] M. Krishnaveni, P. Subashini, T.T. Dhivyaprabha, "Improved Canny Edges Using Cellular Based Particle Swarm Optimization Technique for Tamil Sign Digital Images", International Journal of Electrical and Computer Engineering (IJECE), Vol. 6, No. 5, October 2016.

[13] P. Rajathi, S. Jothilakshmi,"A Static Tamil Sign Language Recognition System", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 4, April 2013.

[14] P. Subha Rajam, G. Balakrishnan, "Recognition of Tamil Sign Language Alphabet using Image Processing to aid Deaf-Dumb People", International Conference on Communication Technology and System Design, Research Gate, December 2012.

[15] Archana S. Ghotkar and Dr. Gajanan K. Kharate, "Study of Vision based Hand Gesture Recognition using Indian Sign Language", International Journal On Smart Sensing And Intelligent Systems, Vol. 7, No. 1, March 2014.