IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Learning And Testing Fingerspelled Signs Using Deep Learning

Mathews Emmanuel, Nimmy Chacko, Pooja Raj M

Assistant Professor,Assistant Professor, MCA Student
Department of Computer Application
Cochin University College of Engineering Kuttanad (CUCEK) , Alapuzha , India.
Department of Basic science Amaljyothi Engineering College Koovappally,Kanjirappally,Kottayam,India
Department of Computer Application
Cochin University College of Engineering Kuttanad (CUCEK) , Kottayam , India

***Abstract**:* Sign Language is one among the media of communication for deaf people. One should learn signs to interact with them. Learning usually takes place in peer groups. There exist very few study materials for sign learning. Because of this, the method of learning signing learning may be a difficult task. Fingerspelled sign learning is the initial stage of sign learning and moreover, are used when no corresponding sign exists or signer is not aware of it. Most of the prevailing tools for sign learning use external sensors which are costly. This paper discusses SignQuiz, which can be an economical web-based fingerspelled sign learning application for Indian signing (ISL) utilizing automatic signing recognition technique. It works in two modes, learning and testing. In learning mode, signs are listed and one can learn the signs by clicking on the required ones. In testing mode, the user is tested for the learned signs. SignQuiz helps to find out signs with none external help. This is the first attempt in ISL for learning finger spelled signs using a deep neural network. The results indicate that SignQuiz is best than the printed medium for fingerspelled sign learning.

Index Terms* - *Assistive technology, sign language, Convolutional Neural Network

I. Introduction

Sign languages provide how to speak without the utilization of voice. This is very import for individuals who have impaired hearing and inability speaking because it provides them an alternative means of communicating with other people. Furthermore, signing may be a universal language which will allow people of a special speaking language to speak which is an extremely important in multicultural settings and times of disaster. Over the past few decades, many efforts are made in creating a sign language recognition (SLR) system. There are two main categories in SLR, namely isolated signing recognition and continuous sign classifications.

There are around 466 million people worldwide with deafness and 34 million of those are children. In developing countries there are only a few schools for deaf students. Unemployment rate among adults with hearing loss is a very high in developing countries. Sign Learning is very difficult for a beginner without the help of trained sign language practitioner. Learning through books are not effective as it is not easy to represent signs in a book using pictures. Though technology based tools exist for sign language learning, they are doing not provide any feedback on signs produced by the user. This makes it difficult to learn signs without any external help. The proposed study aims to develop a system which can recognize static sign gestures, and convert them into corresponding words. A vision-based approach using a web camera is introduced to obtain the data from the signer and can be used offline. The purpose of creating the system is that it will serve as the learning tool for those who want to know more about the basics of sign language such as alphabets, numbers, and common static signs. The proponents provided a white background, and a specific location for image processing of the hand, thus, improving the accuracy of the system and used Convolutional Neural Network (CNN) as the recognizer of the system. The scope of the study includes basic static signs, numbers and ISL alphabets (A–Z). One of the main features of this study is the ability of the system to create words by fingerspelling without the use of sensors and other external technologies. Fig. 1 presents the ISL alphabets that will be fed onto the system.

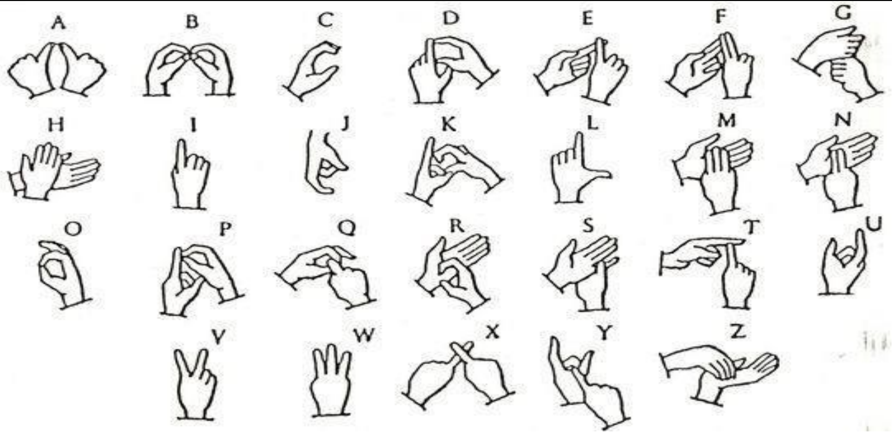

Fig. 1: ISL alphabet signs.

Fig. 2 presents a static gesture for each number provided. The system will be limited with numbers 0–9. Fig. 2. ISL number signs.

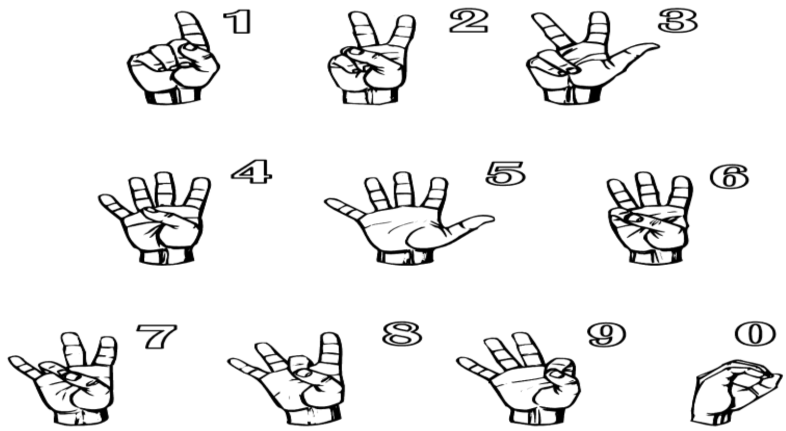

Fig. 2: ISL number signs

II. LITERATURE REVIEW

Different approaches are proposed in trying to unravel the problem of recognizing sign language hand gestures. Some of the first works include [1] which used SVM to classify South African signing (SASL) and [2] which used parallel hidden markov models. In [3], Sole et al. used Extreme Learning Machine (ELM) to find out to classify static hand gestures on the letters of the Auslan dictionary. Although the research showed good preliminary results, the info wont to test the network was restricted to pictures collected from an equivalent day, thus lacking generalization. In [6], Kim et al. demonstrated how to utilize a deep neural network (DNN) to classify frames in image sequences of finger spelled letters. Similar to [6], Kim et al. [7] also used DNN with HoG features. Our work was also able to create a signer-independent classification of images but in contrast to those works our classifier is totally deep neural network including the feature extraction. The work of Pugeault et al. [8] presented the use of depth images from a Microsoft Kinect device. The used multiclass random forest classification and tested their method by varying the input from image only, depth only and combined image with depth. They achieved their best result when depth is combined with image. Their system is additionally fast enough for real-time classification. In [9], Kang et al. also used depth images as its input without the colour image. They also used a DNN for his or her classifier and was also ready to achieve real-time rates. Similar to them, our goal also is to achieve real-time rates but using only color image since only few people have access to depth sensors. In recent years convolutional neural networks had great success in computer vision tasks including image recognition. Starting with Alex Net [10] which popularized deep convolutional neural networks by winning the Image Net Challenge: ILSVRC 2012 [11]. Following Alex Net, new network architectures like [12], [13] and [14] also made significant contributions and improved performance.

III. IMPLEMENTATION

3.1 DATASET

Fingerspelled Indian Sign Language (ISL) signs were captured for training the model used for sign recognition. Capturing was done through mobile cameras, laptop camera and Digital SLR's. Signs corresponding to 26 fingerspelled alphabets were captured. This was collected with the help of signers. Signers comprised students and faculty from Cochin University Science And Technology (CUSAT), Kerala, India. They used signs released by Indian Sign Language Research and Training Center (ISLRTC) as a reference. The validity of the captured signs were confirmed by various sign language practitioners consisting of sign language interpreters, teachers and deaf people. Close to 1000 images were collected for each sign making the entire number of images collected to about 20 1000. Among the captured signs, certain alphabets like "M" and "N" are double handed and certain others like "C" are single handed.

3.2 PROPOSED MODEL

SignQuiz is designed as a web based application that helps to learn signs without any external help. It is designed to work from any web browser so that users can access it without installing any new application. It works in two modes, learning and testing. In learning mode, signs are listed and one can learn the signs by clicking on the required ones. In testing mode, the user is tested for the learned signs. It is designed as a quiz application. User is asked to show a sign and system automatically detects the sign and gives feedback.

From the SignQuiz home page, user can select either learning or testing feature. In the training screen, alphabets from A-Z letters and numbers are listed. In the learning screen user can show signs corresponding to given alphabets for learning. Learning screen shows sign corresponding to English alphabets, current score and finished alphabet list. User can click on the capture button provided and can show the sign. A time delay is provided for sign capturing. After the sign is captured, it is send to the server from the browser. Image is captured within a two second delay to adjust for the lack of experience of the user in showing the sign. Server captures these images and finds out the alphabet corresponding to the sign. This is the output of the softmax classifier. Sign is accepted only if the accuracy is greater than 85%. Otherwise it is treated as an error For simplicity, both learning and training screens are designed as a single screen. Image capturing and score update are done dynamically. This is made possible through client side scripting languages JavaScript and Ajax. Recognition result, score and other details are send back from server to browser in Json format. SignQuiz used pre-trained models for sign classification. Two pretrained models – Nasnet and InceptionV3 were considered.

Learning screen of SignQuiz where user can select alphabets. For each alphabet, images from front and back and video is provided. User can pause and replay the videos. Sign images from front and back side and video will help to clearly understand how a sign is produced. The training screen where user is asked to show signs. Sign to show, current score and completed signs are shown in a box. This gives information about progress.

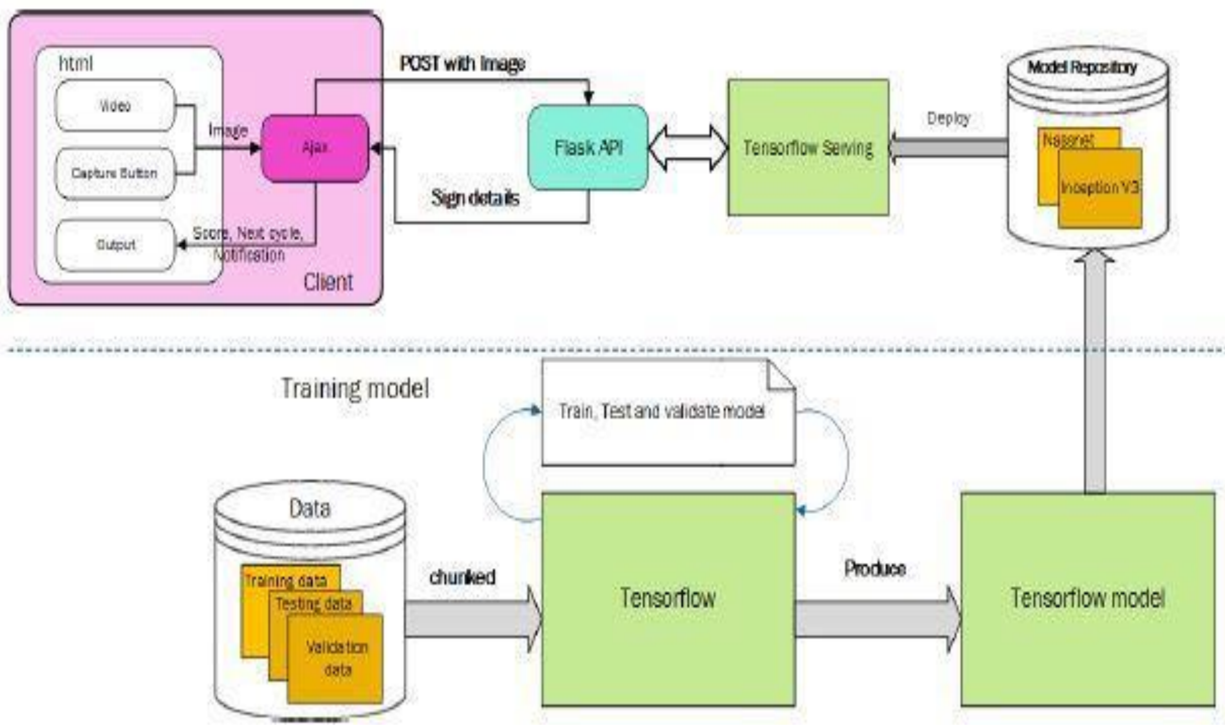

Fig.3: Architecture

3.3 TECHNOLOGY

Convolutional Neural Networks have been extremely successful in image recognition and classification problems, and have been successfully implemented for human gesture recognition in recent years. In particular, there has been work wiped out the realm of signing recognition using deep CNNs, with input-recognition that is sensitive to quite just pixels of the pictures . With the use of cameras that sense depth and contour, the method is made much easier via developing characteristic depth and motion profiles for each sign language gesture [4].The use of depth-sensing technology is quickly growing in popularity, and other tools have been incorporated into the process that have proven successful. Developments such as custom-designed color gloves have been wont to facilitate the popularity process and make the feature extraction step more efficient by making certain gestural units easier to identify and classify [5].

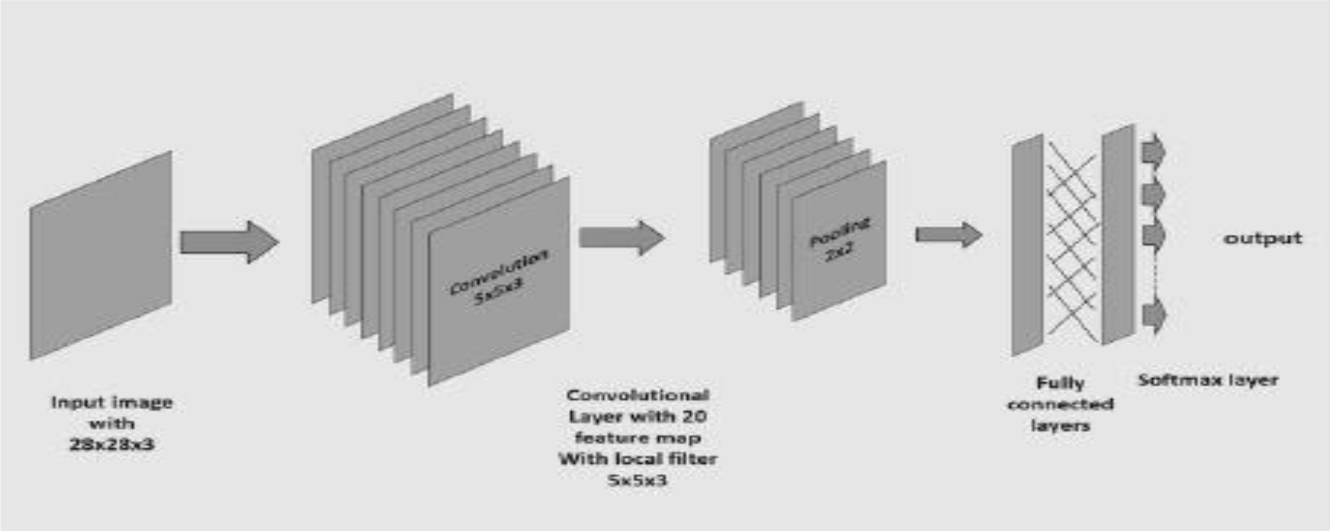

Fig. 4: Convolution with single layer

IV. LIMITATIONS

SignQuiz tests were conducted on those people who volunteered to take part in it. Most of them are supporters of technology based applications. A bias towards technology based applications is obvious. SignQuiz currently included only finger spelled signs for learning. A full-edged sign learning application needs to include more signs. For this purpose more study should be conducted to seek out the effectiveness of gesture detection from videos. More detailed study must be done to seek out the behavior of SignQuiz in signer independent setting. Currently signer independent data is restricted to collecting signs from 14 users utilized for setting the threshold limit. SignQuiz is compared with printed ISL learning materials only. This is due to the absence of any other medium than peer learning for ISL learning

V. CONCLUSION AND FUTURE SCOPE

The SignQuiz, a web-based application for learning to sign making use of Deep Neural Networks (DNN). SignQuiz application can easily be used by both deaf and non-deaf people. Ease of use, availability, low cost of operation are the features that make SignQuiz a useful application for learning fingerspelled signs. By changing the model used, it can support any sign language. With proper training this application can easily include more signs. Usability can be improved if user can select alphabet range of his choice for learning. Getting each user a user account will help to prevent and begin as he wish. This will also help to understand easy or difficult signs based on the global data. Rather than setting sign classification accuracy threshold globally, it is often set for every sign for a better working. More detailed study should be done to set this. To make SignQuiz capture the sign made by the user with none external help, application is meant so on await a couple of seconds after user clicks on the capture button. This will create confusion in a novice user. Rather than putting the delay, showing a timer or automatically understanding that user has shown the sign and capturing it'll be helpful. We feel that results obtained from this study will help to design applications which are helpful in learning sign language.

REFERENCES

[1] S. Naidoo, C. Omlin, and M. Glaser, "Vision-based static hand gesture recognition using support vector machines," University of Western Cape, Bellville, 1998.
[2] C. Vogler and D. Metaxas, "Parallel hidden markov models for american sign language recognition," in Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, vol.1. IEEE, 1999, pp. 116–122.
[3] M. M. Sole and M. Tsoeu, "Sign language recognition using the extreme learning machine," in AFRICON, 2011.IEEE, 2011, pp.1–6.
[4] Agarwal, Anant & Thakur, Manish. Sign Language Recognition using Microsoft Kinect. In *IEEE International Conference on Contemporary Computing* , 2013.
[5] Cao Dong, Ming C. Leu and Zhaozheng Yin. American Sign Language Alphabet Recognition Using Microsoft Kinect. In *IEEE International Conference on Computer Vision and Pattern Recognition Workshops* ,2015.
[6] T. Kim, W. Wang, H. Tang, and K. Livescu, "Signer-independent fingerspelling recognition with deep neural network adaptation," in Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on. IEEE, 2016, pp. 6160–6164..
[7] T. Kim, J. Keane, W. Wang, H. Tang, J. Riggle, G. Shakhnarovich, D. Brentari, and K. Livescu, "Lexicon-free fingerspelling recognition from video: Data, models, and signer adaptation," Computer Speech & Language, vol. 46, pp. 209–232, 2017.
[8] N. Pugeault and R. Bowden, "Spelling it out: Real-time asl fingerspelling recognition," in Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on. IEEE, 2011, pp.1114–1119.
[9] B. Kang, S. Tripathi, and T. Q. Nguyen, "Real-time sign language fingerspelling recognition using convolutional neural networks from depth map," in Pattern Recognition (ACPR), 2015 3rd IAPR Asian Conference on. IEEE, 2015, pp. 136–140.
[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet c classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097–1105.
[11] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein et al., "Imagenet large scale visual recognition challenge," International Journal of Computer Vision, vol. 115, no. 3, pp. 211–252, 2015.
[12] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.
[13] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2818–2826.
[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer