International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.com

IJIEMR Transactions, online available on 29th Aug 2021.

Link :http://www.ijiemr.com/downloads.php?vol=Volume-10&issue=ISSUE-08

DOI: **10.48047/IJIEMR/V10/I08/19-21**

Title:- **SIGN LANGUAGE RECOGNITION USING MACHINE LEARNING**

Volume 10, Issue 08, Pages:263-267

Paper Authors
Dr. Shameena Begum[1], Gottumukkala Manasa [2], Darbha Sowmya Sarvani [3] , Ganta Sireesha [4] , Gollu Sai Krishna [5]

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.com

SIGN LANGUAGE RECOGNITION USING MACHINE LEARNING

Dr. Shameena Begum[1], Gottumukkala Manasa [2], Darbha Sowmya Sarvani [3] , Ganta Sireesha [4] , Gollu Sai Krishna [5]
[1]Professor, Dept. of CSE, [2]17ME1A0577, [3]17ME1A0573, [4]17ME1A0575, [5]17ME1A0576
Ramachandra College of Engineering, A.P., India

ABSTRACT:

The only way the speech and hearing impaired (i.e dumb and deaf) people can communicate is by sign language. The main problem of this way of communication is normal people who cannot understand sign language can't communicate with these people or vice versa. Our project aims to bridge the gap between the speech and hearing impaired people and the normal people. Sign Language Recognition by technology is an overlooked concept despite there being a large social group which could benefit by it. There are not many technologies which help in connecting this social group to the rest of the world. Understanding sign language is one of the primary enablers in helping users of sign language communicate with the rest of the society. Image classification and machine learning can be used to help computers recognize sign language, which could then be interpreted by other people. Convolutional neural networks have been employed in this paper to recognize sign language gestures.

INTRODUCTION

Sign language, as one of the most widely used communication means for hearing-impaired people, is expressed by variations of hand-shapes, body movement, and even facial expression. Since it is difficult to collaboratively exploit the information from hand-shapes and body movement trajectory, sign language recognition is still a very challenging task. This paper proposes an effective recognition model to translate sign language into text or speech in order to help the hearing impaired communicate with normal people through sign language.

Technically speaking, the main challenge of sign language recognition lies in developing descriptors to express hand-shapes and motion trajectory. In particular, hand-shape description involves tracking hand regions in video stream, segmenting hand-shape images from complex background in each frame and gestures recognition problems. Motion trajectory is also related to tracking of the key points and curve matching. Although lots of research works have been conducted on these two issues for now, it is still hard to obtain satisfying result for SLR due to the variation and occlusion of hands and body joints. Besides, it is a nontrivial issue to integrate the hand-shape features and trajectory features together. To address these difficulties, we develop a CNNs to naturally integrate hand-shapes, trajectory of action and facial expression. Instead of using commonly used color images as input to networks like [1, 2], we take color images, depth images and body skeleton images simultaneously as input which are all provided by Microsoft Kinect.

Kinect is a motion sensor which can provide color stream and depth stream. With the public Windows SDK, the body joint locations can be obtained in real-time as shown in Fig.1. Therefore, we choose Kinect as capture device to record sign

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.com

words dataset. The change of color and depth in pixel level are useful information to discriminate different sign actions. And the variation of body joints in time dimension can depict the trajectory of sign actions. Using multiple types of visual sources as input leads CNNs paying attention to the change not only in color, but also in depth and trajectory. It is worth mentioning that we can avoid the difficulty of tracking hands, segmenting hands from background and designing descriptors for hands.

CNNs have the capabilityto learn features automatically from raw data without any prior knowledge.

2. RELATED WORK

Existing System

The researches done in this field are mostly done using a glove based system. In the glove based system, sensors such as potentiometer, accelerometers etc. are attached to each of the finger. Based on their readings the corresponding alphabet is displayed. The main problem faced by this gloved based system is that it has to be recalibrate every time whenever a new user on the finger- tips so that the fingers tips are identified by the Image Processing unit. We are implementing our project by using Image Processing.

Proposed System

In this project we would present a robust and efficient method of sign language detection. Instead of using Data gloves for sign language detection, we would be doing the detection by image processing. The main advantage of using image processing over Data gloves is that the system is not required to be re-calibrated if a new user is using the system. SLR takes input as hand gesture and displays it as text by CNN.

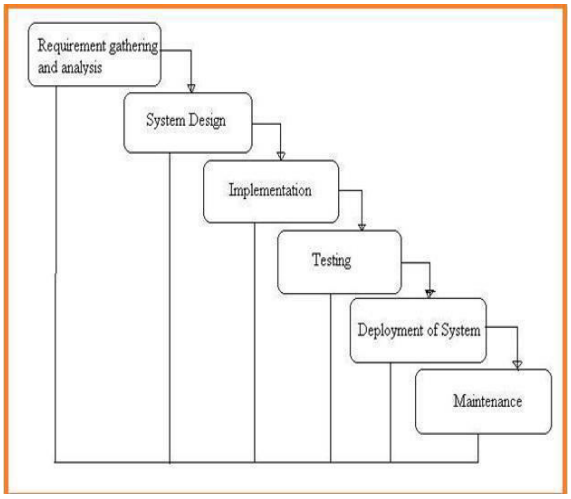


Figure 1.Project SDLC

3. METHODOLOGY:

A diagram is a diagram based on the UML (Unified Modeling Language) with the purpose of visually representing a system along with its main actors, roles, actions, artifacts or classes, in order to better understand, alter, maintain, or document information about the system. As every diagram need not to be included in our project, we tested out what are the best suited diagrams for our project.

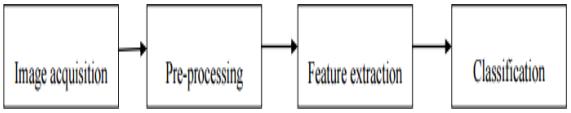


Figure 2.Arcitechure

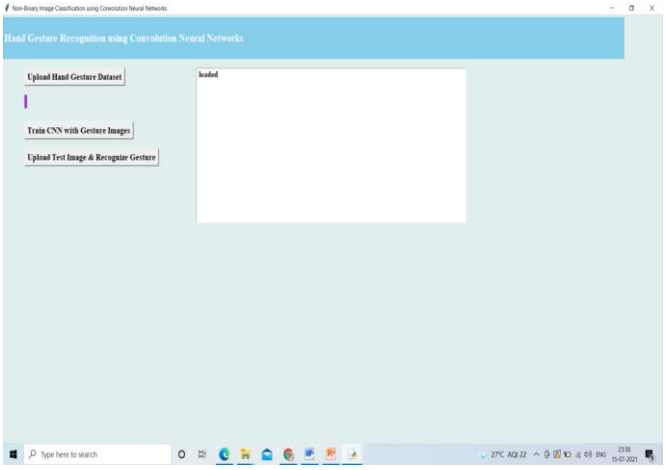
A usage case outline inside the Unified Modeling Language we used in our Project Development is Star (UML) could be a sort of behavioral chart portrayed out by and produced using a Use-case examination. Its inspiration is to gift a graphical layout of the presence of mind gave by a system to the extent performing specialists, their targets (addressed as use

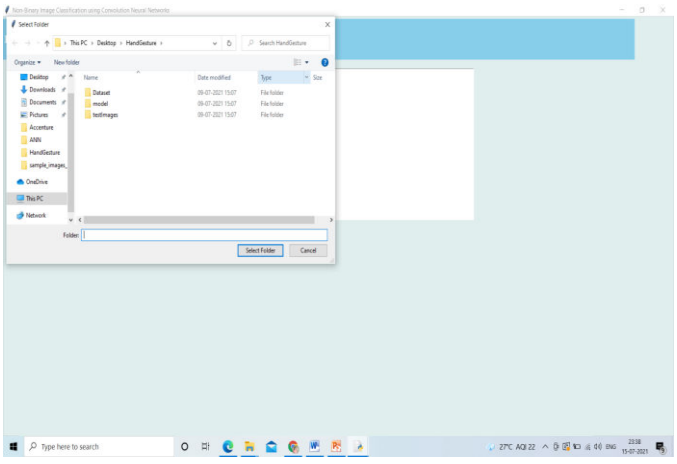
International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.com

cases), and any conditions between those use cases. The most explanation behind a use case diagram is to show what structure limits are played out that on-screen character. Parts of the entertainers inside the systemIn PC code planning, a class plot inside the Unified Modeling Language we used in or Project Development in star (UML) could be a sort of static structure outline the delineates the structure of a system by showing the systems' characterizations, their qualities, operations (or methodologies), and moreover the associations among the groupings. It elucidates that class contains data.It depicts the objects involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario. Sequence diagrams are typically associated with use case realizations in the logical view of the system under development.Unified Modeling Language (UML). These diagrams can be used to portray the dynamic behavior of a particular use case and define the role of each object.

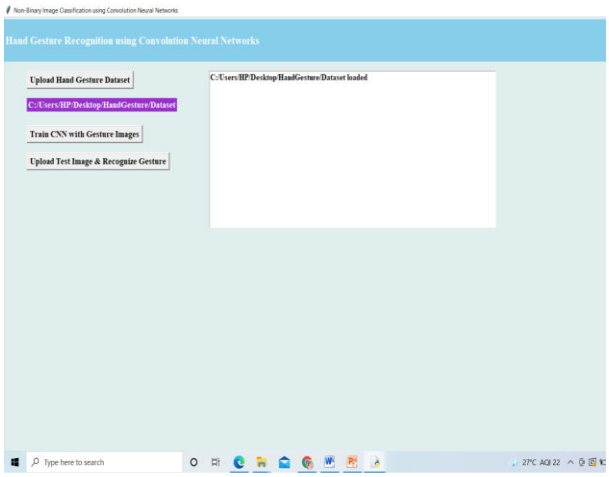
Neural networks, as its name suggests, is a machine learning technique which is modeled after the brain structure. It comprises of a network of learning units called neurons. These neurons learn how to convert input signals (e.g. picture of a cat) into corresponding output signals (e.g. the label "cat"), forming the basis of automated recognition. A convolutional neural network (CNN, or Conv Net) is a type of feed forward artificial neural network in which the connectivity pattern between its neurons is inspired by the organization of the animal visual cortex. CNNs have repetitive blocks of neurons that are applied across space (for images) or time (for audio signals etc). For images, these blocks of neuron scan be interpreted as 2D convolutional kernels, repeatedly applied over each patch of the image. For speech, they can be seen as the 1D convolutional kernels applied across time windows. At training time, the weights for these repeated blocks are 'shared', i.e. the weight gradients learned over various image patches are averaged.

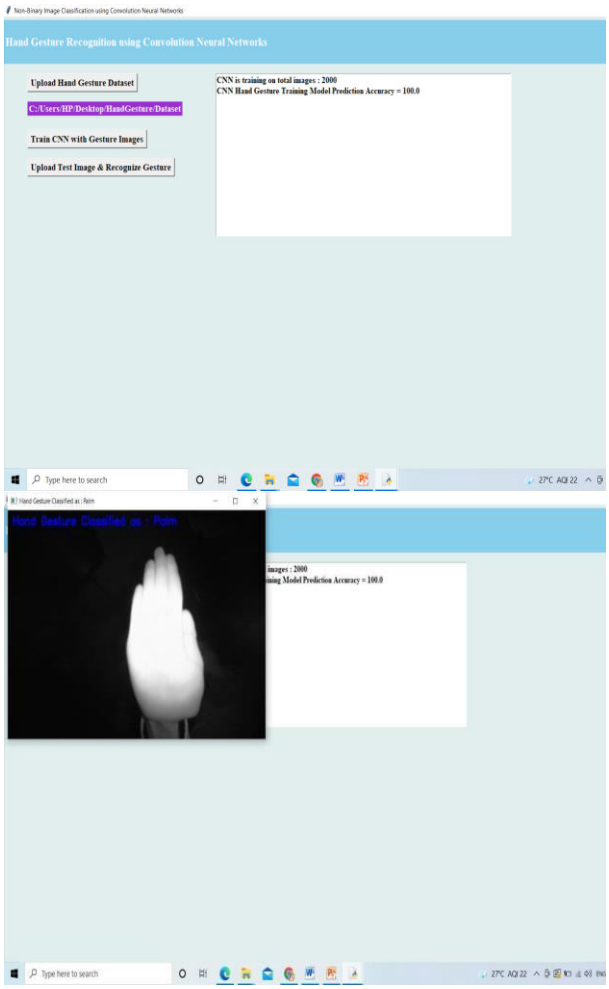
4. STUDY OF RESULTS:





International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.com





5.CONCLUSION :

We developed a CNN model for sign language recognition. Our model learns and extracts both spatial and temporal features by performing 3D convolutions. The developed deep architecture extracts multiple types of information from adjacent input frames and then performs convolution and subsampling separately. The final feature representation combines information from all channels. We use multilayer perceptron classifier to classify these feature representations. For comparison, we evaluate both CNN and GMM-HMM on the same dataset. The experimental results demonstrate the effectiveness of the proposed method.

In future work, proposed system can be developed and implemented using Raspberry Pi. Image Processing part should be improved so that System would be able to communicate in both directions i.e.it should be capable of converting normal language to sign language and vice versa. We will try to recognize signs which include motion. Moreover we will focus on converting the sequence of gestures into text i.e. word and sentences and then converting into the speech it can be heard.

6.REFERENCES :

[1] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097–1105.

[2] Andrej Karpathy, George Toderici, Sanketh Shetty, Thomas Leung, Rahul Sukthankar, and Li Fei-Fei, "Large-scale video classification with convolutional neural networks," in CVPR, 2014.

[3] Yann LeCun, Leon Bottou, YoshuaBengio, and Patrick ´ Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278– 2324, 1998.

[4] HueihanJhuang, Thomas Serre, Lior Wolf, and Tomaso Poggio, "A biologically inspired system for action recognition," in Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on. Ieee, 2007, pp. 1–8.

[5] Shuiwang Ji, Wei Xu, Ming Yang, and Kai Yu, "3D convolutional neural networks for human action recognition," IEEE TPAMI, vol. 35, no. 1, pp. 221–231, 2013.

[6] KirstiGrobel and Marcell Assan, "Isolated sign language recognition using hidden markov models," in Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation., 1997 IEEE International Conference on. IEEE, 1997, vol. 1, pp. 162–167.