INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

ISSN: 2057-5688

# CHIRONOMY TRANSCRIPTION

[1]A. DIVYA SREE, [2]T. SHIRISHA, [3]G. LAHIRI, [4]Y. MANEESHA

[1]Assistant Professor, Dept.of CSE, Teegala Krishna Reddy Engineering College, Meerpet, Hyderabad,

[2,3,4]BTech Student, Dept.of CSE, Teegala Krishna Reddy Engineering College, Meerpet, Hyderabad

shirisha7561@gmail.com, lahirigonela351@gmail.com, 5c7mani@gmail.com

***Abstract***: *Gesture recognition is one of the most popular techniques in the field of computer vision today. In recent years, many algorithms for gesture recognition have been proposed, but most of them do not have a good balance between recognition efficiency and accuracy. In order to offer new possibilities to interact with machine and to design more natural and more intuitive interactions with computing machines, our research aims at the automatic interpretation of gestures based on computer vision. In this paper, we propose a technique which commands computer using six static and eight dynamic hand gestures. The three main steps are: hand shape recognition, tracing of detected hand (if dynamic), and converting the data into the required command. Experiments show 93.09% accuracy.*

***Keywords***: *computer vision, deep learning, hand gesture, neural network, transfer learning, hand gesture recognition*

## I. INTRODUCTION

Gesture recognition is the mathematical interpretation of a human motion by a computing device. Modern research of the control of computers changes from standard peripheral devices to remotely commanding computers through speech, emotions and body gestures [1]. Our application belongs to the domain of hand gesture recognition which is generally divided into two categories i.e., contact-based and vision-based approaches. The second type is simpler and intuitive as it employs video image processing and pattern recognition. The aim is to recognize six static and eight dynamic gestures while maintaining accuracy and speed of the system. The

ISSN: 2057-5688

recognized gestures are to command the computer

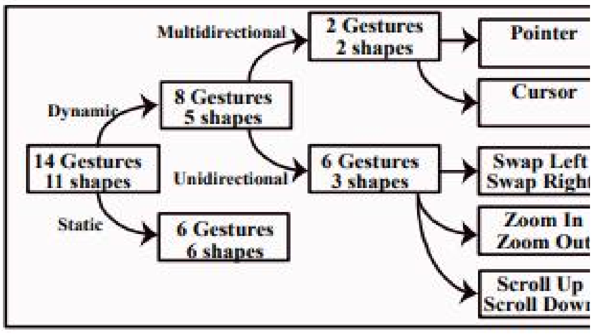



Fig.1 Division of eleven shapes into fourteen gestures

For hand shape recognition, a CNN based classifier is trained through the process of transfer learning over a pretrained convolutional neural net which is initially trained on a large dataset. We are using VGG16 as the pretrained model

## II. LITERATURE SURVEY

Literature review of our proposed system shows that there have been many explorations done to tackle the sign recognition in videos and images using several methods and algorithms. SIMING, He proposed a system having a dataset of 40 common words and 10,000 sign language images. To locate the hand regions in the video frame, Faster R-CNN with an embedded RPN module is used. It improves performance in terms of accuracy. Detection and template classification can be done at a higher speed as compared to single stage target detection algorithm such as YOLO.

The detection accuracy of Faster R-CNN in the paper increases from 89.0% to 91.7% as compared to Fast-RCNN. A 3D CNN is used for feature extraction and a sign-language recognition framework consisting of long- and short-time memory (LSTM) coding and decoding network are built for the language image sequences. On the problem of RGB sign language image or video recognition in practical problems, the paper merges the hand locating network, 3D CNN feature extraction network and LSTM encoding and decoding to construct the algorithm for extraction.

This paper has achieved a recognition of 99% in common vocabulary dataset Let's approach the research done by Rekha, J. which made use of YCbCr skin model to detect and fragment the skin region of the hand gestures. Using P principal Curvature based Region Detector, the image features are extracted and classified with Multi class SVM, DTW and non-linear KNN. A dataset of 23 Indian Sign Language static alphabet signs were used for training and 25 videos for testing. The experimental result obtained were 94.4%for static and 86.4% for dynamic. In, a low-cost approach has been used for image processing. The capture of images was done with a green background so that during processing, the green color can be easily subtracted from the RGB Color Space and the image gets converted to black and white. The sign gestures were in Sinhala language.

The method that thy have proposed in the study is to map the signs using centroid method. It can map the input gesture with a database irrespective of the hands size and position. The prototype has correctly recognized 92% of the sign gestures.

The paper by M. Geetha and U. C. Manjusha, make use of 50 specimens of every alphabet and digit in a vision-based recognition of Indian Sign Language characters and numerals using B-Spine approximations. The region of interest of the sign gesture is analysed and the boundary is removed. The boundary obtained is further transformed to a B-spline curve by using the Maximum Curvature Points (MCPs) as the Control points.

The B-spline curve undergoes a series of smoothening process so features can be extracted. Support vector machine is used to classify the images and the accuracy is 90.00%. In, Pigou used CLAP14 as his dataset. It consists of 20 Italian sign gestures. After pre-processing the images, he used a Convolutional Neural network model having 6 layers for training. It is to be noted that his model is not a 3D CNN and all the kernels are in 2D. He has used Rectified linear Units (ReLU) as activation functions. Feature extraction is performed by the CNN while classification uses ANN or fully

ISSN: 2057-5688

connected layer. His work has achieved an accuracy of 91.70% with an error rate of 8.30%.

A similar work was done by J Huang. He created his own dataset using Kinect and got a total a total of 25 vocabularies which are used in everyday lives. He then applied a 3D CNN in which all kernels are also in 3D. The input of his model consisted of 5 important channels which are color-r, color-b, color-g, depth and body skeleton. He got an average accuracy of 94.2

## III. PROPOSED SYSTEM

This system produces the Arabic audio as an output of identified Arabic hand gesture-based characters respectively. This proposed tool appears successful by focusing on the very important and neglected social issue and gives adequate answer for hard hearing users.

### A) Convolution neural network (CNN)

Convolutional neural networks can also be used for document analysis. This is not just useful for handwriting analysis, but also has a major stake in recognizers. For a machine to be able to scan an individual's writing, and then compare that to the wide database it has, it must execute almost a million commands a minute. It is said with the use of CNNs and newer models and algorithms, the error rate has been brought down to a minimum of 0.4% at a character level, though it's complete testing is yet to be widely seen

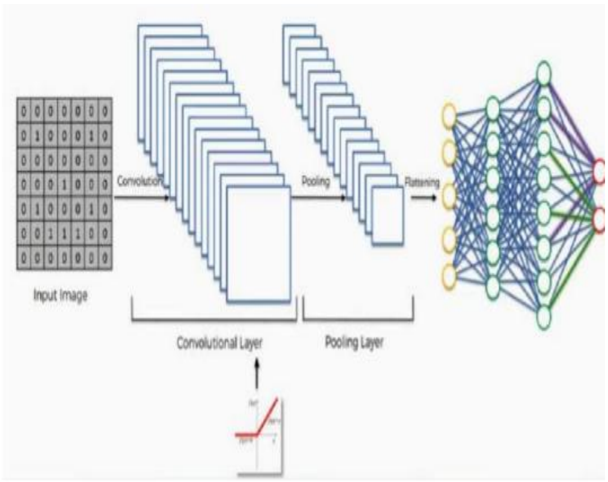



**Fig.2** Layers involved in CNN

### B) GESTURE RECOGNITION SYSTEM

Communication plays a crucial part in human life. It encourages a man to pass on his sentiments, feelings and

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

ISSN: 2057-5688

messages by talking, composing or by utilizing some other medium. Gesture based communication is the main method for Communication for the discourse and hearing weakened individuals. Communication via gestures is a dialect that utilizations outwardly transmitted motions that consolidates hand signs and development of the hands, arms, lip designs, body developments and outward appearances, rather than utilizing discourse or content, to express the individual's musings. Gestures are the expressive and important body developments that speaks to some message or data. Gestures are the requirement for hearing and discourse hindered, they pass on their message to others just with the assistance of motions. Gesture Recognition System is the capacity of the computer interface to catch, track and perceive the motions and deliver the yield in light of the caught signals. It enables the clients to interface with machines (HMI) without the any need of mechanical gadgets. There are two sorts of sign recognition methods: imagebased and sensor- based strategies. Image based approach is utilized as a part of this project that manages communication via gestures motions to distinguish and track the signs and change over them into the relating discourse and content.

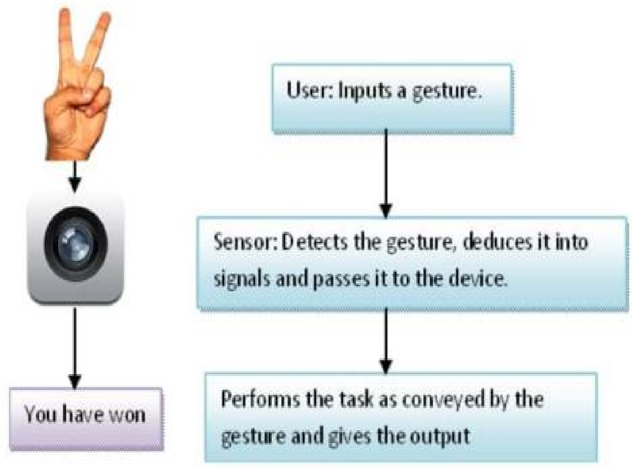



Fig.3 gesture recognition system

### C) HAND SHAPE RECOGNITION USING TRANSFER LEARNING

For hand shape recognition, the classifier is trained through the process of transfer learning over a pretrained CNN that is initially trained on a large dataset. Transfer learning is transferring learned features of a pretrained network to a new problem. The initial layers of the pretrained network can be fixed, the last few layers must be fine-tuned to learn the specific features of the new data set. In our work, VGG16 a CNN architecture is used as the pretrained model. It consists of 13 convolution layers followed by 3 fully connected layers. A convolutional neural network (CNN) 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. We need to recognize eleven hand shapes; hence CNN is trained as a classifier using transfer learning method. To reach the desired output, network model needs to be altered. Therefore, two layers of the model were replaced with a set of layers that can classify 11 classes. All other layers remained unaltered. To avoid over fitting, the Regularization along with a more diverse dataset was introduced. Regularization involves modifying the performance function which is normally chosen to be the sum of the square of the network errors on the training set. The Classifier used over 55 thousand self-created image datasets out of which 70 percent were used for training and rest for testing. If recognized hand gesture is a dynamic hand gesture, then it will further be traced to detect motion.

**D) TRACING OF DETECTED HAND (IF DYNAMIC)**

Recognition of a static gesture requires only the hand shape. Once hand shape is classified as static gesture by the trained classifier, command is given to the computer. Unlike static gesture, dynamic gesture requires both the hand shape as well as the motion of hand. For tracing dynamic hand gestures, hand area is segmented out using HSV (Hue, Saturation, Value) skin color algorithm in a frame, followed by cropping blob area. Centroid of the blob is detected and traced. The main idea in this stage consists in retrieving the coordinates of the traced hand's centre in each frame. These coordinates will be used in order to know which computer command corresponds to which motion. Coordinates will be used differently for each gesture, depending

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE
ISSN: 2057-5688

on detected hand shape. Five out of eleven hand shapes are used for dynamic hand gestures and rest for static hand gestures. These dynamic hand shapes are categorized into unidirectional and multi-directional hand gestures. Unidirectional hand gestures require shape and direction of motion of hand for commanding whereas multidirectional gestures require the position of hand along with its shape. Out of five dynamic hand shapes three are used for unidirectional gestures namely: swap, scroll and zoom, and remaining are used for multidirectional gestures of pointer and cursor. Each unidirectional gesture can further be used for differentiating two hand gestures depending on the direction of motion, e.g., swap can be left or right, depending on the direction of motion of hand. Tracing involves extracting position of hand which is done by skin colour detection, skin cropping, blob detection and centroid extraction. Hence tracing on the whole is a comparatively time-consuming process. This process of tracing can be avoided after certain frames for unidirectional gestures as it only requires the direction of motion which can be derived from the few initial frames. Hence, the direction of unidirectional dynamic gestures can be determined by comparing centroid of initial frames.

## SYSTEM ARCHITECTURE

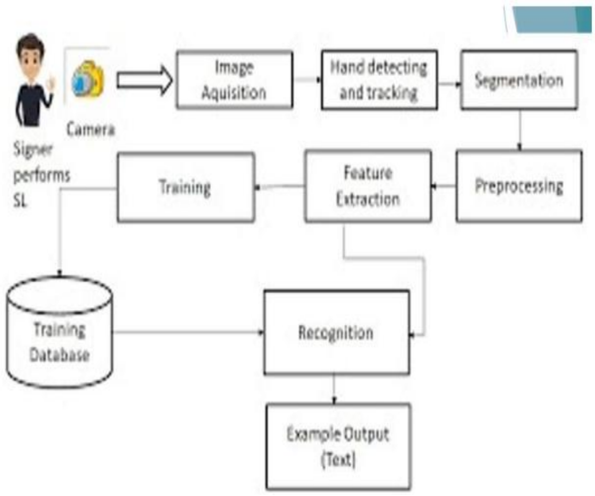



**Fig.4** Proposed system architecture

### E) Training CNN

The created data set is now used to train a CNN. We first load the data using the Kera's Image Data Generator, which enables us to import the train and test set data using the movement from directory function, with the names of each of the number folders

ISSN: 2057-5688

matching the class names for the loaded photograph.

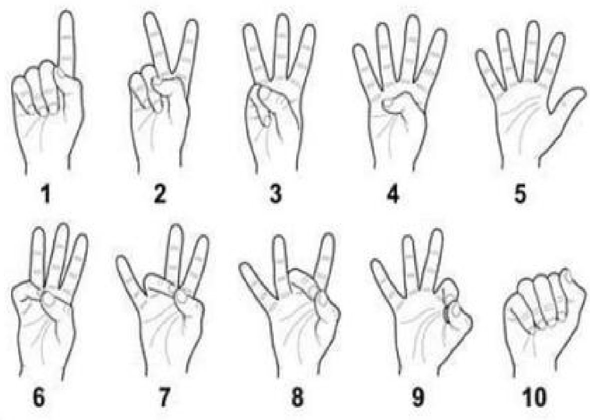



Fig.5 Signs representing Numbers from 1-10

## IV. RESULTS

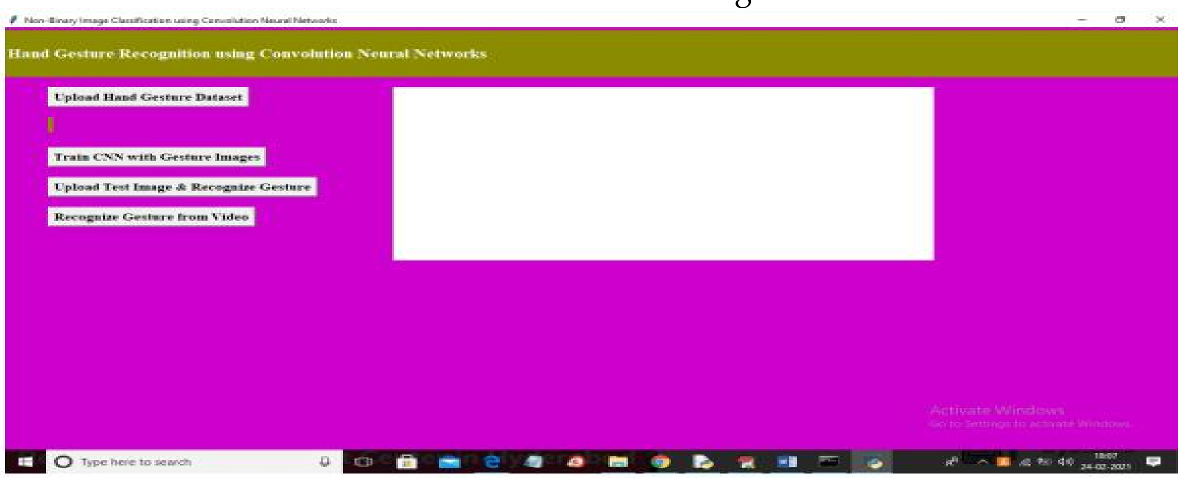
**Data Collection:**

There are various platforms and websites containing many documents like Google Scholar, Scopus, Web of Science, ScienceDirect and many more. These platforms consist of the statistical information which can be used for the Bibliometric Analysis. These platforms contain the research papers as journals, articles, review papers and conference papers. These research papers are easily accessible because of their open access or using the institutional credentials. Here, in this paper for data collection, Web of Science is considered as it has an immense amount of data for Bibliometric analysis

Table.1 Database Search Query

| | |
|---|---|
| Web of Science | "indian sign language" |
| Scopus | "indian sign language" |

To run project double click on run.bat file to get below screen



INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

ISSN: 2057-5688

Fig.6 In above screen click on 'Upload Hand Gesture Dataset' button to upload dataset and to get below screen

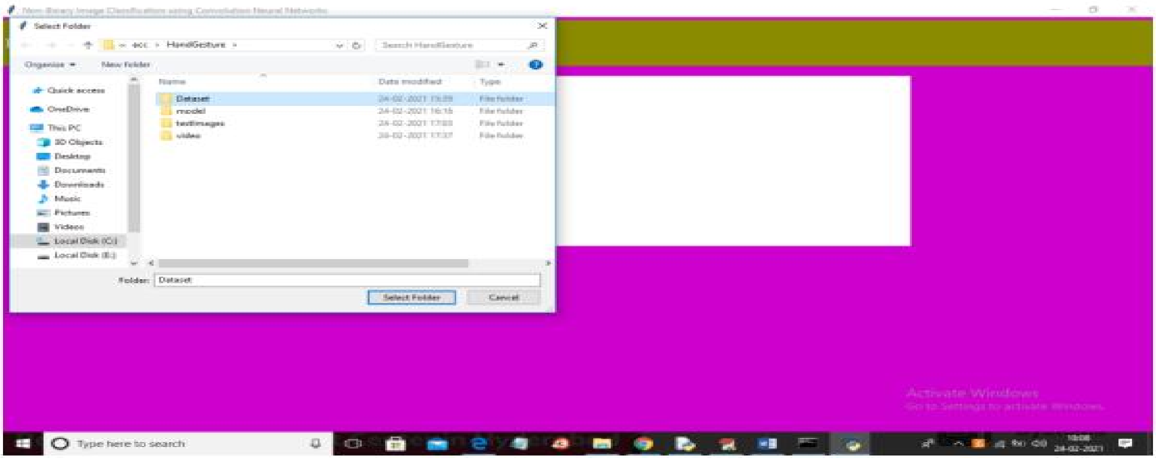



Fig.7 In above screen selecting and uploading 'Dataset' folder and then click on 'Select Folder' button to load dataset and to get below screen

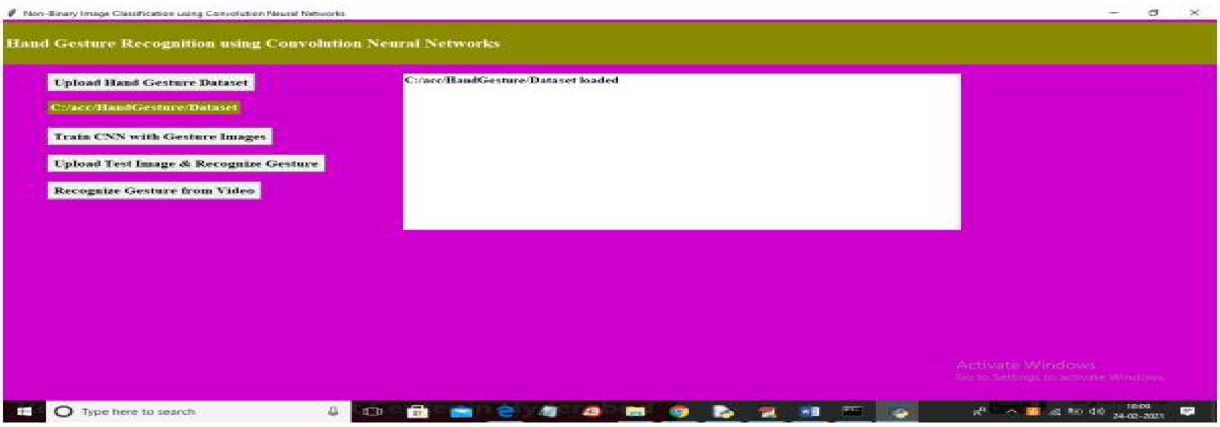



Fig.8 In above screen dataset loaded and now click on 'Train CNN with Gesture Images' button to trained CNN model and to get below screen

ISSN: 2057-5688

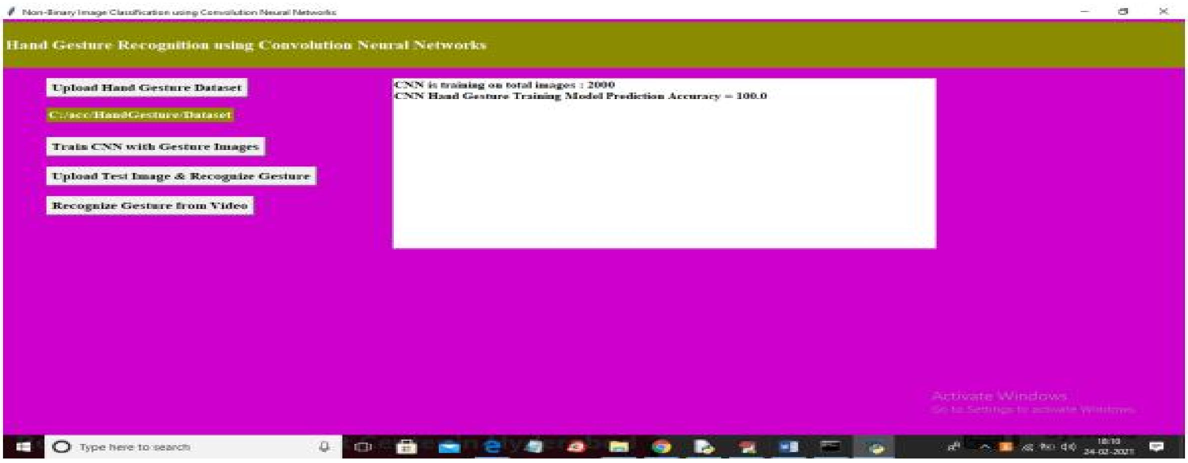



Fig.9 In above screen CNN model trained on 2000 images and its prediction accuracy we got as 100% and now model is ready and now click on ‘Upload Test Image & Recognize Gesture’ button to upload image and to gesture recognition.

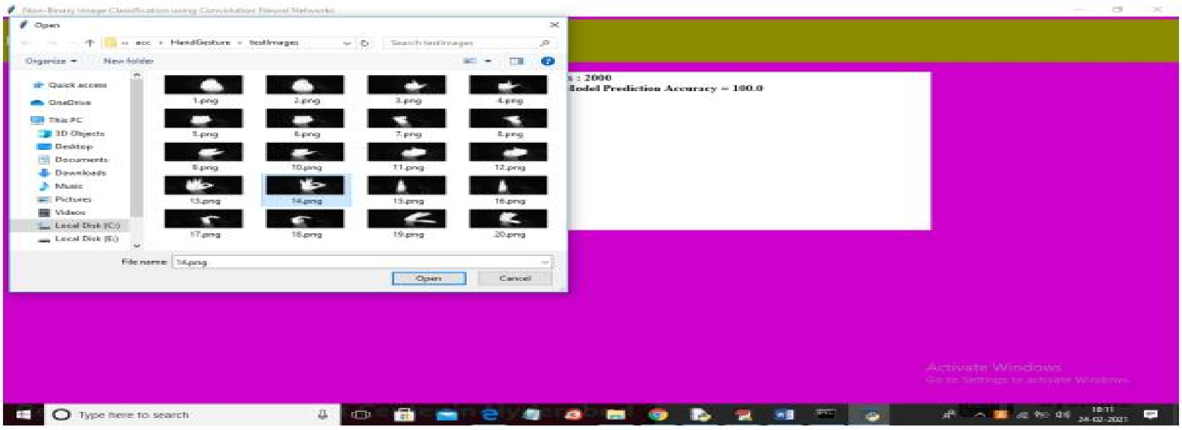


Fig.10 In above screen selecting and uploading’14.png’ file and then click Open button to get below result

ISSN: 2057-5688

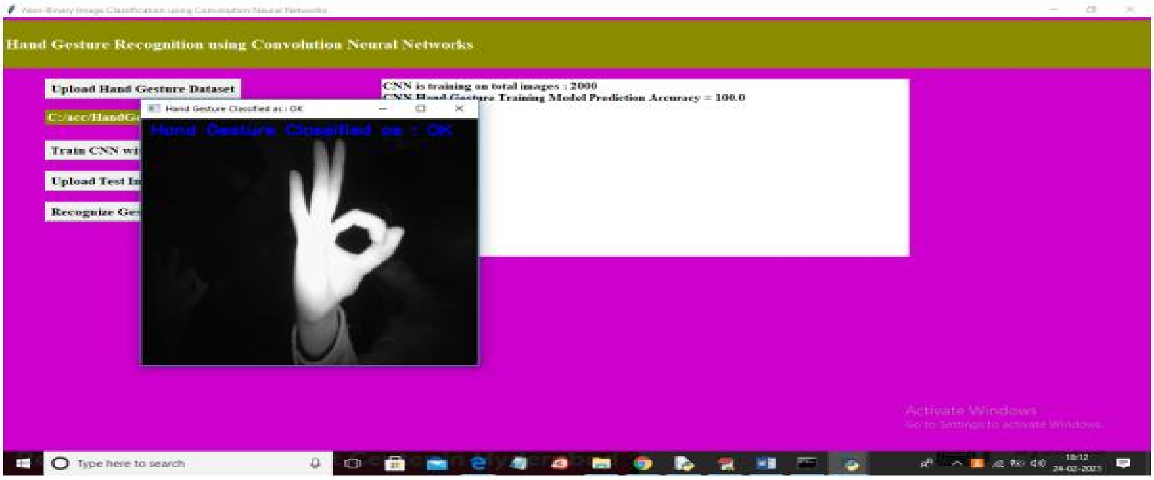



Fig.11 In above screen gesture recognize as OK and similarly you can upload any image and get result and now click on 'Recognize Gesture from Video' button to upload video and get result

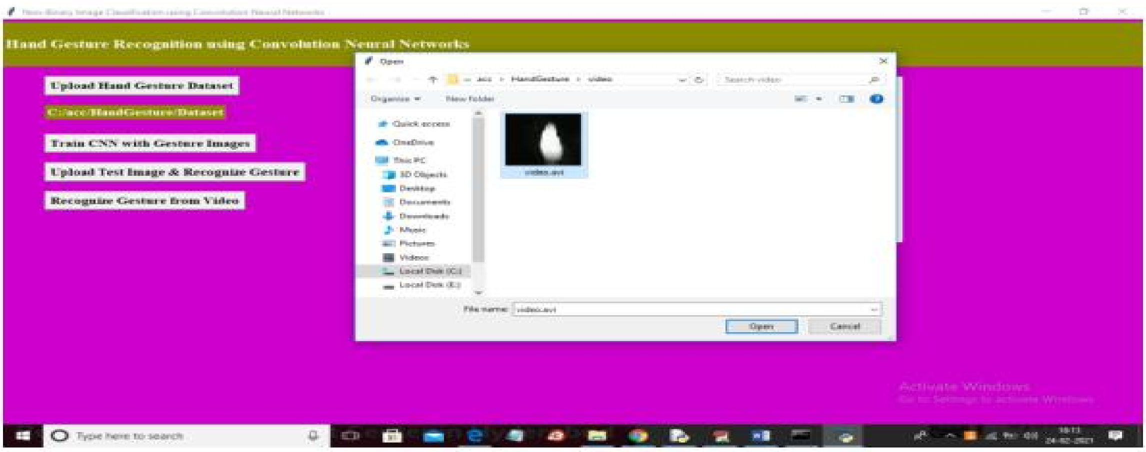



Fig.12 In above screen selecting and uploading 'video.avi' file and then click on 'Open' button to get below result

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE
ISSN: 2057-5688

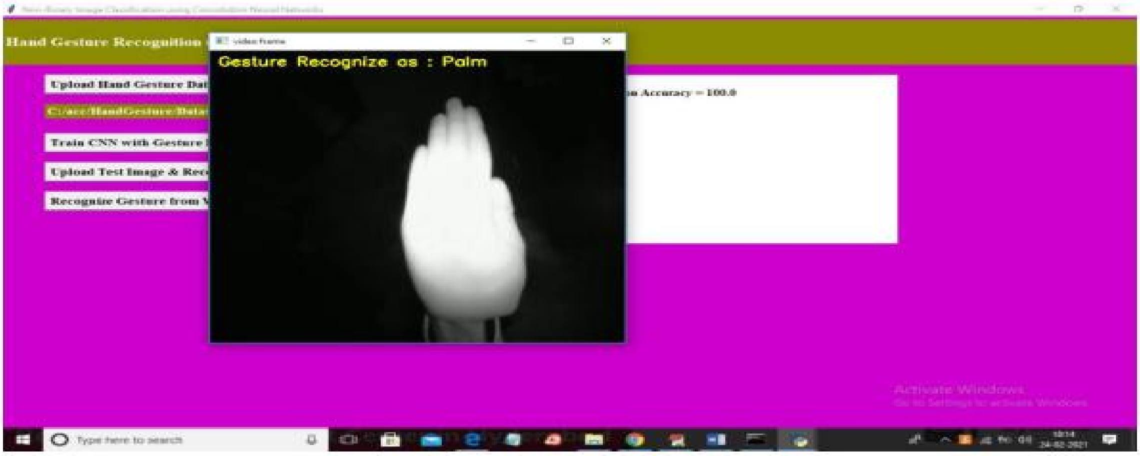



Fig.13 Hand gesture recognize as Palm

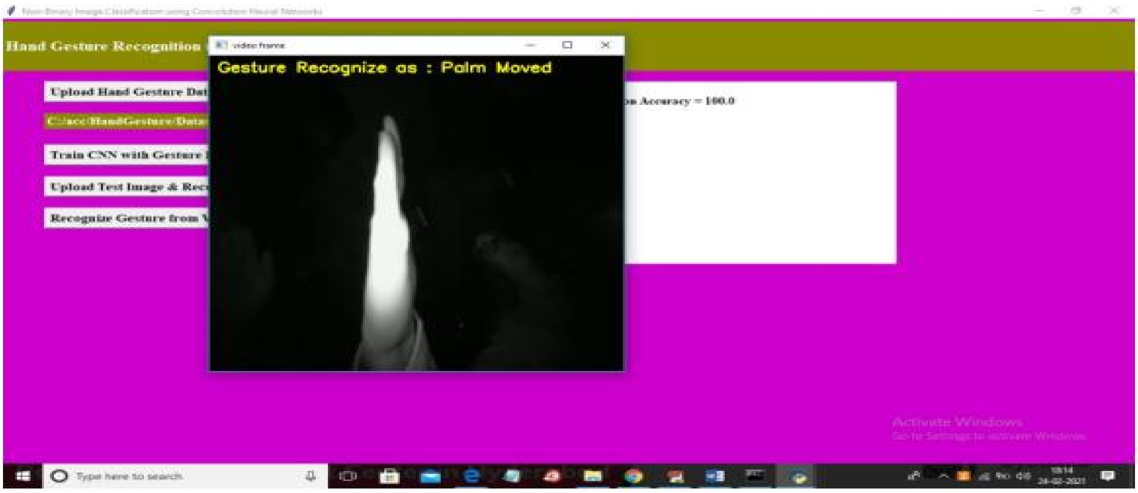



Fig.14 Hand recognize as Palm moved


Fig.15 Hand recognize as Five

In above screen as video play then will get recognition result

## V. CONCLUSION

Sign languages are very broad and differ from country to country in terms of gestures, body language and face expressions. The grammars and structure of a sentence also vary a lot. In our study, learning and capturing the gestures was quite a challenge for us since the movement of hands had to be precise and on point. Some gestures are difficult to reproduce. And it was hard to keep our hands in exact same position when creating our dataset.

## REFERENCES

[1] https://peda.net/id/08f8c4a8511

[2] K. Bantupalli and Y. Xie, "American Sign Language Recognition using Deep Learning and Computer Vision," 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 2018, pp. 4896-4899, doi: 10.1109/BigData.2018.8622141.

[3] CABRERA, MARIA & BOGADO, JUAN & Fermín, Leonardo & Acuña, Raul & RALEV, DIMITAR. (2012). GLOVE-BASED GESTURE RECOGNITION SYSTEM. 10.1142/9789814415958_0095.

[4] He, Siming. (2019). Research of a Sign Language Translation System Based on Deep Learning. 392-396. 10.1109/AIAM48774.2019.00083.

[5] International Conference on Trendz in Information Sciences and Computing (TISC). : 30-35, 2012.

[6] Herath, H.C.M. & W.A.L.V.Kumari, & Senevirathne, W.A.P.B & Dissanayake, Maheshi. (2013). IMAGE BASED SIGN LANGUAGE RECOGNITION SYSTEM FOR SINHALA SIGN LANGUAG

[7] Prasadu Peddi (2021), "Deeper Image Segmentation using Lloyd's Algorithm", ZKGINTERNATIONAL, vol 5, issue 2, pp: 1-7.

[8] N. Srivani, Dr Prasadu Peddi, "Efficient Fr a Geometrical-Model-Based Face Segmentation and Identification in Terms of Identification the Face ", *JFCR*, pp. 1283-1295, Jun. 2022.