

# Gesture Recognition ITSP 2020

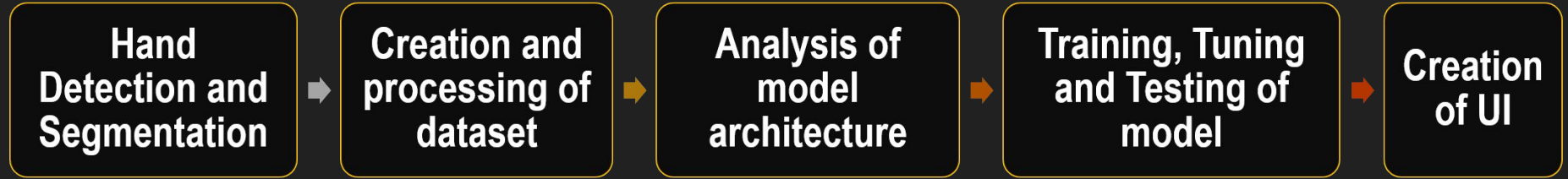
Aakriti  
Divyansh Nankani  
Shirish Chinchankar  
Monu Meena

## Idea behind the project

Human Computer Interaction for ease of communication between man and machine has become a necessity. Hand gesture is the most easy and natural way of communication. We decided to develop a vision based real time gesture recognition system which would work only using a Webcam.

Taking into account the numerous use cases of this technology we decided to implement its functionality for sign language translation, gesture based gaming and dynamic gesture control

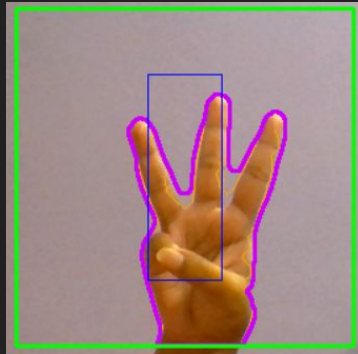
# Workflow



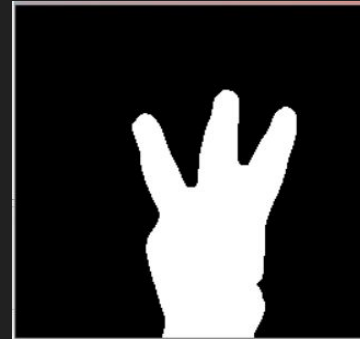
# Project Description

## Hand Detection and Segmentation

The objective of our algorithm was robust skin color detection and removal of static background for segmentation of hand out of any image. This was done firstly by subtracting a calibrated background. Then skin colour feature was extracted and by using back projection, the hand position and orientation was detected in an image. Contours were also detected in the thresholded image.



Original Image



Segmented Image

## Recognition and Classification of gesture

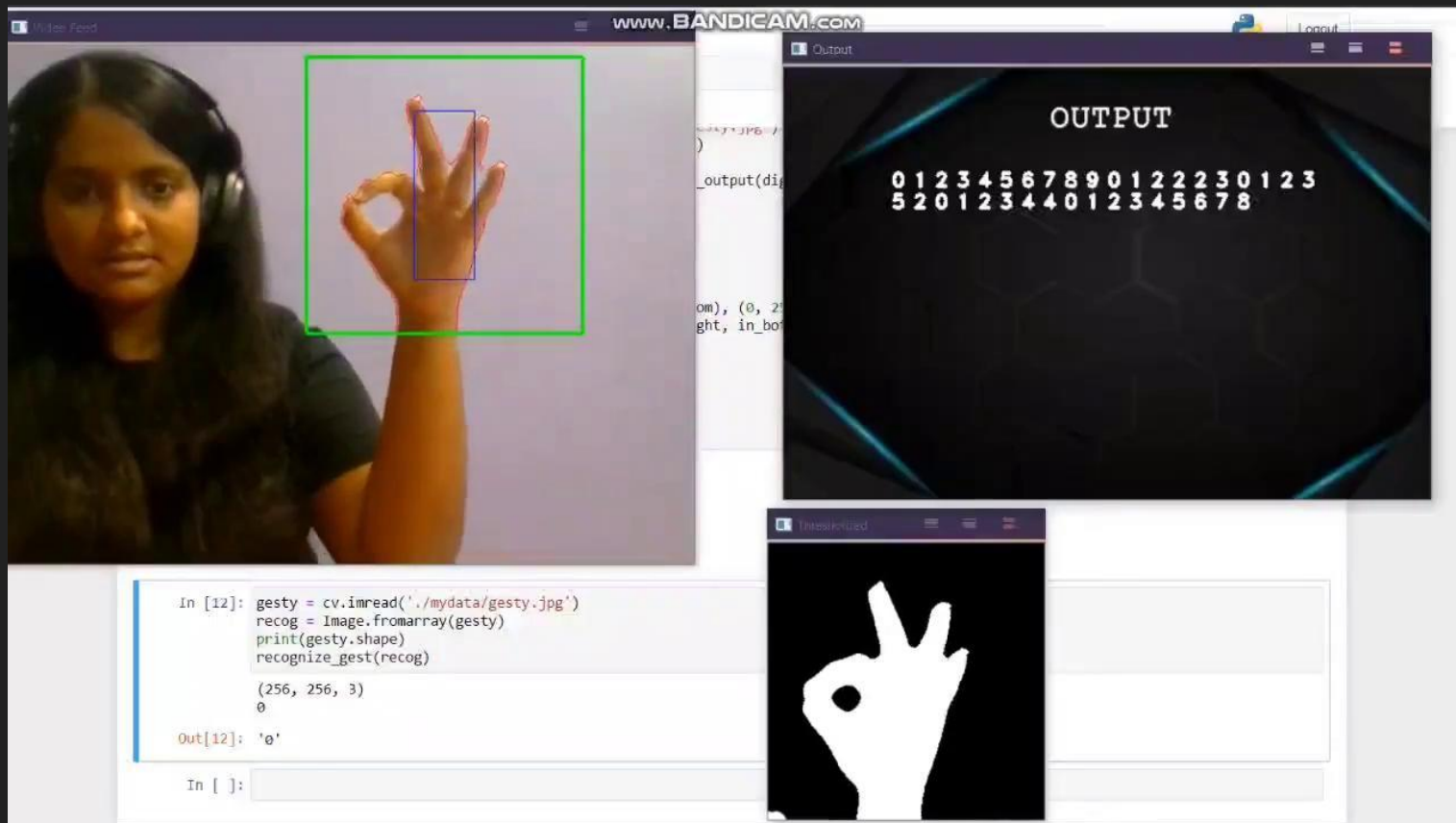
Deep learning was used for classification. For static gesture recognition convolutional neural networks were used and accuracies with various pre-trained models were explored for implementing transfer learning. VGG-16 and GoogleNet gave best results in real time recognition. While implementing dynamic recognition, C3D model based on 3D CNNs and LCRN model based on CNN+LSTM were used to process multiple frames of a video.

Dataset for the recognition was partially created ourselves and partially obtained from the internet and processed.

## UI for Output

Image manipulation with OpenCV was used for Rock-paper-scissors game and sign-language translator. PyAutoGUI was used for gesture control of applications

# Results



# Skills learnt

We learnt how we can use machine learning to connect humans with computers. We acquired knowledge about how Neural Networks, Convolutional Neural Networks, Long Short Term Memory units and Time Distributed Layers function and how they can be implemented in PyTorch and what is their significance. We get introduced to Open CV software and how to employ computer vision to process the response of a computer towards the image captured by the camera.

On a more personal level we learned how to conduct research and explore the existing technologies available in an area of interest as well as ideate innovative methods to tackle a problem at hand. We developed the ability of reading and interpreting research papers as well as analysing the content to deploy it with suitable changes to suit our purpose.

# Challenges faced

Data pertaining to human features such as hand vary a lot and training on online datasets with raw image is very difficult. Skin colour variations can add too much variance to data. Moreover, simple cameras can add too much noise to background in real time detection making feature extraction difficult.

The biggest challenge faced perhaps was the choosing of model architecture and its subsequent training. Most of the time was spent in analysing the accuracies with different changes to the learning rate, the number of convolutions, layers, time distributions and the optimisation methods.



# Conclusion

There are many challenges still associated with the accuracy and usefulness of hand gesture recognition software. For vision based gesture recognition there are limitations on the equipment used and image noise or unsuitable background. Nevertheless, the fact that our program is capable of achieving considerable accuracy on simple laptop's Webcam is remarkable and shows the immense potential of development and research in this area to achieve better accuracy and develop more robust models for Sign Language Translation and Human Computer Interaction using hand gestures