SUMER OF CODE

**Gesture Based Text Creation & Recognition**

Final Report

## Chhavi Yadav [23B3923]

horizontal line

# Introduction

This project started from basics of Machine Learning , OpenCV and python libraries like Pandas, Numpy etc and built all the way up to creating virtual notepad. The final result is a project where we can write on a virtual canvas by making air doodles using finger which would then be shown on a canvas. It extended further where any digit made on the canvas would be recognized using digit recognition models. Some other features such as volume control and brightness control were also added , which could be controlled using finger gestures.

# Weekwise Learning and Assignments

Week 1 - This week covered the basics of python including operators, loops and conditions classes, functions, modules and packages. Essential libraries like numpy, matplotlib and pandas were covered and used in the assignment questions . We also learned about what machine learning is,its terminology its types(supervised and unsupervised). I also learnt the difference between machine learning and deep learning , the later being subset of the former. These topics built a foundation for ML. The assignment included questions implementing each of these topics.

Week 2 - Machine Learning was covered more in depth. I learnt how Convulational Neural Network (CNNs) are built ie their basic structure. Edge detection, padding, layers of CNN,the mathematics behind loss function, and usage of CNN was covered. The assignment was to build a MLP using different functions such as activation function, loss function ad so on. This assignment showed the basic functioning of a model in much more depth. We also followed some tutorials on using tensorflow to build neural nets and cnns.

Week 3 - Continuing of week 2, we learnt about data augmentation which is useful when the training set is limited and we dont want to overfit . We learnt about some famous architectures such as ResNets, VGG and their advantages. We build different models on our own . Using CIFAR100 dataset, we made a CNN. Fine tuned pre built assignments and compared their accuracy. This showed the effect of different attributes on the model. A pneumonia or normal chest x ray classifier ws also built. We learnt to build a model from scratch, how to decide their architecture for better accuracy, change factor such as padding , dropout, learning rate , epochs and batch size to improve the accuracy.

Week 4 - This week was more based on learning , where we covered an advance course on computer vision using python. We used openCV and mediapipe and followed the video to learn how to make hand tracking module, face mesh module, face detection module , pose detection module . These modules were then further used to create programs to control volume and counting fingers using handtracking by computer vision.

Week 5 - This week followed the previous week to implement all the modules created to make brightness controller, drag & drop, screen pause programs. The brightness controller followed the basics of volume controller and used wmi library instead. Screen pause was based on counting fingers on the screen and pausing if the exceed 3, it followed the structure same as that of finger counting and developed further on it. Drag and drop image was based tracking the movement of fingers which could be done by estimating the position of different points on the hands.

Week 6 - The final week used all the knowledge of the previous weeks to create the project. Using computer vision hands on the screen were detected and a virtual canvas was made. This canvas included a header for all the options provided for painting. Using 2 fingers for selection and one for drawing, different colors or eraser were selected and then used upon the image captured by the webcam. The drawing on the screen were further processed upon choosing predict option, where we used digit recognition model trained using mnist dataset. The image captured was first processed and then a prediction was made on which number is drawn on the screen. We used more than 1 models and final ans was voted amongst these models to further improve the prediction accuracy. This virtual notepad also has the volume and brightness controller included in it.

# Conclusion

This project served as a great learning opportunity for computer vision and its practical uses. This project can be further developed by making different finger gestures represent different functions such as taking a screenshot, pausing the screen or showing the result on the screen itself.