ACM-W 5th National Hackathon - 2020

Hand Talk

Technocrats

Team Introduction

TEAM TECHNOCRATS

Sri Sivasubramaniya Nadar College of Engineering

1

Lakshmi Priya B

Nivedhitha D

Sakthi Sairaj

Problem Statement

17. Gesture Recognition for Human Computer Interaction

- Alternative way of effectively interacting with a computer system by using hand gestures for input and controls
- Simulating the functions of the keyboard alphabets, cursor and common controls like switching slides, etc using hand gesture motions
- Help users interact with the computer from a considerable distance without requiring any device like a mouse or keyboard
- Eliminate or at least decrease the hardware requirement for effective human-computer interaction

Probable Domains of Impact

Gesture controlled interaction has a plethora of professional, recreational and supportive applications

- In the prevailing pandemic situation, it prevents contact with surfaces, for example, billing stations and manufacturing warehouses
- It facilitates interaction with a computer from a considerable distance, for, example, changing slides while taking a presentation
- No need for extra hardware and mimics real actions, heavy machinery operations and game playing

Support for deaf and mute people to communicate with a computer generated voice

4

Implementation Plan

- Plan of Action
- Technology Stack
- Workflow
- Product Pipeline

Plan of Action

STEP 1

Neural Network module to load serialized face detector and face embedding model from disc to authorize the user

STEP 2

Define the classification categories for gestures and collect training images to create a clean dataset

STEP 3

Train and test the gesture recognition CNN classifier using dlib from scikit-learn and OpenCV

STEP 4

The gesture selected is used to perform a select action

STEP 5

The corresponding phrase or keywords associated with the gesture is recognized by the speech module

STEP 6

The keywords associated with the gestures are converted to complete sentences using BERT

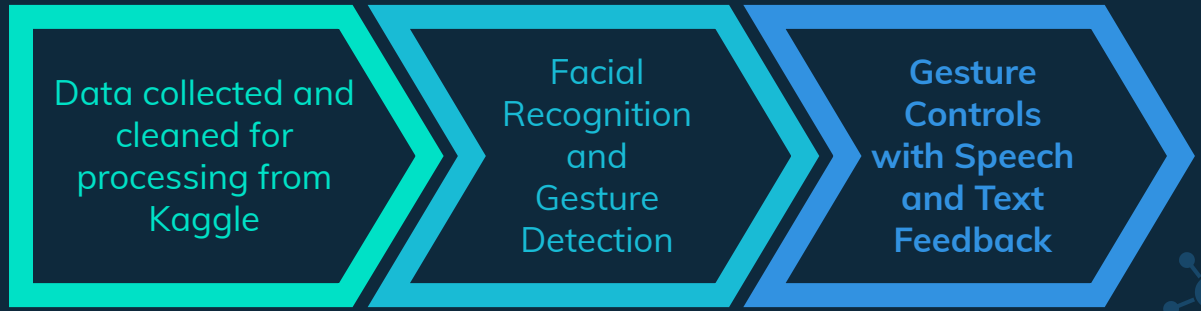
Technology Stack

Packages or tools	Components used	Purpose
Python 3.9	• Interpreter • Related packages pywin32	Primary language of development of the solution
OpenCV, scikit-learn	• dnn • blobFromImage	• Deep Neural Network module to load serialized face detector and face embedding model from disc • To construct a blob from image
Numpy	Arrays and vectors	Vectorization and image representation
sentence_transformers	Sentence Transformer	For BERT model
Keras	SqueezeNet	Pre-trained CNN Algorithm for classification

Workflow

1. Collect live web camera feed
2. Run face identification and gesture recognition processes parallelly
3. If the person performing the gestures is not authorized, end both processes
4. Else, perform the necessary gesture recognition and conversions

Product Pipeline

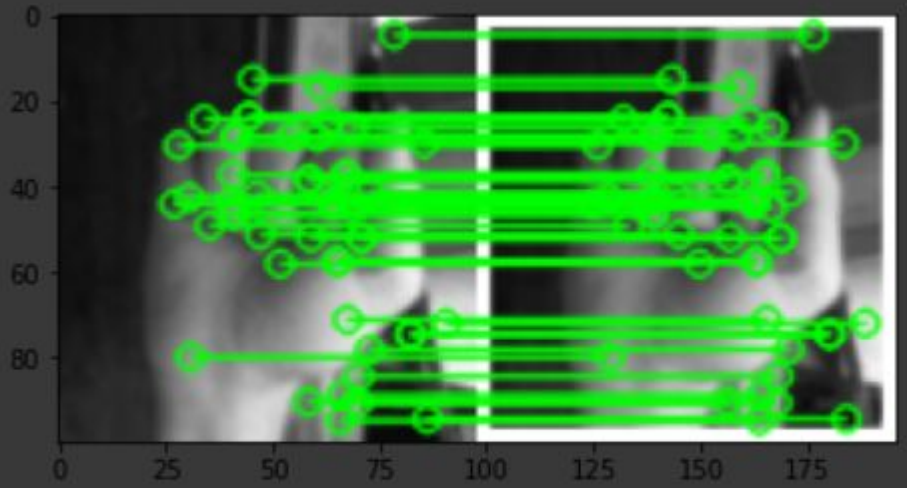


5

Demo

Code and results obtained during the development lifecycle highlighting the successes and bugs

Approach 1



['a/a_1.jpg', 'a/a_1.jpg']

MATCH!!!!

Approach 1

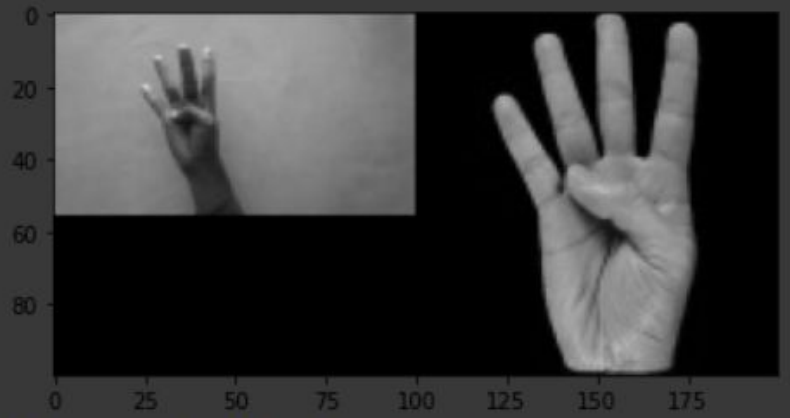
Problem

Using just a SIFT detector and FLANN matcher didn't deliver significant gesture recognition performance

Resolution

A CNN or convolutional neural networks algorithm was employed for gesture image classification

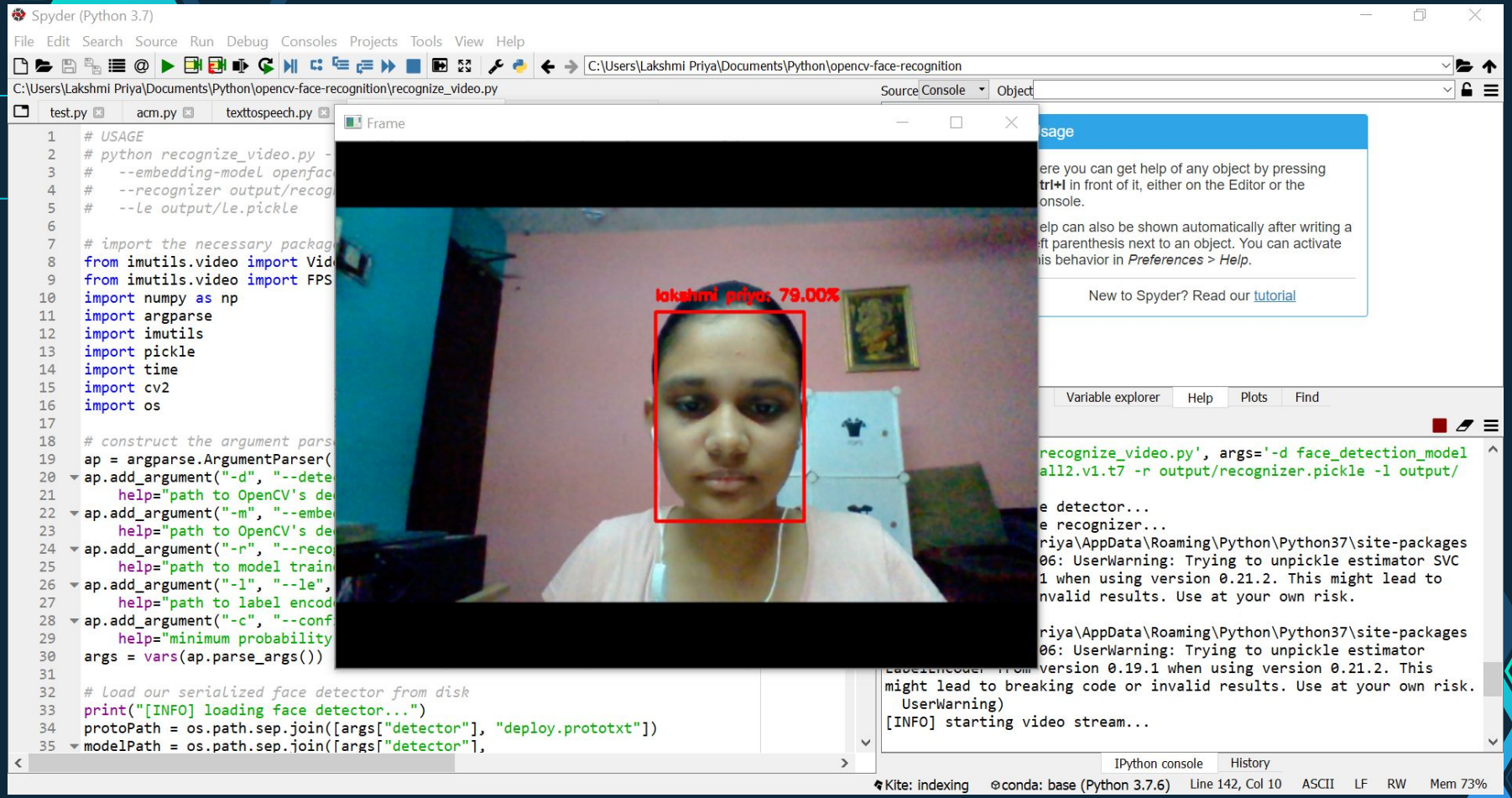
```
Not enough matches are found - 0/20
```




```
['b/b_1.jpg', 'b/b_100.jpg']

NO MATCH



TESTING

*****CONSOLIDATED TESTING REPORT******
Testcase  0 :  PASSED: MATCH
Testcase  1 :  FAILED: NO MATCH
```

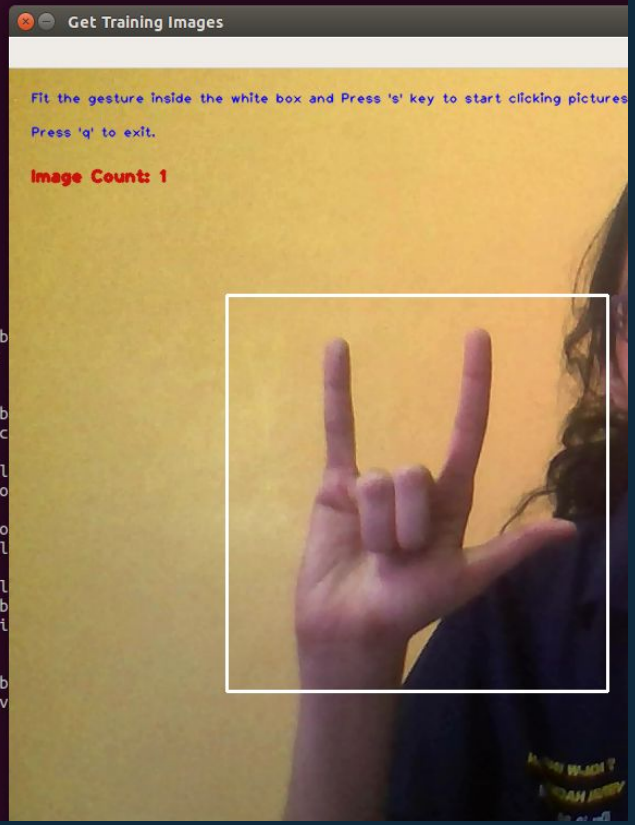
Approach 2

- Facial recognition for authorization and security using a SVM
  - To identify and verify the presence and location of a human face in an image and authorize
  - To extract the 128-d feature vectors (called “embeddings”) that quantify each face in an image
- Gesture classification and controls
  - A CNN image classifier takes the video feed and classifies the selected frame into one of the possible gesture categories that it was trained to identify
- The gesture will produce the corresponding action along with speech and text feedback

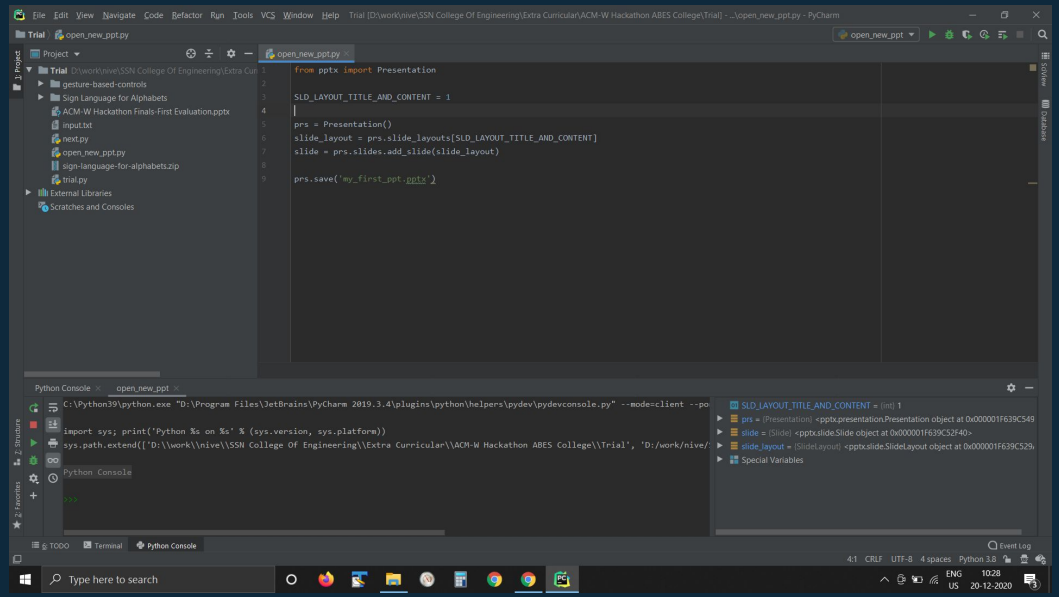

Spyder (Python 3.7)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\Lakshmi Priya\Documents\Python\opencv-face-recognition
C:\Users\Lakshmi Priya\Documents\Python\opencv-face-recognition\recognize_video.py
test.py
acm.py
texttospeech.py
Frame
lakshmi priya: 79.00%
Variable explorer
Help
Plots
Find
IPython console
History
Kite: indexing
conda: base (Python 3.7.6)
Line 142, Col 10
ASCII
LF
RW
Mem 73%

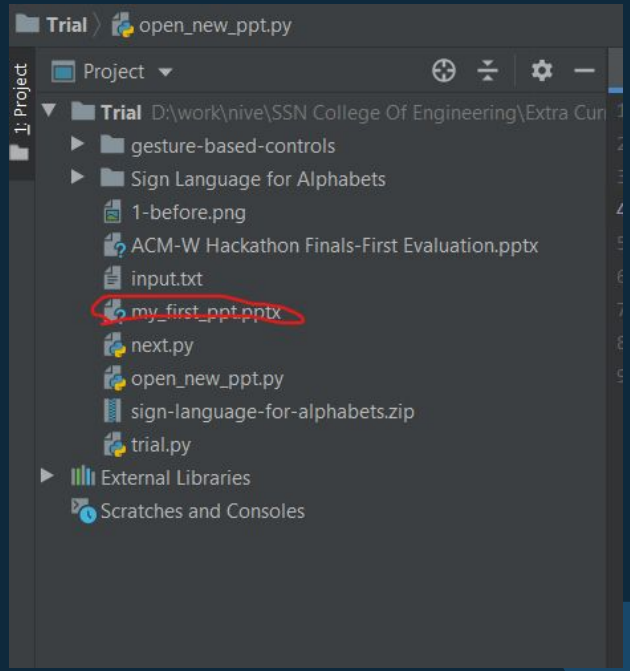
Get Training Images

```
4
Predicted Gesture: See you later
1
Predicted Gesture: I love you
1
Predicted Gesture: I love you
5
Predicted Gesture: Yes
1
Predicted Gesture: I love you
^CTraceback (most recent call last):
  File "model_test.py", line 126, in <module>
    k = cv2.waitKey(10)
KeyboardInterrupt

parsermouth@parsermouth:~$
parsermouth@parsermouth:~$ python3 model_test.py
Using TensorFlow backend.
WARNING:tensorflow:From /home/parsermouth/.local/lib
from tensorflow.python.framework.ops) is deprecated
Instructions for updating:
Colocations handled automatically by placer.
WARNING:tensorflow:From /home/parsermouth/.local/lib
nsorflow.python.ops.nn_ops) with keep_prob is deprec
Instructions for updating:
Please use `rate` instead of `keep_prob`. Rate shoul
2020-12-20 10:41:28.880753: I tensorflow/core/platfo
 not compiled to use: AVX2 FMA
2020-12-20 10:41:29.162407: I tensorflow/core/platfo
2020-12-20 10:41:29.194116: I tensorflow/compiler/xl
vices:
2020-12-20 10:41:29.194279: I tensorflow/compiler/xl
WARNING:tensorflow:From /home/parsermouth/.local/lib
rom tensorflow.python.ops.math_ops) with keep_dims i
Instructions for updating:
keep_dims is deprecated, use keepdims instead
WARNING:tensorflow:From /home/parsermouth/.local/lib
python.ops.math_ops) is deprecated and will be remov
Instructions for updating:
Use tf.cast instead.
1
Predicted Gesture: I love you
```



Code to open a PPT from python





USP and Novelty

- Gesture to Speech conversion is still not extensively deployed in the market as support software for the deaf and mute, but, big tech giants have published algorithms for the same
- Our model binds together various well-established modules like face identification, image classification, text to speech conversion, text commands to action stimulation into one cohesive whole and uses all these to create the “Hand Talk” for the disabled

7

Scope

Further improvements and other ideas

Scope

- Using infrared sensors for face recognition
- Cursor movement

Thank You

Any Questions?