IRIS

NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA, SURATHKAL
ರಾಷ್ಟ್ರೀಯ ತಂತ್ರಜ್ಞಾನ ಸಂಸ್ಥೆ ಕರ್ನಾಟಕ,ಸುರತ್ಕಲ್
राष्ट्रीय प्रौद्योगिकी संस्थान कर्नाटक, सुरत्कल
P.O. SRINIVASNAGAR, MANGALORE-575 025

CCTV Attendance Marking - IRIS Labs Proposal

13/08/2019

A Naga Neeramitra Reddy
17CO101

Rishi Sharma
17CO135

Member 1 Name and Contact Information

Name: A Naga Neeramitra Reddy

Email: neeramitra.reddy@gmail.com

Website: neeru1207.github.io

Blog: NA

GitHub: https://github.com/neeru1207

Member 2 Name and Contact Information

Name: Rishi Sharma

Email: rishi94075@gmail.com

Linkedin: **https://www.linkedin.com/in/rishi-sharma-278999152/**

Blog: https://medium.com/@rishi94075/the-karate-kid-5c4d0ddfe67

GitHub: **https://github.com/kampaitees**

About Us

Education

Member 1:

Department: Computer Science

Degree level: Btech

Roll Number: 17CO101

Graduation year: 2021

Member 2:

Department: Computer Science

Degree level: Btech

Roll Number: 17CO135

Graduation year: 2021

Our previous experience related to the project:

Member 1:

I have worked on a project related to live webcam face recognition using OpenCV and Tensorflow in Python. Link to the project:

https://github.com/neeru1207/GUI-Face-recognizer

I have also built an RNN to predict Google stock prices based on past values. Link:

https://github.com/neeru1207/RNN-to-predict-Google-Stock-Prices

Member 2:

I have been working in AI/ML/DL for more than a year.

Last summer, I did an internship in Karnataka state police computer wing in collaboration with IISC Bangalore. They wanted me to build a website for criminal image search. So for this task, I built a website which takes in an image of criminal and outputs all similar images available in the database.

Even when the images were of bad quality with low resolution and scratches, my model worked perfectly. Link to the project:-
https://github.com/kampaitees/Web-App-for-Image-Search

In the same summer, I also did a work from home internship where I was 1 among 36 candidates who were selected globally for doing this project. The task was to detect anomalies from the images of Mars which were provided by the mentor who was a NASA Scientist who provided us with data. Data consisted of huge images taken by the rover. I used Data Augmentation as deep learning requires huge datasets. Link to certificate :-

https://github.com/kampaitees/Certificates/blob/master/Mars_Rishi_Omdena%20Certification.png

Why this project is important to us?

As we have been exploring Deep Learning/Computer Vision and would love to have a career in the same, this project would be a great opportunity to solve a real world challenge using our skills and knowledge while being an amazing learning experience. As this project (If successful) would drastically change the attendance system, it would be a great achievement for us and hopefully a valuable addition to NITK IRIS. It would also serve as an amazing project to showcase our skills in future career endeavours.

About The Project

1. CCTV Attendance Marking

Attendance marking in a classroom during a lecture is not only burdensome but also a time consuming task, not to mention the infamous proxy problem. This project introduces an automated approach using Computer Vision and Deep learning. Developing a real time face recognition approach which automatically detects and registers students attending a lecture while being time and resources efficient at the same time would be the goal.

2. Approach

1. **Image Processing:**

A number of **image processing techniques**, in addition to **enhancement techniques**, can be applied to improve the data usefulness. Techniques include convolution, edge detection, filters, trend removal, gamma correction and image analysis.

These techniques make the images better for processing through models and drastically decreases the time to process them through various deep learning models.

2. **Face detection:**
Faces are detected using DLIB library which is a state-of-the-art library for detecting faces from the images. If DLIB is not able to detect faces from the input image, then the image is passed through OpenCV's HaarCascade classifier using the frontal face default haar cascade which while being more accurate is a bit slower than DLIB.

3. **Face extraction:**
As we have now detected the face from the image using DLIB and OpenCV, now we will extract only the faces previously detected by the above algorithms, since face is the only region of interest. This step will ensure better time and resources efficiency.

4. **Facial Alignment:**
As of now we have first detected the face and extracted it, but there may be a possibility that the face is not vertical, it may be tilted at some angle which can affect the performance of models as it will be trained on vertical images so facial alignment will be done. Facial alignment is done using Dlib which has state-of-the-art deep learning models to do facial alignment. The model we are going to use for this task is **'68 Landmarks Predictor'** which will give the 68 landmarks of our face then with some preprocessing face will be made vertical which is the requirement of our deep learning model.

5. **Training Deep Convolutional Network:**
As of now we have well aligned face of student, but here comes the most important part we need a model which will predict whether given input image is an image of particular student or not. So we are going to use **Siamese Network** which will predict the 128 dimensional vector given a face of student, which is nothing but 128 important features of image which will make it distinguishable from other images. Here we will also going to use the concept of **Transfer Learning** where we will use some network which is already trained on different images. This already trained models is way better than training model from scratch as it has some knowledge of images.

6. **Predicting encodings of aligned face:**
As of now, we have vertically aligned faces of the students and a pre-trained deep learning model which predicts 128 important features given an image. We will now pass the face of

student to a pre-trained deep learning model which will give the 128 dimensional vector which represents the encoding vector of that image.

7. **Calculating distance between images:**
The above 5 steps will be repeated with all the images and the 128-dimensional encodings of all the students face will be stored in database for faster calculation because if we are not going to store the encodings in database then we have to calculate the encodings of all images at real time which will definitely be time consuming. So given a test input image of student, it's encodings will be calculated and will be compared with each of the encodings already stored in the databases. If the image matches with any student in the database, attendance of that student will be updated on iris automatically. For comparing two images either '**Spatial Cosine'** distance or **'Mahalanobis'** distance will be used as a metric as it will restrict the distance between images between 0 and 1.

Libraries and Tools that will be used:

- Tensorflow
- Keras
- Opencv
- Dlib
- Numpy
- Pillow
- Scipy

Software development Methodology:

We will use agile SDLC model because

- The small project development can be shown in small intervals, say 2 weeks.
- Both team members and Scrum masters are going to work closely.
- Early and frequent changes to the product.
- Documentation requires lesser priority than software development.

3. Deliverables

- OpenCV HaarCascade detector performance report and blog.
- DLIB face detector performance report blog.
- DLIB facial landmarks predictor performance report and blog.
- Facenet performance report and blog.
- Baidu's model performance report and blog.
- Deep Face performance report and blog.
- Best performance model will be used for the final product.

4. Plan

22 Aug - 15 Oct : Go through Stanford and Andrew Ng's Lectures where they use vanilla convolutional neural network for facial recognition, will show the results and write a blog.

16 Oct - 15 Nov : Go through survey papers to see which models could be better than a vanilla convolutional neural network, will show the results and write a blog.

16 Nov - 15 Dec : Look up and find optimal techniques necessary for the project's image processing. Eg: Gamma Correction, Image augmentation, Image segmentation

16 Dec - 15 Feb : Use Facenet/Baidu model/Deep face for prediction, will compare the results and pick the best one based on accuracy and speed.

16 Feb - 20 Feb : Final Project and a Presentation giving a comprehensive overview of the entire project.