LaTeX Proposal for Computer Vision

Murtadha Marzouq
Graduate Student
mmarzouq@charlotte.edu

Param Patel
Graduate Student
ppate211@charlotte.edu

Haochen Ye
Graduate Student
hye5@charlotte.edu

Sam Aldehayyat
Undergraduate Student
saldehay@charlotte.edu

Yuepei Yu
Undergraduate Student
yyu20@uncc.edu

Abstract

In educational settings, assessing classroom occupancy is integral for attendance keeping purposes. This project aims to develop a sophisticated computer vision system aimed at automating the attendance process by accurately detecting and recording the presence of students within images taken in a classroom by an instructor. We will utilize well-known datasets and methods for examining the images which will be introduced further along in the paper. It will contain a detailed explanation of the problem, planning the solution, implementing it and the lessons learnt throughout this process.

Keywords: Computer Vision, Object Detection, Image Processing, Student Counting, Classroom Management.

1. Introduction

Attendance is an indispensable metric for a student's academic journey, where consistent participation correlates with academic success. The process of monitoring and maintaining attendance records presents significant challenges, which mostly originates from constraints in time and resources. By harnessing the benefits of computer vision, this project addresses the problem of automating the attendance-taking process in classrooms using a tool that is at every teacher's disposal, their cell phone camera.

2. Problem Statement

Attendance enforcement and record keeping in the class room is often overlooked because of added time and resources. Additionally, the COVID-19 pandemic has made it difficult to enforce attendance in the classroom. In this project, we will develop a system that will allow students to check in and out of a classroom using face recognition.

3. Data

1. The data for this project will be collected from the University of North Carolina's Canvas API (Will use Murtadha's Account) to fetch the number of students in the Computer Vision course. This information will be stored to be references to give a percentage of students attending.
2. The dataset that will be used to train our model is the Faces in the Wild (LFW) dataset. This dataset contains more than 13,000 images of faces collected from the web. Each face has been labeled with the name of the person pictured. 1680 of the people pictured have two or more distinct photos in the dataset. The only constraint on these faces is that they were detected by the Viola-Jones face detector.
3. Images of the students in the class will be collected using a webcam. The images will be taken in the classroom and will be used to train the model to recognize the students in the class.

The dataset used to test the model consists of photographs taken by group members during each class session. These photos were captured with the consent of the participants to record attendance for the day. Additionally, a small amount subset of a wild dataset will be incorporated as supplementary learning material.

Our approach involves training the model on a portion of the wild dataset to help it recognize and understand human faces. Subsequently, we will use our own images as test data to assess whether the model has been trained effectively. The primary objective is to determine the number of human faces in these images by performing face detection.

The dataset will be split into two parts: training and testing. The training dataset will be used to train the model to recognize the faces of the students in the class. The testing dataset will be used to test the model's accuracy.

The training dataset will consist of 80% of the images in the dataset. The testing dataset will consist of the remaining

20% of the images in the dataset.

asdsad

4. Methodology & Computer Vision Algorithm

4.1. Methodology

1. Train a model to recognize faces using the LFW dataset
2. Use the trained model to recognize the students in the class.
3. Use the Canvas API to fetch the number of students in the class.
4. Compare the number of students in the class to the number of students recognized by the model to get a percentage of students in the class.
5. Log the percentages and present them in a graph to provide a visual representation of the attendance in the class.

The model will be trained using the LFW dataset (Partially). The model will be used to recognize the students in the class. The model will be able to recognize the students in the class and log the time they entered and exited the classroom. The model will also be able to identify the number of students in an image and cross-reference it with Canvas's people table (number of students) to provide a percentage of students in the classroom. The percentages will be logged and presented in a graph to provide a visual representation of the attendance in the class.

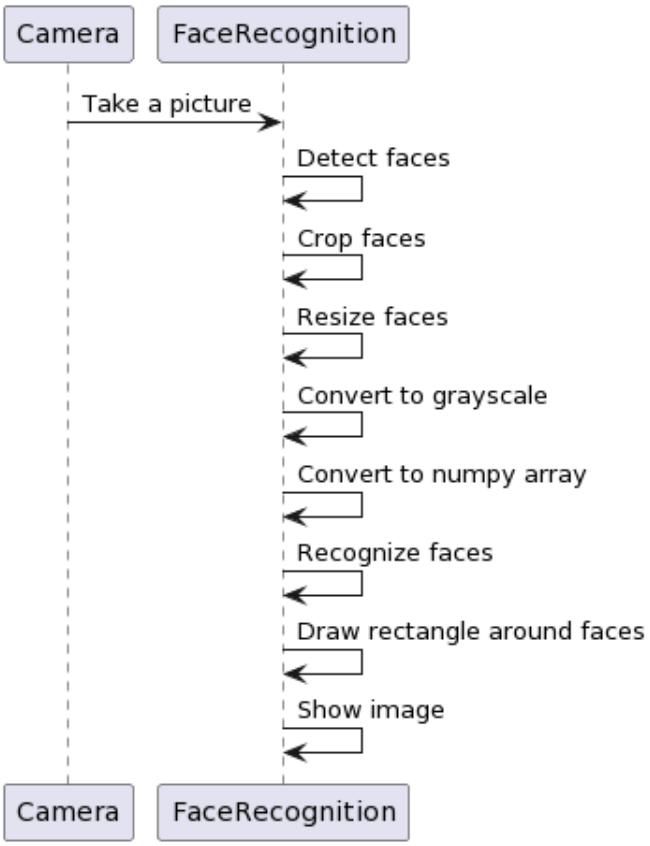


Figure 1. Flowchart of the system

4.2. Computer Vision Algorithm

Using YOLO [1], a neural network for object detection, we will train a model to recognize faces. The model will be trained using the LFW dataset. The model will be used to recognize the students in the class. The model will be able to recognize the students in the class and log the time they entered and exited the classroom. The model will also be able to identify the number of students in an image and cross-reference it with Canvas's people table (number of students) to provide a percentage of students in the classroom. The percentages will be logged and presented in a graph to provide a visual representation of the attendance in the class.

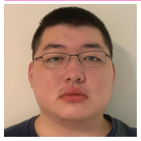
5. Team

Murtadha Marzouq

Resume Link

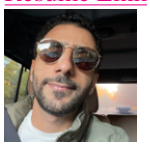
Param Patel

Resume Link



Haochen Ye

Resume Link



Sam Aldehayyat

Resume Link

Yuepei Yu

Resume Link

This team had been build with the following fellowship in minds:

1. Liberté, Egalité, Fraternité
2. Diversity and Inclusion
3. Teamwork
4. Respect
5. Integrity

6. Results

Prior to loading any images, we implemented a pre-processing step to stitch images to ensure that all the images are the same size. The images were stitched together using the OpenCV.

Figure 2. Pre-Processed Image

After the image is pre-processed, it is fed into the model. The image below shows the image before it is fed into the model.

Figure 3. Before

The image below shows the results of the model. The model was trained for 100 epochs and the loss was 0.0001. The model that was used was pre-trained on yolov6-n variant of the YOLO family.

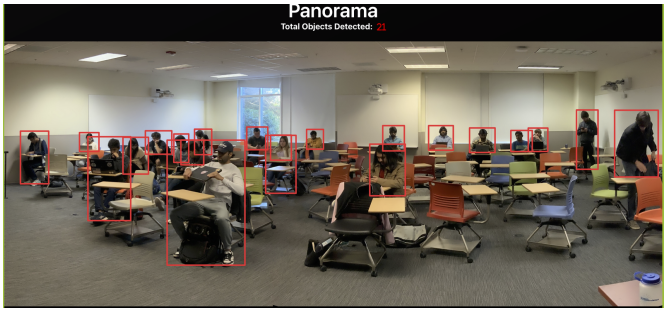


Figure 4. After

7. Conclusion

References

[1] Awais Adnan Misbah Ahmad, Imran Ahmed. Overhead view person detection using yolo. *IEEE*, pages 0627–33, 2019. 2