A Project Report on

Exam Hall And Classroom Surveillance using MODnT

Submitted in partial fulfillment of the requirements
for the award of the degree of

Bachelor of Engineering

in

Computer Engineering

by

Name of the Students
Kaushal Attal, Yash J. Jain, Darshan Jundare, Riddhi Chitre

Student ID's
18102015, 18102047, 12102030, 18102058

Under the Guidance of

Prof. Sachin Takmare
Prof. Mayuri Jain

Department of Computer Engineering
NBA Accredited
A.P. Shah Institute of Technology
G.B.Road,Kasarvadavli, Thane(W), Mumbai-400615
UNIVERSITY OF MUMBAI

Academic Year 2021-2022

Approval Sheet

This Project Report entitled ***"Exam Hall and Classroom Surveillance using MODnT"***
Submitted by ***"Kaushal Attal"(18102015) "Yash J. Jain"(18102047) "Darshan Jundare"(18102030) "Riddhi Chitre"(18102058)***is approved for the partial fulfillment of the requirement for the award of the degree of ***Bachelor of Engineering*** in ***Computer Engineering*** from ***University of Mumbai***.

Prof. Mayuri Jain
Co-Guide

Prof. Sachin Takmare
Guide

Prof. Sachin Malawe
Head, Department of Computer Engineering

Place:A.P.Shah Institute of Technology, Thane
Date:

CERTIFICATE

This is to certify that the project entitled ***"Exam Hall and Classroom Surveillance using MODnT"*** submitted by ***"Kaushal Attal"(18102015) "Yash J. Jain"(18102047) "Darshan Jundare"(18102030) "Riddhi Chitre"(18102058)*** for the partial fulfillment of the requirement for award of a degree ***Bachelor of Engineering*** in ***Computer Engineering***,to the University of Mumbai,is a bonafide work carried out during academic year 2021-2022.

Prof. Mayuri Jain
Co-Guide

Prof. Sachin Takmare
Guide

Prof. Sachin Malawe
Head Department of Computer Engineering

Dr. Uttam D.Kolekar
Principal

External Examiner(s)

1.

2.

Place:A.P.Shah Institute of Technology, Thane
Date:

Acknowledgement

I have great pleasure in presenting the report on **Exam Hall and Classroom Surveillance using MODnT**. I take this opportunity to express my sincere thanks towards my guide **Prof. Sachin Takmare** & Co-Guide **Prof. Mayuri Jain** Department of Computer Engineering, APSIT thane for providing the technical guidelines and suggestions regarding line of work. I would like to express my gratitude towards his constant encouragement, support and guidance through the development of project.

I thank **Prof. Sachin Malawe** Head of Department,Computer Engineering, APSIT for his encouragement during progress meeting and providing guidelines to write this report.

I thank **Prof. Amol Kalugade** BE project co-ordinator, Department of Computer Engineering, APSIT for being encouraging throughout the course and for guidance.

I also thank the entire staff of APSIT for their invaluable help rendered during the course of this work. I wish to express my deep gratitude towards all my colleagues of APSIT for their encouragement.

Student Names: Kaushal Attal, Yash J. Jain, Darshan Jundare, Riddhi Chitre

Student IDs: 18102015, 18102047, 18102030, 18102058

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

_____

(Signature)

_____

(Name of Student and Student ID)

Date:

Contents

List of Figures

List of Abbreviations

CCTV:	Closed-Circuit TeleVision
YOLO:	You Only Look Once
OpenCV:	Open ComputerVision
MODnT:	Multiple Object Detection and Tracking
GPU:	Graphics Processing Unit
CNN:	Convolutional Neural Network
EOD:	End Of Day
Recogn:	Recognition
MOD:	Multiple Object Detection
MOT:	Multiple Object Tracking

Chapter 1

Introduction

This report is based on the ideas, research conducted and information gathered, on how to utilize computer vision and deep learning for applications such as human surveillance inside a college premises, particularly for uses such as Exam Monitoring and Attendance inside Classroom monitoring.
At times inside our exam halls, we have heard that cheating is an offence for which we can be barred for years at a time, and that all of our actions inside the room are captured on video surveillance using CCTV.
This project aims to ensure that whenever there are cases of cheating happening inside a classroom authorities are given proper proof of the deed and the image(s) of the conspirators so that there is hard evidence.
This project aims to use YOLOv4 and Deep-Learning alongside an entire network of CCTV cameras, to bring the solution, a real time multiple object detection and tracking system that performs exam surveillance and also counts students present inside a classroom during the normal days to prevent any proxies from happening. (We have also kept in a future addition to this project post launch, and for others to add into, by making it a complete virtual attendance marking system).

Chapter 2

Project Concept

This chapter includes the following in its inclusive:

- Abstract
- Objectives
- Literature Review
- Problem Definition
- Scope
- Technology Stack
- Benefits for environment and society

2.1 Abstract:

In several applications of computer vision and image processing, the inception of the processing starts with object detection and subsequently tracking, if the need arises. In recent years, there has been extensive research in the field of object detection and tracking. Many remarkable algorithms have been developed for object detection and tracking, including color segmentation, edge tracking and many more. These algorithms and libraries include widely known YOLO , OpenCV, Kalman Filtering, Hungarian Algorithm, Darknet (an open source neural network) etc.

Our research is aimed at using combinations of above mentioned algorithms and libraries, moulded into one **Real Time Multiple Object Detection and Tracking (MODnT)**, for applications within our college to help reduce cases of cheating/copying in examinations and also to reduce/mitigate the opportunistic students from marking proxies for their peers, hereby creating a more disciplined environment inside college premises.

With these aforementioned algorithms and applications using real camera equipments being applied directly at server level access for the executeable file to directly tap into video feeds, monitor and assess ongoing situations, alert concerned individuals of higher authority within respective departments for any cases of exam malpractices using screenshots as evidence with concerning individuals conducting the malpractices inside exam premises.

2.2 Objectives:

- Exam invigilation:
  Detecting Scenarios Of Cheating in exams between two/more individuals.
  Send Real-Time Snapshot to concerned higher authority(HoD or Exam In-charge)

- Mitigating Attendance Proxies:
  Detecting number of students in the class vs actual count from register.
  Future Implementation Post Launch - Facial Recogn. To mark attendance virtually using facial recognition system against a student record database.

The above objectives are identified after several first hand experiences inside examination halls and classrooms by us students and our faculties.

2.3 Literature Review:

Object detection is one of the fundamental problems of computer vision. It forms the basis of many other downstream computer vision tasks, for example, instance segmentation, image captioning, object tracking, and more. We will be using Object Detection for Video Analytics.
We have decided to use a very specific use case application for Object Detection, ie. People Detection and counting. Deep learning networks and the computing power of GPU's, the performance of object detectors and trackers has greatly improved, achieving significant breakthroughs in object detection. Upon further research, it was deduced that, we are performing Object Detection and Tracking using Deep Learning as it is significantly more robust to occlusion, complex scenes, and challenging illumination. Though DL algorithms and methods require large amounts of training data, we have MS COCO dataset readily available.
Upon further researching which algorithms are currently the best when it comes to object detection, it was found that YOLOv4 (You Only Look Once v4) is the best object detection algorithm according to the MAP benchmark of MS COCO dataset.

2.4 Problem Definition:

Education Institutions have 2 major areas of interests/concerns such as:
Cheating/Copying during Examinations and;
Attendance Proxies by students.
These have been issues for the ages, that may finally be governed more effectively thanks to modern computer vision techniques using surveillance cameras(aka CCTV).
This project allows us to reduce the occurrences of these very namely incidents, promoting better levels of transparency, honesty and discipline within students as any malpractices we are aiming for within this project will implicate those students caught breaking the rules with proper evidential footage.
The solution we are proposing via this project, is utilization of College CCTV cameras and college central server to implement the solution within classrooms and labs to take a head start on our goal to reduce cheating and proxies.
We identify these cheating cases by creating vision cones for each student inside the exam hall, and if the vision cones are found to overlap/intersect each other because of two students sitting on consecutive benches looking at each others answer sheets, the system will create a snapshot of the identified students whose vision cones are overlapping and will send this snapshot to respective department exam in-charges to conduct further proceedings.

2.5 Scope:

This Real-Time Multiple Object Detection and Tracking project (RT-MODnT) has been taken up by my group to create a more disciplined learning environment inside college classrooms without creating any unnecessary disturbances/distractions inside the classroom during lecture and exam hours.
The project will primarily consist of a 3 phased manner:

- Phase 1:
  The first phase of implementation is to create the foundation of the project, create an object detection and tracking solution that works and identifies objects in still images. (currently underway)

- Phase 2:
  The second phase of this program is to upgrade the solution with the help of faculties and implement this as a pilot project within 1 classroom/computer lab based on permissions given by authorities, using live video feeds directly from CCTV cameras.

- Phase 3:
  The third and final phase will be to implement this pilot project across entire floors, including classrooms and labs, whilst also storing the recorded, detected and tracked videos to storage on the server or someplace where only authorities will have access.

2.6 Technology Stack:

We will be implementing this project with the following technologies, algorithms and libraries:

- Python3 - most adequate language for deep learning
- YOLOv4 (with an option to upgrade to v5 if it provides more functionality) + DeepSORT - Kalman Filtering + Hungarian Algorithm + CNN's

We also aim to keep open a possibility of using OpenCV along with a custom database to create an entire virtual attendance marking feature that may be taken up as a further addition to the project should anyone choose to continue the project.

2.7 Benefits For Environment and Society:

The benefits to society, which in our case are inclusive of the entire college ecosystem, students,faculties,administration etc:

- Reduced labour and error rates for teachers in larger capacity classrooms.
- Some students become more disciplined.

And for the environment, the college atmosphere becomes more transparent between students and faculties.

Chapter 3

Project Design

3.1 Proposed System Architecture:

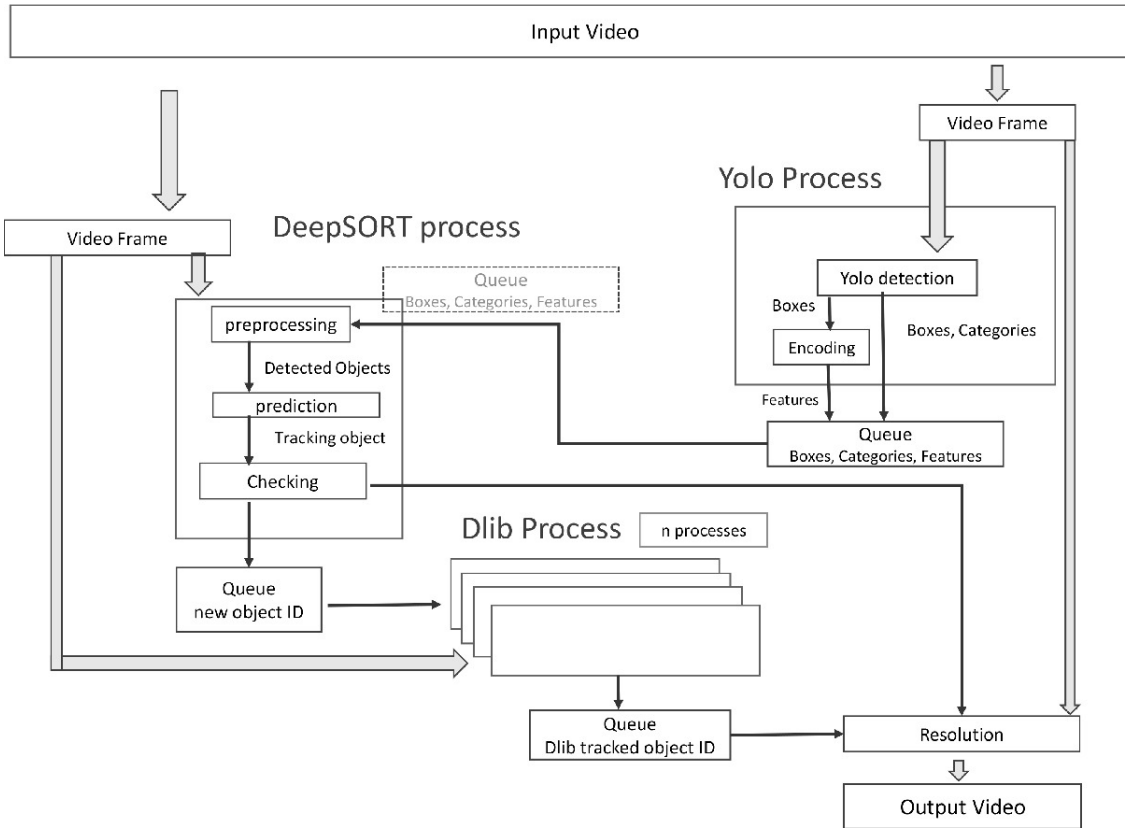

Figure 3.1: Real Time Multiple Object Detection and Tracking

The above figure demonstrates the proposed system architechture we have come up with as our solution, and it helps us maintain a consistency of processes and flows to maintain. The DeepSORT process end of the system will indulge in image/video frame preprocessing, detecting objects inside the image and tracking the objects in subsequent frames and images. The Yolo Process end of the system will indulge in creating identity boxes for the objects in the image/video frame and also provide an adequate feature set for the the DeepSORT process to be able to accurately identify and classify our targeted objects inside the frame.

3.2 Class Diagram:

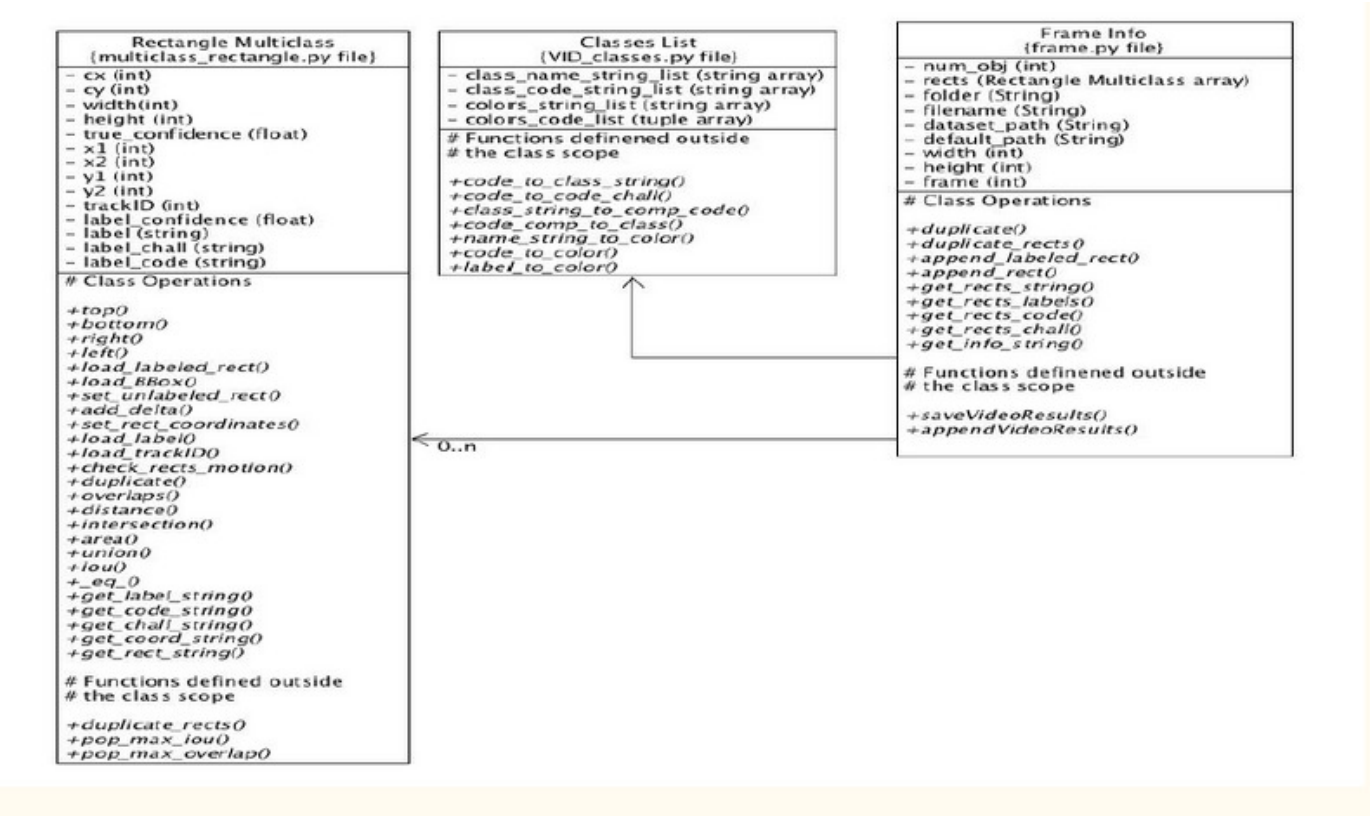

Figure 3.2: Class Diagram depicting all frame info, class operations, lists.

Though Filenames may differ whilst development the above class diagram depicts the relations between our image frame and the YOLO object recognizers.

3.3 Modules:

3.3.1 Module 1:

Our first module of the project/solution is create the basic MODnT program that works on the COCO datasets for training and the is used for still-image object detection, for the tracking part of the start the project we will feed in sample video already stored on system storage to test the functionality of the proposed solution.

3.3.2 Module 2:

Module 2 will involving using networking and using a personal camera eg,DSLR to produce live video feeds for the system to detect object and track our targeted objects in a way to confirm that the system is working. Inside Module 2 of the project we aim to add the functionality inside the tracking of an object to detect a field of view (view cones) threshold for each student inside the examination room/classroom as a healthy parameter to gauge whether or not we have objective cases on our hands.

3.3.3 Module 3:

Module 3 of the project involves upgrading the networking capabilities of the solution to be able to access, monitor and use it's MODnT application across multiple cameras within the same room. Alongside this we want to be able to create a more robust logic too ensure there are going to be very few cases of false negatives with the system so that there aren't any unnecessary alerts sent out.

3.3.4 Module 4:

The storage module, depending upon a days worth of footage that gets recorded, and analyzed we will decide the storage solution and time period before footage desruption to allow for memory overwrite, here admins will have access to ouutput footage folder and may at any time download the file while it still exists for further investigation.

3.4 Design - Flow of Modules:

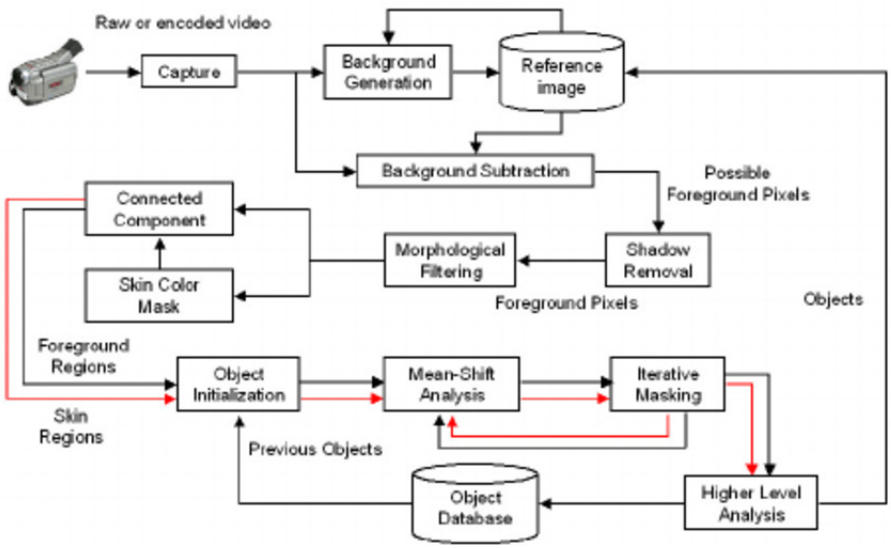

Figure 3.3: Multiple Object Tracking Data Flow Diagram

The above diagram demonstrates the processes that take place during the object tracking part of the algorithm. The background has to be subtracted in order to isolate the foreground object to be identified, boxed correctly and then tagged for tracking.
The mean shift algorithm can be used for visual tracking. The simplest such algorithm would create a confidence map in the new image based on the color histogram of the object in the previous image, and use mean shift to find the peak of a confidence map near the object's old position. The confidence map is a probability density function on the new image, assigning each pixel of the new image a probability, which is the probability of the pixel color occurring in the object in the previous image.

3.4.1 Use Case Diagram:

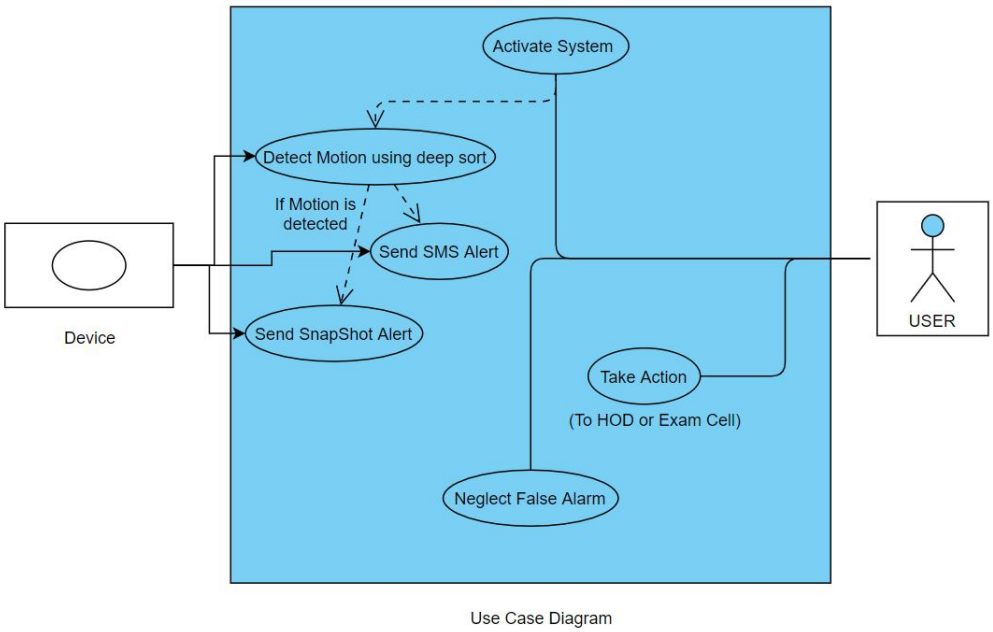

Figure 3.4: Use Case Diagram

In the above use case diagram, we are able to demonstrate which user and which equipment has what functionality present inside this solution, where the admins can access stored footage at EOD and also have the ability to alert HoD's or Exam-Incharge's of their respective department when coming across finds of our objectives.

3.4.2 Gantt Chart:

Gantt chart

Kaushal Attal | November 10, 2021

	January				February				March				April		
	Week 1	Week 2	Week 3	Week 4	Week 1	Week 2	Week 3	Week 4	Week 1	Week 2	Week 3	Week 4	Week 1	Week 2	Week 3
Phase 1	Development Of Phase 1 - Basic Backbone														
creating a basic backbone for the solution a multiple object detection and tracking software that works on still images and pre-recorded videos			Testing and Implmentation		Documentation of Phase 1 development Cycle										
Phase 2					Upgrading System and Development of system phase 2										
upgrading the system to work on a singular live camera and implement the project in a classroom/laboratory								Pilot Implementation, Testing		Documentation of phase 2 development cycle and results					
Phase 3										Upgradation from singular view to multi-view, development of system phase 3					
final upgrades to system for it to work multi view inside same room, and for it to work across multiple cameras at the same time across majority of college floors													System integration, implementation, trial runs and handover		
														Documentation of phase 3 development cycle, results, and project completion	

Figure 3.5: Gantt Chart - proposed project timeline

The Gantt chart as seen above is our proposed timeline for the Surveillance solution, with each phase, duration and month properly mentioned according to our goals.

3.4.3 Flowcharts:

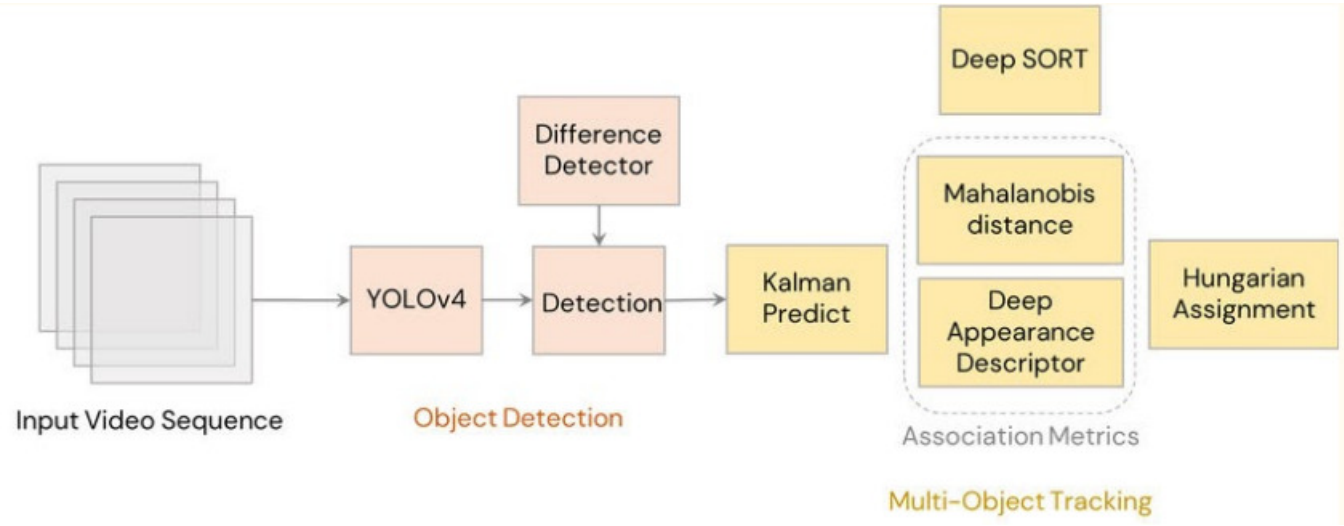

Figure 3.6: YOLOv4 Multiple Object Detection Block Diagram

The above flowchart is a diagrammatic representation of the end system we want to implement, with YOLOv4 working on a live video feed camera and giving output video. And seen above is a block diagram of the same application for a more comprehendable interpretation of how the flow of the YOLOv4 MODnT will take place.

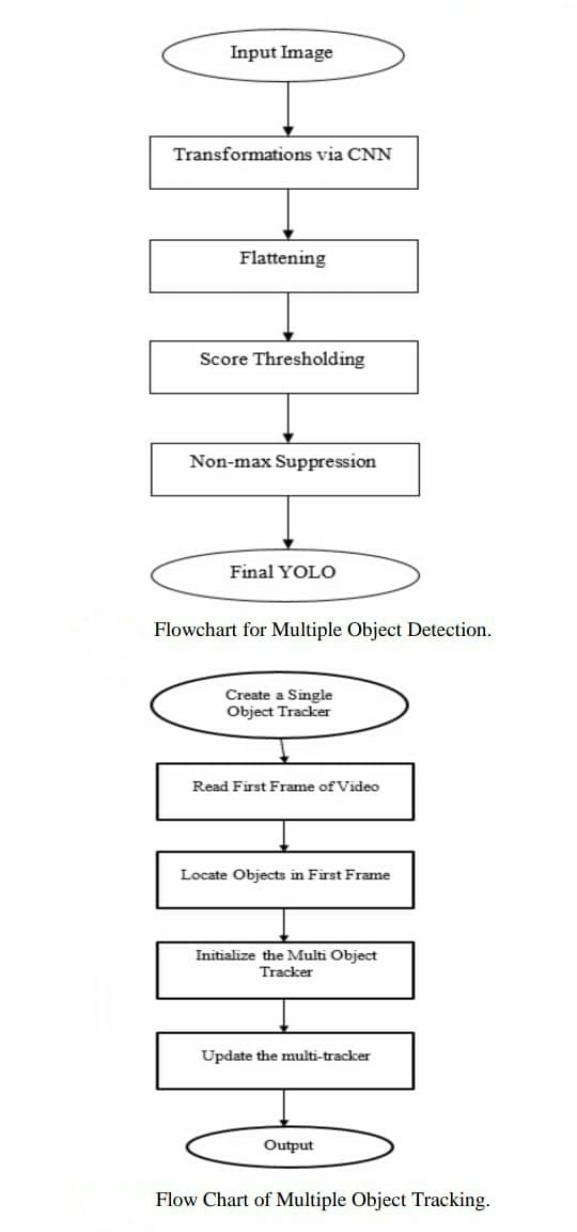

Figure 3.7: Individual flowcharts for MOD and MOT

Above present, are individual level flowcharts for both Multiple Object Detection part of the solution and Multiple Object Tracking part of the solution.

Chapter 4

Planning For Next Semester:

Our planning for the next semester involves conducting month-wise phased development. The first phase is already underway in which we are currently creating the basic backbone of the solution, ie, the Multiple Object Detection and Tracking. We aim to complete this phase as soon as possible and report it to our guides by mid-end of January '22.

The second phase of our development will commence once we have approval of first phase being completed successfully and then we will work on upgrading the system to work with live video camera feeds and then implement this project on pilot basis inside 1 classroom/laboratory. We expect this phase to be done by end of Feb '22.

The third phase will include upgrading the solution from single camera capacity to multi-camera capacity within same room and also across the network of CCTV camera's present inside our college. The phase should ideally by completed by End of March or at latest 2nd week of April so that we have adequate time remaining to focus on documentation.

Bibliography

[1] Object Detection from Video Sequences Using Deep Learning: An Overview- Dweepna Garg and Ketan Kotecha.

[2] Mitchell, T.M.: Machine Learning, p. 421. McGraw-Hill Science/Engineering/Math (1997)

[3] Chen, Y., Yang, X., Zhong, B., Pan, S., Chen, D., Zhang, H.: CNNTracker: online discriminative object tracking via deep convolutional neural network. Appl. Soft Comput. 38, 1088–1098 (2015) [Elsevier]

[4] LeCun, Y., Bengio, Y., Hinton, G.: Deep learning.

[5] Deep Learning-based Multiple Pedestrians Detection Tracking Framework-Xuan-Phung Huynh Dept. of Computer Engineering Sejong University, Yong-Guk Kim* Dept. of Computer Engineering Sejong University

[6] https://www.viso.ai/deep-learning - Object Detection in 2021

[7] https://www.pyimagesearch.com - multiple references for projects and technologies