YANGON TECHNOLOGICAL UNIVERSITY
Department of Mechatronic Engineering

Customized Deep Learning Models for Artificial Intelligence Surveillance

Defense Seminar

25-5-2020
Monday

Supervised by

Dr. Phyu Phyu Htun

Presented by

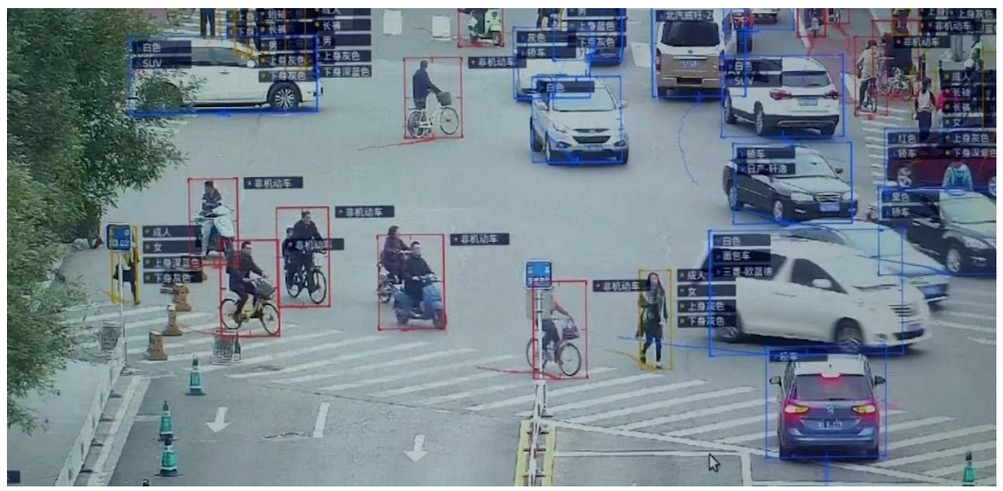
Aw Thura (VI-McE-11)

Outline

- Introduction
- Problem Statement
- Literature Review
- Objective and Goals
- System Overview
- Thesis Roadmap
- Timeline and Milestones

Introduction

- A.I. Surveillance is increasingly being used to monitor public and private spaces throughout the world
- Example applications include
  - Smart road surveillance
  - Facial recognition systems
  - Smart policing
- Facial recognition system is integrated successfully in immigration check points and offices in Myanmar
- The potential to implement smart road and security surveillance systems



Problem Statement

- Surveillance cameras are monitored manually and the officers usually have many cameras to cover.
- Most CCTVs are left idle and it makes CCTV reviewing a time consuming process.
- The recorded data is also vulnerable if stored locally.

Literature Review

- J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once : Unified, real-time object detection. (Redmond et al 2015)
- J. Redmon and A. Farhadi. Yolo9000: Better, faster, stronger. (2016)
- J. Redmon Ali Farhadi YOLOv3: An Incremental Improvement (2018)
- A. Bochkovskiy C.Y. Wang H.Y.M. Liao YOLOv4: Optimal Speed and Accuracy of Object Detection (2020 April)

Literature Review

Model	Features & Capabilities	Limitations
Original YOLO (2015)	• 24 convolutional layers, Divides the image into 7x7 grid cells. • Can perform classification and localization on all 49 grid cells simultaneously. **mAP** score is **63.4%** at with frame rate of 45 FPS.	• Each grid cell can only have one object, Thus, YOLO can detect only 49 max objects. • Bad for small objects that appear in groups • Relatively high localization error.
YOLO v2 (2017)	• 19 convolution layers, Detects over 9000 classes using Hierarchical classification • Uses anchor boxes, Added batch normalization on all convolutional layers, **2% improvement** in mAP	• Computationally expensive • Still struggles with small clusters

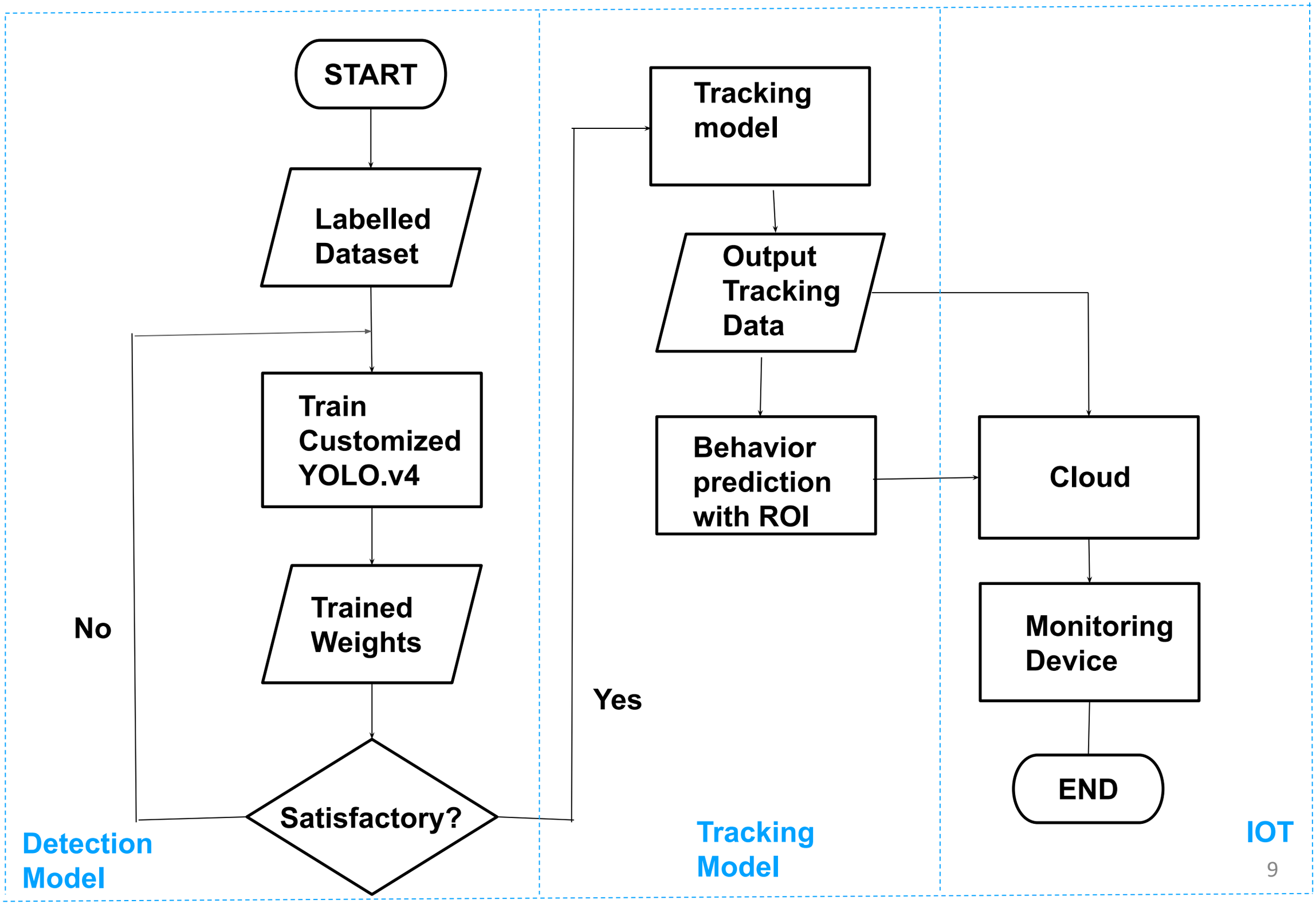
Literature Review

Model	Features & Capabilities	Limitations
YOLO v3 (2018)	• 53 convolutional layers, enables multilabel prediction. • Uses shortcut connecters to achieve better performance on small objects	• Worse performance on medium and larger size objects compared to YOLO 9000 • Sometimes struggles to get the boxes perfectly aligned with the object
YOLO v4 (2020 April)	• 53 conv layers, state of the art results • Data augmentation, which increases the generalization ability of the model • Uses Spatial Pyramid Pooling to maintain aspect ratio of the image while pooling, results in better accuracy	• None

Objectives & Goals

- To develop and experiment customized YOLO v4 models that can detect various classes such as pedestrians, types of vehicles, and animals to track them.
- To develop a tracking model with deepSORT that can track movement of the detected objects and to identify suspicious behavior.
- To implement the developed models for real time surveillance.
- To store the detected data on cloud and monitor the data via network.

System Overview



Thesis Roadmap

Detection Model → Tracking Models → IoT integration

- Developing appropriate model (custom trained YOLO v4)
- Gathering datasets for training
- Training the model
- Testing and evaluation

- Developing appropriate model (most likely deepSORT)
- Perform behavior prediction based on Region of Interest
- Generate tracking output data

- Making the data real time compatible with a device
- Data storage in cloud
- Live monitoring via the network

Timeline and Milestones

Duration	Tasks
May 1, 2020 - June 30, 2020	**Developing and Testing Detection models**
July 1, 2020 - July 31, 2020	**Developing and Testing Tracking models**
Aug 1, 2020 - Aug 31, 2020	**Behavior Prediction**
Sep 1, 2020 - Sep 31, 2020	**IoT integration**

End

Thank You For Your Attention.