Autonomous Intelligence-based Security & Surveillance System (AISS)

by:

Faizan Ahmad

Muhammad Hammad

Umar Nawaz

M. Momin Abbas

Supervised by:

Col. Dr. Imran Makhdoom

Submitted to the faculty of Department of Electrical Engineering,
Military College of Signals, National University of Sciences and Technology, Islamabad,
in partial fulfillment for the requirements of B.E Degree in Electrical (Telecom) Engineering.

June 2023

In the name of Allah,

the Most Benevolent, the Most Courteous

CERTIFICATE OF CORRECTNESS AND APPROVAL

This is to officially state that the thesis work contained in this report

"Autonomous Intelligence-based

Security and Surveillance System"

is carried out by

Faizan Ahmad, Muhammad Hammad, Umar Nawaz, and M. Momin Abbas

under my supervision and that in my judgement, it is fully ample, in scope and excellence, for the degree of Bachelor of Electrical (Telecom.) Engineering in Military College of Signals, National University of Sciences and Technology (NUST), Islamabad.

Approved by

____

Supervisor

(Col. Dr. Imran Makhdoom)

Department of EE, MCS

Date: __________________

DECLARATION OF ORIGINALITY

We hereby declare that no portion of work presented in this thesis has been submitted in support of another award or qualification in either this institute or anywhere else.

ACKNOWLEDGEMENTS

Allah the Almighty is the sole guidance in all domains and to whom we all are eternally and humbly grateful.

Our beloved parents, colleagues, and most of all our supervisor, Col. Dr. Imran Makhdoom, whose guidance was pivotal in the implementation of this project.

Last, but not the least, the group members worked steadfastly through all adversities and made the experience beneficial and memorable.

Plagiarism Certificate (Turnitin Report)

This thesis has ____ similarity index. The Turnitin report endorsed by the Supervisor is attached.

Faizan Ahmad

NUST Serial no. 325357

Muhammad Hammad

NUST Serial no. 325400

Umar Nawaz

NUST Serial no. 325402

M. Momin Abbas

NUST Serial no. 325362

Signature of Supervisor

ABSTRACT

Creating a vigilant and reliable surveillance system is foremost necessary for ensuring our safety for ourselves at our residences and workplace. As we are observing the significant and consequential developments in the fields of computer vision and machine learning, it is fair and wise to employ emerging technologies in our surveillance and security systems. Such techniques will enhance the system's efficiency significantly. Therefore, we are proposing an **Autonomous Intelligence-based Security and Surveillance System (AISS)** that can autonomously detect perimeter intrusion over the boundary walls of an institution. The proposed solution, AISS, will autonomously and promptly raise the alarm and notify the security staff upon detecting a breach over the boundary. The primary objective of our project is to eliminate the likelihood of human error, which can occur due to fatigue, stress, and various other reasons.

Table of Contents

List of Figures

4

Chapter 1: Introduction

Science and engineering are concurrent when it comes to the advancement of human life. A scientist must apply the knowledge to formulate better solutions to our existing problems through research and experiments. Whereas an engineer must apply the skills to materialize the solutions for the public and ensure those are reliable and sustainable.

Throughout history, we have been able to achieve remarkable progress due to the ingenuity and finesse of engineering. The application and realization of apparently arduous scientific solutions have been the principal objective for engineers. By delivering the job, we have seen a radical change in people's productivity. Automation is a feat of engineering, and this has already shown its brilliance in improving people's lives. Where long manual labour isn’t feasible and takes a toll on human effectiveness, automation has been the best solution. Equipped with machine learning algorithms, it can now perform at par with human excellence. Such solutions are instrumental in solving problems promptly and efficiently. Also, it creates more time for humans to perform their tasks.

As many tasks require high performance throughout time, human performance can't keep up due to fatigue, and this causes a toll on his health, too. Constant surveillance is a demanding job upon which we cannot comprise for safety concerns. By using such automated systems, we can increase the productivity of our security personnel. The personnel can utilize their time in routine patrolling and inspection. Furthermore, it will enhance the personnel’s response in case of a breach.

1.1 Overview

As time progresses, we are realizing the expanding potential of automated systems. Not only it has revolutionized production, but it can also benefit us in such tasks that require high concentration, such as surveillance and detection. Additionally, society is expressing an increased interest in security and safety. This has resulted in the exploration of deploying intelligent and smart CCTV systems. These systems would be able to detect and identify anomalous behaviours. This can aid us in giving better and more effective responses to the situations being developed upon detecting anomalous behaviours. Surveilling numerous CCTV cameras installed on the boundary walls of a building and detecting an anomaly is an exhaustive and intensive job. Naturally, the personnel assigned to monitor the cameras would feel lethargic and be easily distracted by something else. Studies indicate the relation between longer working hours and the degradation of personnel's performance. The personnel performs poorly in his job and at multiple times would not be effectively monitoring the cameras due to monotony [1]. This can give the malicious actors a window of opportunity to breach the boundary and commit crimes when the employee would be looking elsewhere or not focusing on a particular camera.

We must utilize an efficient mechanism to improve the quality of labour. Hence, to satisfy the employers' requirements and reduce the staff's workload, a machine learning algorithm can be deployed with a CCTV system that can autonomously detect a breach and notify the security personnel accordingly. Consequently, cutting down the need for constant monitoring and the resulting fatigue.

1.2 Problem Statement

Creating reliable and vigilant surveillance systems is necessary for ensuring safety. Currently, the systems installed require manual monitoring of the CCTV feed. Subsequently, the likelihood of human error increases due to fatigue. The following are the problems of existing surveillance systems:

- The constant monitoring of multiple cameras is tiresome.
- The staff could be distracted due to monotony.
- The staff must keep on monitoring diligently to notify if the breach occurs, which can increase fatigue and the likelihood of error.
- Miscommunication can occur as the staff hastens to notify a developing situation.
- The existing sensor-based systems are expensive and require high maintenance of the hardware.
- The sensor-based systems are susceptible to interference by external factors. For instance, thunderstorms, electromagnetic disturbances, et cetera.

1.3 Proposed Solution

The primary goal of this project is to eliminate human errors and automate surveillance. In this project, the live feed from the cameras would be passed from the machine learning algorithm and autonomously detect a breach over the boundary. This would improve employee productivity and provide an affordable surveillance system. The security staff can identify a breach in minimal time with the aid of AISS. The solution is quite simple to install and requires minimal maintenance. An interactive user interface is also provided to facilitate this process.

1.4 Objectives

1.4.1 General Objectives

To create a sophisticated system built upon Machine Learning (ML) techniques that reduce manual labour and provide an affordable, easy-to-integrate, and efficient method of surveillance system.

- Development of an autonomous surveillance system.
- Implement machine learning techniques to improve results.
- Increase the productivity of the team.
- Contributing to improving work efficiency.

1.4.2 Academic Objectives

- To develop a system that can accurately detect anomalous behaviour in real-time, such as a person climbing a ladder, using advanced computer vision techniques.
- To investigate the use of machine learning algorithms to improve the accuracy of object detection and reduce false positives.
- To optimize the system's performance by exploring different techniques and processing methods.
- To evaluate the system's effectiveness in different environments and scenarios, such as indoor or outdoor settings, day or night conditions, and different weather conditions.
- To explore the ethical and legal implications of autonomous surveillance systems and propose guidelines for their responsible use.
- To contribute to the development of open-source tools and resources for autonomous surveillance systems implemented in OpenCV, such as tutorials, code samples, and libraries.

1.5 Working Principle

The project primarily operates on object detection and activity recognition embedded with machine learning algorithm. Upon detection of anomalous activity, it will generate an alarm and send e-mail notifications to the concerned staff. The project is completed in multiple modules – each module is a precursor for the next one. The modules of this project are as follows:

- Custom Dataset Generation
- Dataset Annotations
- Preparing the best Algorithm
- Training the Algorithm
- Activity-based Decision
- GUI Desktop Application
- Alarm Generation and Notification

1.5.1 Dataset Collection and Generation

This is the primary part of the project, as we have to detect suspicious/anomalous behaviour and raise an alarm about that. We have used existing and custom-made datasets for this purpose. It includes images of anomalous and non-anomalous behaviours, i.e., to differentiate between the behaviours and detect them accurately.

1.5.1.1 COCO Dataset

This is the ***Common Objects in Context*** (COCO) dataset that include everyday objects in our lives. To annotate the images, we used Roboflow to aid the algorithm to detect the classes.

1.5.1.2 Custom Dataset

For the detection of anomalous behaviour, we have prepared a custom dataset where a person brings a ladder and climbs upon it. Our dataset is specifically designed for recognizing the climbing gesture of the person upon which breach will be detected.

1.5.2 Training the Algorithm

For creating and training an algorithm on our datasets, we have used the **YOLOv5** algorithm. ***You Only Look Once*** (YOLO) computer-vision models trained on COCO datasets, which can be used in real-time. Object detection is done as a regression problem. We trained the model by specifying the epochs, batch size, weights, and et cetera.

1.5.3 Activity-based Decision

The decision-making criterion for detecting anomalous behaviour is done by setting up a restricted zone on the camera feeds. By defining such zones for different camera feeds, the system will detect a breach upon a person entering that zone. Upon detection, a bounding box will appear to facilitate viewing the intruder.

1.5.3.1 Person and ladder detection decision

The algorithm detects a person and ladder separately, which is important for detecting them both as a breach.

1.5.3.2 Breach detection decision

Although the algorithm will detect the person and the ladder, it would detect it as a breach only when the person starts climbing and enter the defined restricted zone.

1.5.4 GUI Desktop Application

The graphical user interface (GUI) is developed using **PyQt5** and **OpenCV**. PyQt5 is among the popular choices opted by developers for developing GUI applications using Python. It provides widgets and tools for creating modern and interactive graphical user interfaces. Moreover, PyQt5 also provides a range of tools for working with multimedia, such as video files, etc. This makes PyQt5 the best option for developing our project's graphical user interface.

The GUI serves the purpose of viewing and managing the CCTV feeds installed at the boundary walls. In addition to that, when an anomaly occurs in any feed, that camera screen will pop up for ease of viewing the said action.

1.5.5 Alarm Generation and Notification

For raising an alarm, a restricted zone is defined to detect perimeter intrusion. As an intruding person enters the defined zone. The siren is attached to the control room and will be activated upon breach. This module is designed for facilitating the security personnel to duly encounter the intruder. To notify concerned authorities, a notification will be sent with the image of the intruder.

1.6 Scope

This project is suitable for border security, critical infrastructure protection, (i.e., law-enforcement agencies), hospitals, educational institutes, and city public areas. This is a sophisticated system using ML and image processing techniques to detect and notify any anomalous behaviour (wall breach in this case) and autonomously monitors the camera feeds. This can also be customized for different scenarios of an anomaly to satisfy the users' requirements.

1.7 Deliverables

1.7.1 Anomaly Detection Algorithm

The anomaly detection algorithm serves as the autonomous surveillance and decision-maker. It will apply its algorithm to different environments and detect an anomaly and raise an alarm to notify the security personnel.

1.7.2 Desktop Application

This will be the graphical user interface for the security personnel to monitor the CCTV feed themselves and receive notification of an anomaly occurrence. It will also give the image of the breach to all concerned personnel.

1.8 Relevant Sustainable Development Goals

The socio-economic issues of security and expensive technical systems are addressed. The UN SDGs 9, 11, 16, and 17 are related to our project, which is as follows:

- **SDG 9:** Industry, Innovation and Infrastructure
- **SDG 11:** Sustainable Cities and Communities
- **SDG 16:** Peace, Justice and Strong Institutions
- **SDG 17:** Partnership for the Goals

1.9 Structure of Thesis

- Chapter 2 contains the literature review and the background this thesis is based upon.
- Chapter 3 has the details of the design and development of this project.
- Chapter 4 introduces a detailed analysis and evaluation of the algorithm.
- Chapter 5 consists of the conclusion of this project.
- Chapter 6 highlights the future work required for the commercialization of this project.

Chapter 2: Literature Review

With the progression of time, existing solutions are updated and enhanced to overcome the issues faced during the operation of the solutions. These updates and enhancements are done upon the occurrence of a new idea or development. For the development of such processes, a comprehensive study is vital to identify the unresolved problems our research can address and formulate the framework to achieve that solution. Our research underwent the following stages:

- Industrial Background
- Existing solutions and their drawbacks
- Proposed solution

2.1 Industrial background

The processes that are done manually cost time and efficiency of the workforce. Therefore, productivity drops and employees are susceptible to making errors upon more workload [1]. As highlighted in the Problem Statement, we aim to automate an intensive task and diminish the likelihood of human error.

To begin with, we are aware of low efficiency because of strained manual labour. We know that it not only costs the employers their expenses, but the employees are also physically strained and, subsequently, they face the stress of performing poorly at their job. This drives productivity down into the ground. Unfortunately, Pakistan falls behind the world in productivity. With the availability of cheap workforce, we often discard the idea of applying automation to enhance our work and provide relief in employee workload. With the application of Machine Learning (ML) techniques, we can aid in the provision of better workplace standard operating procedures (SOPs) that can execute their jobs effectively and improve the workplace environment for the employees.

2.2 Existing solutions and their drawbacks

Numerous solutions are in existence that are solving the said issue. Each solution has its own pros and cons; each with its own uniqueness. Following are the solutions which are applied in such systems:

- Fibre optic-based solutions
- Buried and mounted seismic motion sensor-based solutions

2.2.1 Fibre optic-based solutions

Such systems have a fibre optic deployed along the boundary wall or fence. It sends pulses of laser light across the optic line. It detects disturbance upon the minute changes in light's reflection. [2] Although a good solution, it is quite extensive and expensive to install it over the boundary. It requires regular maintenance, which costs resources during its operation.

2.2.2 Seismic motion sensor-based solutions

Similar to the above-mentioned systems, this is more extensive to install as the user would need to dig the ground for installing sensors. [3] In addition to maintenance costs, the installation costs are quite higher. The operation of such systems is complicated and requires a technician with specific knowledge to maintain them.

Chapter 3: Design and Development

3.1 YOLOv5 Algorithm

You Only Look Once (YOLO) is a state-of-the-art object detection algorithm developed by Ultralytics. YOLOv5 is one of the latest iterations of this family of algorithms, which incorporates numerous improvements including speed, accuracy, and efficiency. The key improvement is the use of data augmentation techniques — namely, mosaic data augmentation. This process combines multiple images into a single image and randomly crop and augments image results. This version has several optimizations improving the model's speed and efficiency. It has a lightweight backbone network that greatly reduces computational costs. Further on, YOLOv5 has multiple variants. These variants operate in similar methods but differ in the number of parameters and layers.

For the detection of a person, a ladder, and then a breach, we have used the YOLOv5s, the *small* variant. This algorithm is a suitable option as it does real-time object detection on limited computational resources. Despite having a lighter network, higher accuracy is achieved. The algorithm is a quick and efficient method to detect objects. It functions as a single-stage object detector, whose architecture comprises of three components: *backbone*, *head*, and *neck*. The variants of YOLOv5 operate in similar method and only differ in the number of parameters and layers.

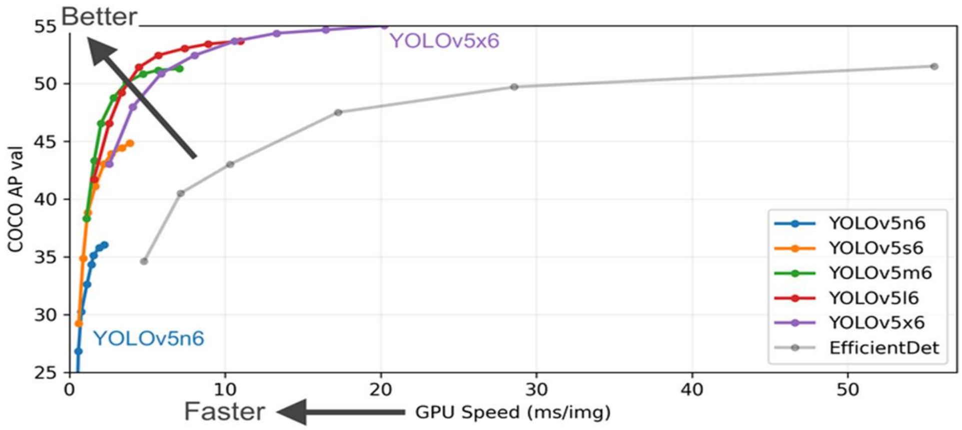


Figure 1: Performances of YOLOv5 variants

Model	size (pixels)	mAP^{val} 0.5:0.95	mAP^{val} 0.5	Speed CPU b1 (ms)	Speed V100 b1 (ms)	Speed V100 b32 (ms)	params (M)	FLOPs @640 (B)
YOLOv5n	640	28.0	45.7	**45**	**6.3**	**0.6**	**1.9**	**4.5**
YOLOv5s	640	37.4	56.8	98	6.4	0.9	7.2	16.5
YOLOv5m	640	45.4	64.1	224	8.2	1.7	21.2	49.0
YOLOv5l	640	49.0	67.3	430	10.1	2.7	46.5	109.1
YOLOv5x	640	50.7	68.9	766	12.1	4.8	86.7	205.7
YOLOv5n6	1280	36.0	54.4	153	8.1	2.1	3.2	4.6
YOLOv5s6	1280	44.8	63.7	385	8.2	3.6	12.6	16.8
YOLOv5m6	1280	51.3	69.3	887	11.1	6.8	35.7	50.0
YOLOv5l6	1280	53.7	71.3	1784	15.8	10.5	76.8	111.4
YOLOv5x6 + TTA	1280 1536	55.0 **55.8**	72.7 **72.7**	3136 -	26.2 -	19.4 -	140.7 -	209.8 -

Figure 2: Metrics of YOLOv5 variants
(The one used in our project is highlighted in yellow)

3.1.1 Algorithm Architecture

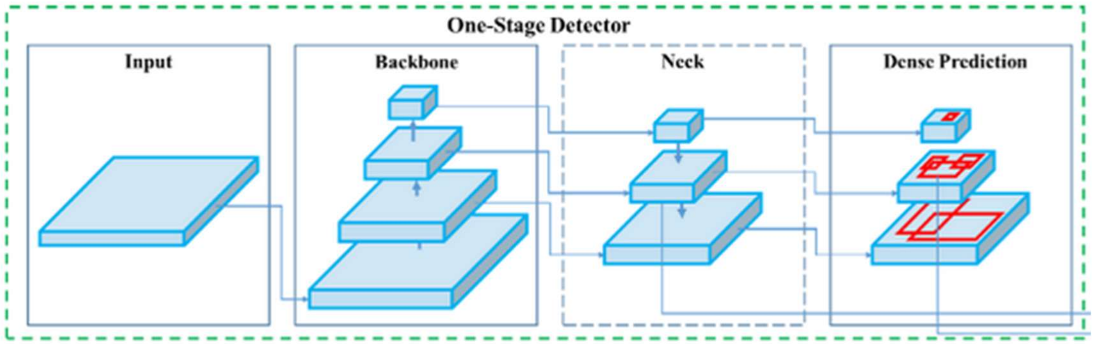


Figure 3: Single-stage detector's architecture

First, in the backbone, **Darknet53** is a convolutional network upon which a **Cross Stage Partial** (CSP) network strategy is applied. In the CSP network, the excessive amount of redundant gradient information is reduced as it removes the gradient flow. CSPNet is employed to segregate base layer's feature map into two parts. Afterwards, those are merged via a cross-stage hierarchy. This process lowers the number of parameters; therefore, the computational cost is reduced and increases the inference speed – significant for our real-time operation.

Then, in the neck, **Spatial Pyramid Pooling** (SPP) and **Path Aggregation Network** (PANet) are applied. The SPP performs aggregation (i.e., accumulates) of the received input information and returns a fixed length output for that. Primarily, it separates the important features while maintaining the network's speed. On the other hand, PANet performs these two important tasks: (a) improves the flow of information, and (b) the localisation of pixels for mask prediction.

Finally, in the head, three convolutional layers are used to predict the location of the bounding boxes (*x, y, height, width*), prediction scores, and objects' classes. Following are the equations used for this purpose in the YOLOv5 algorithm:

$$b_x = (2 \cdot \sigma(t_x) - 0.5) + c_x \qquad b_w = p_w \cdot (2 \cdot \sigma(t_w))^2$$
$$b_y = (2 \cdot \sigma(t_y) - 0.5) + c_y \qquad b_h = p_h \cdot (2 \cdot \sigma(t_h))^2$$

3.1.2 Hyperparameters for the Model

The hyperparameters set for the model are as follows:

```
IMG_SIZE = 640            BATCH_SIZE = 64            EPOCHS = 20

MAX_SEQ_LENGTH = 50       NUM_FEATURES = 2048
```

The input size (i.e., the frame size of the videos passed) for training, validation, and testing for the model was set at 640×640 pixels. The algorithm was processing 64 images at a time. The training data was passed through the algorithm 20 times. The maximum input length is set at 50; i.e., any input longer than 50 tokens was truncated. The algorithm extracted 2,048 features from the input dataset images.

By applying these hyperparameters, we were able to achieve sufficient and effective results for object detection. These parameters were optimal for getting desired results.

3.2 Dataset and Detection

For this step, we initially prepared the custom dataset for our project and then proceeded to train the model on our dataset. For dataset preparation, we had to collect images of the person setting the ladder on various walls and then climbing on it. The making of the custom dataset is quite beneficial since we make the model learn the specific types of objects (ladder and breach in our case).

3.2.1 Preparation of Dataset

For this task, we made our custom dataset. Especially for the breach by climbing a ladder, as we had to train the model on specific actions. We prepared around 6,000 images by climbing the ladder at different locations around our campus' boundary walls. It includes both, anomalous and non-anomalous images for training the model to differentiate and recognize the required action. The COCO datasets were also used to train the person detection. Hence, we got the availability to create ladder dataset.

Figure 4: Anomalous activity dataset image

Figure 5: Non-anomalous activity dataset image

Afterwards, we annotated the person and ladder in these images using Roboflow's annotation tool. For a person, we used bounding boxes. Whereas, for the ladder, we used polygon annotation as we require an accurate detection of a ladder due to its shape. Later, augmentation of the dataset was done to improve the accuracy of our model. In this process, we cropped the images and adjusted their brightness and contrast. To pass the dataset on the model, the dataset was split into training, validation, and testing sets. The training set is used for teaching the model, the validation is to fine-tune the model and select perfect hyperparameters, and finally, the testing is for the evaluation of the model. The training of our detection model is done using deep learning frameworks, such as TensorFlow, PyTorch, and Keras.

3.2.2 Person and Ladder Detection

We prepared over 1,500 images for anomalous behaviour which included a person and a ladder. Further on, we annotated the person and the ladder in these images using Roboflow. For person and ladder each, the dataset images were split into training, validation, and testing sets for our model. Hence, the model will consider these two as a class and detect them separately. The dataset for the ladder is necessary for training the model to detect the climbing gesture upon which it will only detect the person doing an anomalous activity.

Figure 6: Person and ladder detection

3.2.3 Breach Detection

As mentioned above, the model to detects the person and ladder separately. This is to differentiate between anomalous and non-anomalous behaviour. There may be a scenario where a person is taking a ladder elsewhere. Therefore, we annotated the climbing gesture as a breach to train the model better for detecting it. In addition to that, we have also set a logical zone along the boundary wall which serves as a checkpoint. Here, as the person starts climbing the ladder and crosses the logical zone, the breach will be detected and an alarm will be generated, too.

Figure 7: Breach detection

3.3 GUI Desktop Application

To execute an autonomous surveillance system, an efficient and user-friendly application compliments the detection model. The application fulfils the requirement of managing the CCTV feeds. The security staff can easily view and monitor certain feeds as well. The design and development of the GUI are done using primarily **PyQt5** and **OpenCV**.

The GUI will autonomously alert in case of a perimeter breach. It will mark a ***breached area*** with a display box around the person breaching the perimeter. This will grasp the attention of the security staff present in the control room.

3.3.1 Interface

As the application executes, a page consisting of sixteen cameras' feeds are displayed. At the bottom of this page, the application provides the functions given to the user. At the bottom, there are functions of **STATUS**, which shows the camera's links/port numbers, and **REFRESH**, which refreshes the camera feeds. In addition to that, as the model is being executed at the backend, when an anomalous action is detected at any camera, that camera's feed will pop up in front of the user for his ease to monitor the situation as well.

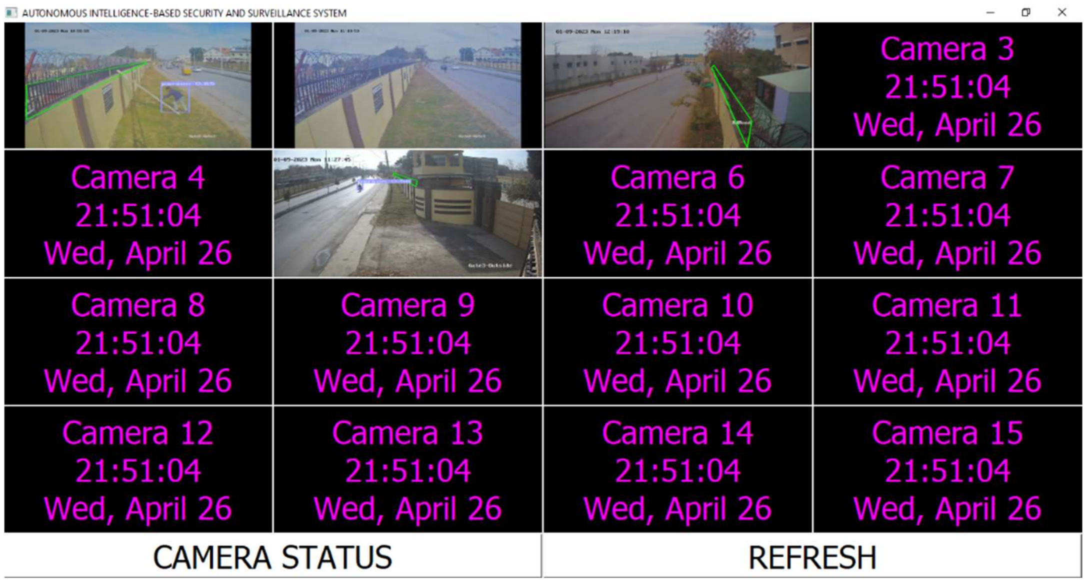
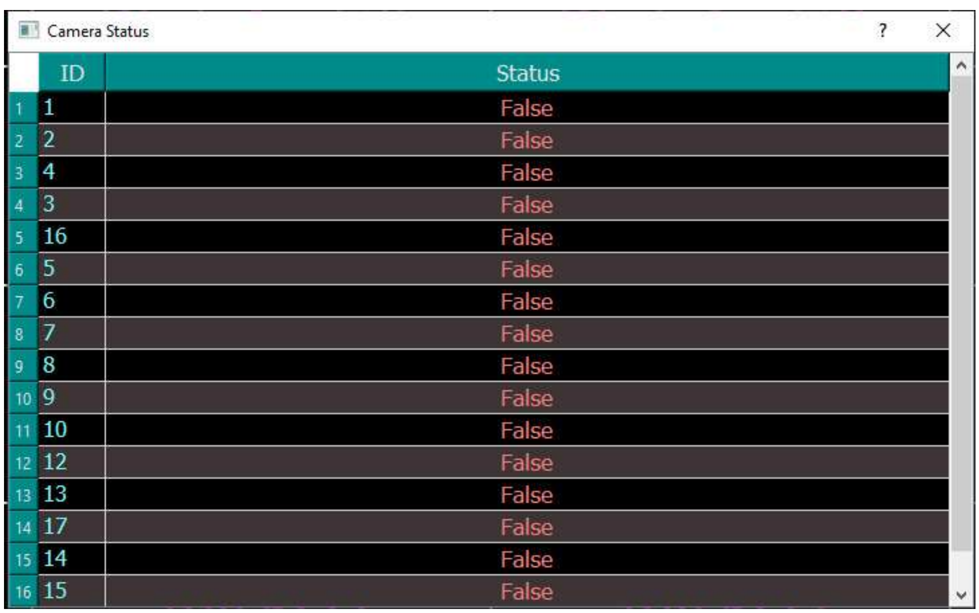


Figure 8: Graphical interface of AISS


Camera Status

	ID	Status
1	1	False
2	2	False
3	4	False
4	3	False
5	16	False
6	5	False
7	6	False
8	7	False
9	8	False
10	9	False
11	10	False
12	12	False
13	13	False
14	17	False
15	14	False
16	15	False

Figure 9: Status check for the cameras' link
(Before adding camera feeds)

3.3.2 Alarm Generation

In this module, the system will be monitoring the camera feed. As the person is climbing the ladder and reaches the top of the boundary wall, the system will raise the alarm siren attached to the control room and other designated offices. This is to grasp the attention of the security personnel and put them on alert. By this, we can facilitate the security personnel to put up a prompt response to deal with the intruders. When a person enters a ***no-man zone*** marked on our system, an alarm will be generated. Upon that, a bounding box will appear around the intruding person and remain on him until he exits the camera's view.

3.3.3 Notification

Like the alarm generation, this module is designed for facilitating the security personnel to duly encounter the intruder. In a notification, the security personnel and other concerned staff will receive an alert with the image of the intruder attached. This includes information such as the image of the intruder and the location of the breach. This is a beneficial step as there will be no haste in notifying the personnel regarding the developing situation. This reduces the risk of miscommunication, and the personnel can promptly converge on the area of intrusion. Therefore, the administration can devise better and sound SOPs regarding security situations.

Chapter 4: Code Analysis and Evaluation

As mentioned in the previous chapter, YOLOv5s is a single-stage object detector that comprises *backbone*, *neck*, and *head* components. The backbone is a pre-trained network that extracts rich features for images. This curtails the spatial resolution of an image and increases the feature resolution. The neck extracts feature pyramids helping the model to generalize objects of various shapes and sizes. The body finalizes the operation and gives the output in the form a bounding box, prediction score, and classes. Furthermore, CSP Darknet53 is used as the backbone; SPP and PANet as the neck. Whereas YOLOv4's head is used as YOLOv5's head.

4.1 Detection and Results

As we trained our model for detecting a person and a ladder, we achieved favourable and satisfactory results regarding the detection of person, ladder, and breach. Initially, the prediction scores were mostly in range of 40% to 60%; thus, we further enhanced our custom dataset.

Following are the results of our model:

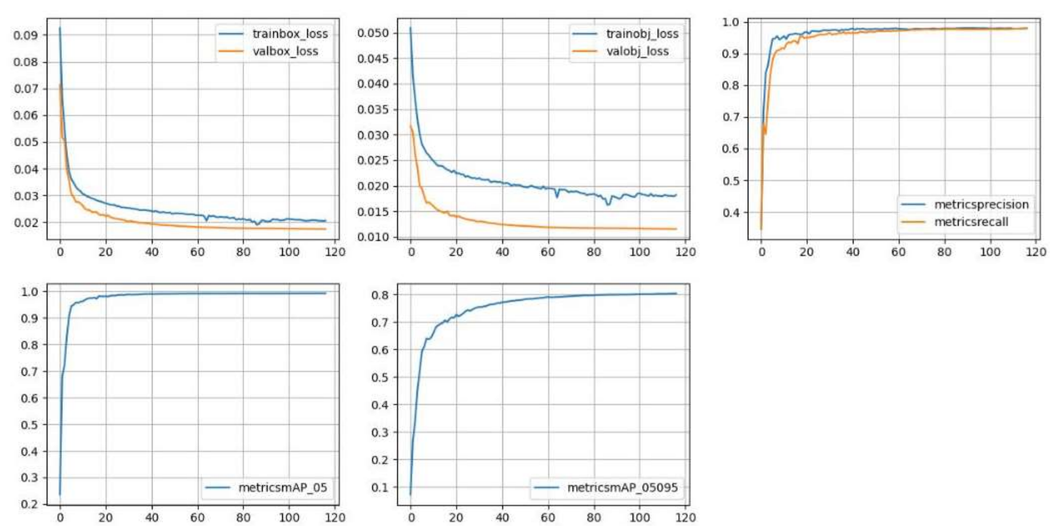


Figure 10: Overall results of AISS model

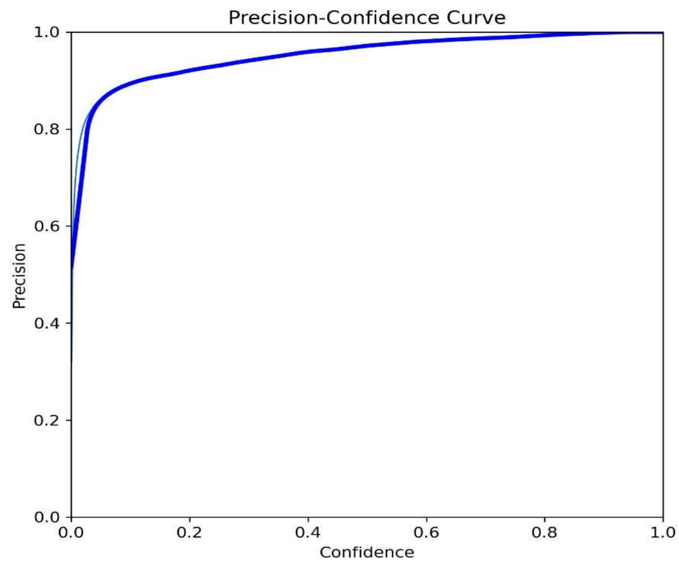


Figure 11: Precision-Confidence curve of AISS model

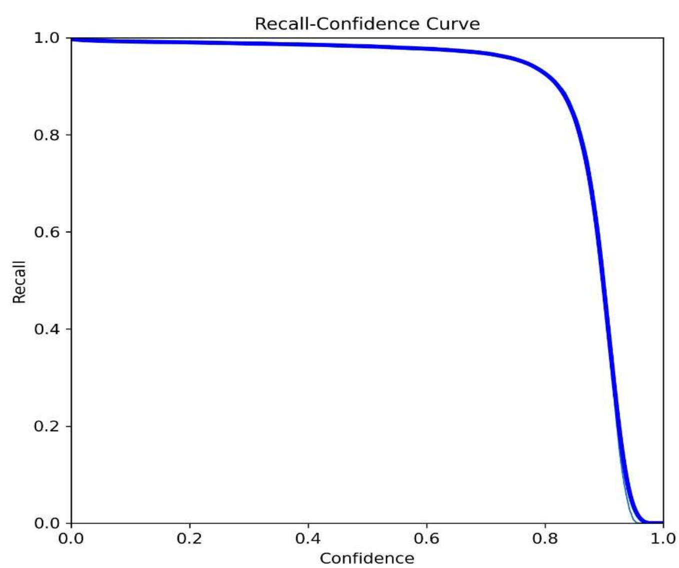


Figure 12: Recall-Confidence curve of AISS model

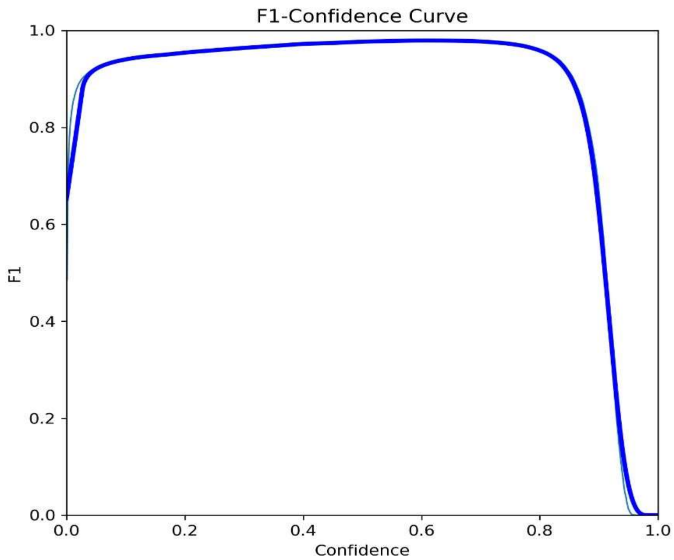


Figure 13: F1-Confidence curve of AISS

As seen in the results of our model, the precision is at 80.4%, recall is at 79.5%, and mean average precision (mAP) is at 82.5%. The prediction score of person and ladder are ~85% and ~80%, respectively, in validation and those scores are 85% and 80%, respectively, in testing. The model has performed the object detection on the video stream resized to 640×640 pixels. The results indicate that the model has a higher accuracy, but still requires further optimization and improvement. For that purpose, a greater number of images at different angles are necessary for the model's better training. The images of ladder and person are ought to be taken separately, too, which improves the object detection of each class. The newer images must be taken in environments as the system will be implemented in different environments as well. This process will aid the model to deduct the number of false positives it gives; thus, improving the performance metrics in terms of precision, recall, and F1 score.

As mentioned in **Chapter 3.1.2**, the hyperparameters were set as such due to computational limitations since we were training the model on our computers. However, to further optimize the algorithm, we can to increase **batch size** and **epochs** given the limited computational resources we have. By this approach, we can further improve the detection results.

4.2 Graphical User Interface (GUI)

For our project, we had to employ the functions that aid the algorithm to process visual and computational data. Since we have written the code in Python, the following necessary libraries were used:

i. **numpy** is used for computing large-scale calculations.
ii. **cv2** is used for image and video processing.
iii. **torch** for pre-processing, model building, and optimization were done using this library.
iv. **threading** managing threads for enhanced processing of algorithm.

This code defines a program that captures video from multiple cameras simultaneously and runs object detection on each frame of the video using YOLOv5. The program has a graphical user interface (GUI) created using PyQt5 library. The program imports **os** and **PyQt5** libraries, too. It then sets the width at 480 pixels, height at 640 pixels of the video frames, and the delay between each frame capture. Th capture delay is set at 10 seconds in order for smooth execution of the code. The program loads a **YOLOv5** object detection model from a local path using **torch.hub.load()** function.

The program then defines a class called **NewWindow** that inherits from **QDialog**. This class creates a window that will display the video stream from each camera. It contains a **QLabel** object that will display the video frames. The program also defines a Thread class that inherits from **QThread**. This class captures the video stream from each camera and runs object detection on each frame. It emits a **pyqtSignal** that sends the processed frame, camera index, and a Boolean indicating whether the camera is active.

The program also defines a main function that creates a **QApplication** object and initializes a list of camera links. It then creates a list of **NewWindow** objects that will display the video stream from each camera. It then creates a list of Thread objects, one for each camera. The **clickable()** function is a helper function that makes any widget clickable. It creates a filter object that emits a **clicked** signal whenever the widget is clicked. The **TableStatus** class creates a **QTableWidget** and sets its properties such as the number of rows and columns, column headers, font, color, and selection mode. The **updateTable()** method updates the table with the camera links and their status. It sets the background and foreground colors of the cells and resizes the table columns and rows to their contents. The **AddImage** class creates a **QMainWindow** with a label, line edit, and button widgets.

Chapter 5: Conclusion

We have discussed at length about our project that provides a reliable solution regarding security and surveillance. This is a cost-effective and simple-to-install solution for clients as they only require to deploy the machine learning algorithm on their existing CCTV system. More importantly, it reduces the strain on the security staff who otherwise has to constantly monitor the CCTV feed around the clock. The proposed system is also computationally cost-effective as we have achieved better accuracy with a lightweight network and has a lower number of parameters. Thus, the clients don't need to speed extra money and require a separate, high-speed system to operate the model.

We achieved success in training our model to detect anomalous gestures (e.g., climbing), raise the alarm upon breach, and notify the concerned staff and personnel timely via call. Further on, the information and image of the breach are also sent perfectly to the staff. The model gave accurate results – i.e., up to 70% – as we had curated our own custom dataset, which was instrumental in detecting a breach. This indicates that our model is working accurately, easily detecting anomalous behaviour, which is trespassing in our case.

As mentioned above, the main intent of this system is to curtail the stress and likelihood of human error regarding security and surveillance. It cuts down the burden and lethargy the staff feel while monitoring the CCTV feeds. This system aids the staff in their duties to maintain the security of their institution's perimeters. The graphical user interface (GUI) also facilitates the staff to easily manage the camera feeds and promptly notify a developing situation. The security staff can also devise better standard operating procedures (SOPs) on patrolling and surveillance of the boundaries of institutions.

Chapter 6: Future Work and Development

Although our model has high accuracy, there is still some room for improvement. Therefore, we have to increase the number of images and quality of the custom dataset curated. This includes taking images the anomalous behaviour from different angles in different environments. Secondly, we also require taking separate images ladder, too, for improving its detection.

Additionally, we will also add different types of anomalous behaviours and train the model upon them. In other words, we will curate more custom datasets of anomalous behaviours; for instance, placing a bag alongside the wall, throwing an object inside the perimeter, and et cetera.

Moreover, we felt the need to improve our user interface for commercial aspects. The improved usage of the interface will aid the security staff to easily perform and manage their tasks.

In conclusion, the future work and development to be done for our project are the following:

i. More images for our custom dataset.
ii. Images are to be taken in different environments for better training of our model.
iii. New classes of anomalous behaviours in the dataset.
iv. Automatic setting of the logical zone over the walls.
v. Separate alarm circuitry.
vi. Improving the notification display.

References and Citations

1. Pikaar, Ruurd, et al. "Human Factors Guidelines for CCTV system design." *Proceedings 19th Triennial Congress of the IEA, Melbourne*. 2015.

2. Allwood, G., Wild, G., & Hinckley, S. (2016). Optical fibre sensors in physical intrusion detection systems: A review. *IEEE Sensors Journal*, *16*(14), 5497-5509.

3. B. Mukhopadhyay, S. Anchal and S. Kar, "Detection of an Intruder and Prediction of His State of Motion by Using Seismic Sensor," in IEEE Sensors Journal, vol. 18, no. 2, pp. 703-712, 15 Jan.15, 2018, DOI: 10.1109/JSEN.2017.2776127.

4. Nguyen, V.T., Nguyen, T.B., & Chung, S. (2016). ConvNets and AGMM based real-time human detection under fisheye camera for embedded surveillance. *2016 International Conference on Information and Communication Technology Convergence (ICTC)*, 840-845.

5. Chen, W., Chen, P., Lee, W., & Huang, C. (2008). Design and Implementation of a Real Time Video Surveillance System with Wireless Sensor Networks. *VTC Spring 2008 - IEEE Vehicular Technology Conference*, 218-222.

6. Haryono, H. P., & Hidayat, F. (2022, August). Trespassing Detection using CCTV and Video Analytics for Safety and Security in Railway Stations. In 2022 International Conference on ICT for Smart Society (ICISS) (pp. 01-04). IEEE.

7. Xiao, Y., Chang, A., Wang, Y., Huang, Y., Yu, J., & Huo, L. (2022, May). Real-time Object Detection for Substation Security Early-warning with Deep Neural Network based on YOLO-V5. In *2022 IEEE IAS Global Conference on Emerging Technologies (GlobConET)* (pp. 45-50). IEEE.

Students FYP

ORIGINALITY REPORT

8%	6%	2%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDEN

PRIMARY SOURCES

1	Submitted to Higher Education Commission Pakistan Student Paper
2	Submitted to The University of the West of Scotland Student Paper
3	iq.opengenus.org Internet Source
4	mahb.listedcompany.com Internet Source
5	www.ijser.in Internet Source
6	Quoc Bao Phan, Tuy Nguyen. "A Novel Approach for PV Cell Fault Detection using YOLOv8 and Particle Swarm Optimization", Institute of Electrical and Electronics

8 Submitted to University of Glasgow
Student Paper

9 Submitted to Liverpool John Moores University
Student Paper

10 uir.unisa.ac.za
Internet Source

11 www.coursehero.com
Internet Source

12 assets.researchsquare.com
Internet Source

13 repository.sustech.edu
Internet Source

14 www.mdpi.com
Internet Source

15 etd.aau.edu.et
Internet Source

16 "Intelligent Communication, Control and Devices", Springer Science and Business Media LLC, 2018
Publication