

**VANDALISM VIDEO ANALYSIS EMPLOYING COMPUTER VISION  
TECHNIQUE**

**BY**

**BRITNEY MUK YUEN KUAN**

**A REPORT**

**SUBMITTED TO**

**Universiti Tunku Abdul Rahman**

**in partial fulfilment of the requirements**

**for the degree of**

**BACHELOR OF COMPUTER SCIENCE (HONOURS)**

**Faculty of Information and Communication Technology**

**(Kampar Campus)**

**JAN 2022**

UNIVERSITI TUNKU ABDUL RAHMAN

**REPORT STATUS DECLARATION FORM**

**Title:** VANDALISM VIDEO ANALYSIS EMPLOYING COMPUTER VISION  
TECHNIQUE

**Academic Session:** JAN 2022

I BRITNEY MUK YUEN KUAN  
(CAPITAL LETTER)

declare that I allow this Final Year Project Report to be kept in  
Universiti Tunku Abdul Rahman Library subject to the regulations as follows:

1. The dissertation is a property of the Library.
2. The Library is allowed to make copies of this dissertation for academic purposes.

Verified by,



(Author's signature)



(Supervisor's signature)

**Address:**

95, JALAN ABADI 5,  
TAMAN MUTIARA RINI,  
81300, SKUDAI, JOHOR.

Leung Kar Hang  
Supervisor's name

**Date:** 17 April 2022

**Date:** 18 April 2022

<b>Universiti Tunku Abdul Rahman</b>			
Form Title : <b>Sample of Submission Sheet for FYP/Dissertation/Thesis</b>			
Form Number: <b>FM-IAD-004</b>	Rev No.: <b>0</b>	Effective Date: <b>21 JUNE 2011</b>	Page No.: <b>1 of 1</b>

**FACULTY/INSTITUTE\* OF Information and Communication Technology**  
**UNIVERSITI TUNKU ABDUL RAHMAN**

Date: 17 April 2022

**SUBMISSION OF FINAL YEAR PROJECT /DISSERTATION/THESIS**

It is hereby certified that Britney Muk Yuen Kuan (ID No: 18ACB03450 ) has completed this final year project/ dissertation/ thesis\* entitled “ Vandalism Video Analysis Employing Computer Vision Technique ” under the supervision of Prof. Leung Kar Hang (Supervisor) from the Department of Computer Science , Faculty/Institute\* of Information and Communication Technology .

I understand that University will upload softcopy of my final year project / dissertation/ thesis\* in pdf format into UTAR Institutional Repository, which may be made accessible to UTAR community and public.

Yours truly,



---

Britney Muk Yuen Kuan

## DECLARATION OF ORIGINALITY

I declare that this report entitled “**VANDALISM VIDEO ANALYSIS EMPLOYING COMPUTER VISION TECHNIQUE**” is my own work except as cited in the references. The report has not been accepted for any degree and is not being submitted concurrently in candidature for any degree or other award.

Signature :



Name : BRITNEY MUK YUEN KUAN

Date : 17/4/2022

## **ACKNOWLEDGEMENTS**

I would like to take this opportunity to express my sincere thanks and appreciation to my supervisors, Maylor Leung Kar Hang who have been providing advices and helps throughout the development of this project. His valuable feedbacks and comments definitely help me to think outside of the box and finally developed the program with desirable quality.

In addition, I would also like to appreciate my parents who always provide me unconditional love and support throughout my academic life, and my brother and sister who would lend a hand to me when I'm depressed. Last but not least, I am grateful for all the people who support me. Thank you very much.

## **ABSTRACT**

In this advance era of technology, the computer vision technique is involved regularly in surveillance system compare to last decade. This project is carried in the field of image processing to solve and improve the problem of vandalism activity occurred in the town that lead to massive destruction and repair costs, also to protect public and private properties by preventing vandalism.

This project is a development of an intelligent surveillance system that can detect vandalism events and the proposed novel method is implemented with the technique of YOLO detection, suspicious characteristic detection, and background changes detection. In the proposed system, the method monitors the changes inside the captured scene. When there is a human enter the scene and there are significant changes, indicating damage, a vandalism event is declared and warning alert will be triggered to scare the vandals off. On the other hand, the characteristics and behaviour of the suspicious vandal will also be monitored. Warning is flagged when the probability of vandal behaviour is exceeding a threshold and the early warning will be given out to prevent the vandalism events.

The method is tested on the UCF\_Crime dataset with around 50 different videos containing vandalism scenes such as spraying paint, breaking windows, defacing public property and etc

## TABLE OF CONTENTS

TITLE PAGE .....	i
REPORT STATUS DECLARATION FORM .....	ii
FYP THESIS SUBMISSION FORM .....	iii
DECLARATION OF ORIGINALITY .....	iv
ACKNOWLEDGEMENTS .....	v
ABSTRACT.....	vi
TABLE OF CONTENTS.....	vii
LIST OF FIGURES .....	ii
LIST OF TABLES.....	iv
LIST OF ABBREVIATIONS.....	v
Chapter 1 : Introduction .....	1
1.1    Problem Statement and Motivation.....	1
1.2    Project Objectives .....	2
1.3    Project Scope.....	3
1.4    Impact, significance and contribution.....	5
1.5 Background information .....	6
1.6    Report Organisation .....	7
Chapter 2 : Literature Review .....	8
2.1    Comprehensive Review .....	8
2.1.1    Anomalies/Outliers .....	8
2.1.2    Conventional Integrated Image Processing Methods .....	10
2.1.3    Deep-Learning Based Methods for Video Anomaly Detection.....	15
2.1.4    Comparative Analysis of Video Anomaly Detection Method.....	18
2.2    Automated Surveillance System .....	19
2.2.1    Real-Time Detection of Suspicious Behaviours In Shopping Malls [9].....	19

2.2.2	Fast and Robust Occluded Face Detection in ATM Surveillance [13].....	21
2.2.3	A Study of Deep Convolutional Auto-Encoders for Anomaly Detection in Videos [15]	22
2.2.4	Real-world Anomaly Detection in Surveillance Videos [3] .....	24
2.2.5	Summary of the Reviewed System .....	27
Chapter 3 : System Methodology/Approach.....		28
3.1	Real Case Scenario.....	28
3.1.1	Scenario A.....	28
3.1.2	Scenario B .....	29
3.1.3	Scenario C.....	30
3.1.4	Discussion .....	31
3.2	Design Specifications.....	32
3.2.1	Methodologies and General Work Procedures .....	32
3.2.2	Assumptions.....	33
Chapter 4 : System Design.....		34
4.1	System Design / Overview.....	34
4.2	System Components Specifications .....	36
4.2.1	Initialization and Pre-processing.....	36
4.2.2	YOLO Detection [19] .....	37
4.2.3	Loitering Detection .....	39
4.2.4	Significant Background Changes Detection .....	40
4.2.5	Background Updating .....	42
Chapter 5 : System Implementation.....		44
5.1	Hardware Setup.....	44
5.2	Software Setup .....	44
5.3	System Operation.....	45



5.3.1	Vandalism Video Analysis .....	45
5.3.2	Detailed Window (Significant background changes detection).....	47
Chapter 6 : System Evaluation and Discussion .....		48
6.1	System Testing and Result.....	48
6.1.1	Scenario A.....	48
6.1.2	Scenario B.....	51
6.1.3	Scenario C.....	54
6.1.4	Scenario D.....	56
6.1.5	Scenario E.....	58
6.2	Project Challenges.....	60
6.3	Objectives Evaluation .....	61
Chapter 7 : Conclusion and Recommendation.....		62
7.1	Conclusion .....	62
7.2	Recommendation .....	63
Bibliography .....		64
Appendices.....		A-1
A.1	Experimental Result.....	A-1
A.2	Final Year Project 2 Weekly Report .....	A-3
A.3	Poster.....	A-13
A.4	Plagiarism check result .....	A-14
A.5	FYP 2 Checklist .....	A-18

## LIST OF FIGURES

Figure 1.1 List of videos from UCF_Crime dataset .....	3
Figure 1.2 Sample videos involved vandalism .....	4
Figure 2.1 Classification of anomalies.....	9
Figure 2.2 Classification of the training and learning frameworks .....	9
Figure 2.3 Two main approaches for the video anomaly detection.....	10
Figure 2.4 Fundamental image processing steps [5].....	11
Figure 2.5 Classification of deep learning-based methods for the detection and localization of video anomalies [7].....	15
Figure 2.6 Video pre-processing and human tracking (Arroyo et al., 2015).....	19
Figure 2.7 Occlusion management [9].....	20
Figure 2.8 Loitering detection [9].....	20
Figure 2.9 Examples of face occlusion detection results [13] .....	22
Figure 2.10 Architecture of the proposed CAE [15].....	23
Figure 2.11 Reconstruction error on detecting anomalies [15].....	24
Figure 2.12 Flow diagram of the proposed anomaly detection approach [3] .....	25
Figure 2.13 Evolution of score on training video over iterations [3].....	26
Figure 3.1 Video sequence of scenario A .....	28
Figure 3.2 Video sequence of scenario B .....	29
Figure 3.3 Video sequence of scenario C .....	30
Figure 3.4 Methodology on vandalism detection .....	32
Figure 4.1 System diagram .....	34
Figure 4.2 Initialization.....	36
Figure 4.3 YOLO detection .....	37
Figure 4.4 Loitering detection.....	39
Figure 4.5 Significant background changes detection .....	40

Figure 4.6 Background Estimation .....	43
Figure 5.1 Vandalism video analytic interface .....	45
Figure 5.2 Result plane .....	45
Figure 5.3 Analysis board .....	46
Figure 5.4 Detailed Window for Significant background changes detection .....	47
Figure 6.1 Scenario A Result .....	49
Figure 6.2 Scenario B Result .....	52
Figure 6.3 Scenario C Result .....	55
Figure 6.4 Scenario D Result .....	57
Figure 6.5 Scenario E Result .....	59

## LIST OF TABLES

Table 2.1 Studies on object detection, adapted from [5] .....	11
Table 2.2 Studies on face detection, adapted from [5] .....	12
Table 2.3 Studies on facial component detection, adapted from [5] .....	13
Table 2.4 Studies on object shape and appearance detection, adapted from [5] .....	13
Table 2.5 Studies on decision making, adapted from [5] .....	14
Table 2.6 Selection of video anomaly detection methods .....	18
Table 2.7 Table of Comparison .....	27
Table 5.1 Hardware Tools for Development .....	44
Table 5.2 Software Tools for Development.....	44

## LIST OF ABBREVIATIONS

AI	Artificial Intelligence
AO	Accuracy-Oriented
C3D	Convolutional 3D
CAE	Convolutional Auto-Encoders
CCTV	Closed-circuit television
CV	Computer Vision
DNN	Deep Neural Network
GCH	Global Colour Histogram
HAR	Human Action Recognition
HOG	Histogram of Oriented Gradients
IDE	Integrated Development Environment
IVSS	Intelligent Video Surveillance System
LBP	Local Binary Pattern
LSAP	Linear Sum Assignment Problem
MHT	Modified Hough Transform
Mil	Multiple Instance Learning
ML	Machine Learning
OCC	One-Class Classification
PO	Processing-Time-Oriented
R-CNN	Region Based Convolutional Neural Networks
SVM	Support Vector Machine
TCNN	Tube Convolutional Neural Network
YOLO	You Only Look Once

### **Chapter 1 : Introduction**

#### **1.1 Problem Statement and Motivation**

Physical vandalism, the intentional action of damaging property belongs to the other people e.g. breaking windows, spraying paint, destroying public facilities, and etc have been occurring frequently worldwide. The vandal behaviours are defined as unlawful destructing or damaging of public or private property [1]. Vandalism is becoming an increasingly important crime to remark. One of the significant vandalisms occurred in recent years was the 2019-20 Hong Kong protest. The bill for repairing public facilities had hit US\$8.4 million and it also vitiate the quality of life of the society when vandalism escalates to a more serious crime [2]. Therefore, surveillance systems are essential on watching and recording scenes of streets and buildings to lower down the crime rate. However, such system requires continuous monitoring by humans which is time consuming and waste of human resource. Furthermore, there is the possibility of failure on detecting the crime event due to boredom or fatigue of security staff. Therefore, an instant automated detection on vandalism has become a demanding issue to address such problems. With advanced modern technologies, video surveillance camera employing computer vision techniques can analyse and classify the vandalism scenes and give early warning in order to prevent vandalism. The real-time vandalism analysis and detection can greatly decrease the need for human labour and increase the productivity by avoiding human error.

## 1.2 Project Objectives

The objectives of this project are:

✓ **To develop a real-time automated vandalism detection surveillance system**

As the problem statement mentioned earlier, the surveillance system requires manpower for continuous monitoring and the fatigue of security staff may lead to failure of detecting abnormality. Thus, a real-time automated vandalism detection surveillance system can reduce the human labour on continuous monitoring and benefit the communities by handling the vandalism events to enhance security.

✓ **To develop the vandalism detection system by accomplishing the following sub-objectives:**

○ **To predict the potential vandals by identifying characteristics.**

To give off early warning for the certain risk situation or suspicious behaviours.

○ **To detect the significant static changes of the vandalism-prone object**

*If there are such changes, representing damage. A vandalism event is declared*

To give off alarm and prevent further destruction of property.

### 1.3 Project Scope

The scope of this project covers the analysis and classification of vandalism scenes based on the UCF\_Crimes dataset [3] and to explore solutions to detect vandalism employing computer vision techniques. Throughout the planning, a piece of software to detect a few categories of vandalized events and to give early warning of such incidences would be developed in this project.

There are 50 videos involving scenes of vandalism from the UCF\_Crime dataset as listed in Figure 1.1, all the videos are around one to four minutes long which focus the vandalism incident. Besides, Figure 1.2 shows some of the sample vandalism event from the dataset such as graffiti, red paint splasher, glass breakage and etc.

Name	Size	Length	Name	Size	Length
Vandalism001_x264.mp4	12,319 KB	00:00:47	Vandalism026_x264.mp4	2,737 KB	00:00:32
Vandalism002_x264.mp4	11,653 KB	00:01:01	Vandalism027_x264.mp4	3,891 KB	00:00:23
Vandalism003_x264.mp4	11,631 KB	00:00:49	Vandalism028_x264.mp4	28,464 KB	00:02:29
Vandalism004_x264.mp4	20,399 KB	00:01:34	Vandalism029_x264.mp4	4,286 KB	00:00:15
Vandalism005_x264.mp4	3,911 KB	00:00:54	Vandalism030_x264.mp4	12,217 KB	00:00:48
Vandalism006_x264.mp4	2,785 KB	00:00:27	Vandalism031_x264.mp4	12,418 KB	00:01:00
Vandalism007_x264.mp4	6,289 KB	00:00:38	Vandalism032_x264.mp4	1,819 KB	00:00:22
Vandalism008_x264.mp4	103,253 KB	00:06:59	Vandalism033_x264.mp4	18,128 KB	00:01:15
Vandalism009_x264.mp4	12,751 KB	00:00:44	Vandalism034_x264.mp4	14,571 KB	00:00:58
Vandalism010_x264.mp4	25,439 KB	00:02:11	Vandalism035_x264.mp4	6,811 KB	00:00:39
Vandalism011_x264.mp4	103,164 KB	00:06:54	Vandalism036_x264.mp4	12,277 KB	00:00:48
Vandalism012_x264.mp4	4,432 KB	00:00:28	Vandalism037_x264.mp4	20,149 KB	00:01:40
Vandalism013_x264.mp4	8,719 KB	00:01:19	Vandalism038_x264.mp4	8,003 KB	00:01:34
Vandalism014_x264.mp4	3,368 KB	00:00:14	Vandalism039_x264.mp4	3,274 KB	00:00:30
Vandalism015_x264.mp4	14,433 KB	00:01:39	Vandalism040_x264.mp4	7,051 KB	00:00:34
Vandalism016_x264.mp4	17,583 KB	00:03:13	Vandalism041_x264.mp4	36,260 KB	00:02:42
Vandalism017_x264.mp4	7,951 KB	00:00:33	Vandalism042_x264.mp4	28,635 KB	00:01:59
Vandalism018_x264.mp4	27,890 KB	00:03:05	Vandalism043_x264.mp4	19,132 KB	00:01:48
Vandalism019_x264.mp4	6,237 KB	00:01:08	Vandalism044_x264.mp4	27,523 KB	00:02:33
Vandalism020_x264.mp4	8,554 KB	00:00:48	Vandalism045_x264.mp4	35,774 KB	00:02:25
Vandalism021_x264.mp4	10,313 KB	00:00:45	Vandalism046_x264.mp4	17,090 KB	00:01:10
Vandalism022_x264.mp4	28,123 KB	00:02:11	Vandalism047_x264.mp4	10,660 KB	00:01:22
Vandalism023_x264.mp4	51,568 KB	00:03:30	Vandalism048_x264.mp4	56,067 KB	00:03:59
Vandalism024_x264.mp4	8,555 KB	00:00:33	Vandalism049_x264.mp4	65,805 KB	00:04:26
Vandalism025_x264.mp4	34,344 KB	00:02:23	Vandalism050_x264.mp4	7,156 KB	00:00:30

*Figure 1.1 List of videos from UCF\_Crime dataset*





*Figure 1.2 Sample videos involved vandalism*

### **1.4 Impact, significance and contribution**

One and the most problem to be concerned, the property damage caused by vandalism, especially those properties that no one has direct responsibility on it, or those area that seem to be less guarded, are frequent targets of vandalism. The end product developed in this project will be a robust vandalism detection system with analysis of the motion pattern and characteristic of people that can recognise the potential vandal and prevent them from vandalising the private and public properties. The real-time automated vandalism detection surveillance system will replace the job of security guard in continuous monitoring and also reduce the failure on detecting abnormality due to fatigue of the security guard. More importantly, this project will help to address the community's crime problem and develop a safer city to live, work, and shop.

By consider some cases that unable to be detect suspicious vandal, the system will also monitor the object property and vandalism will be declared when human enters the captured scene and causes an unauthorized damage. As such, it can prevent the further destruction and preserve the condition of the property with greatest possible extent by interrupt the vandalism event from the instant rising warning alert. Therefore, with this two-level defence system the surveillance can be more reliable and comparative among the other systems.

### **1.5 Background information**

Surveillance systems are the monitoring of behaviour, action, and activity, designed to secure a particular area, street or blind spot. One of the most popular surveillance systems is the video surveillance, also known as Closed-circuit television (CCTV), every change in the scene will be captured by the camera and transmit into video sequence for ongoing monitoring.

However, the need for continuous human supervision is one of the major disadvantages of all the camera-based monitoring system. Also, an individual generally loss his/her focus on the screen after a number of hours and could lead to a reduction in the effectiveness of crime detecting due to such poor monitoring. Thus, intelligent Video Surveillance System (IVSS) employing computer vision technique are the essential improvements that can provide huge benefit to the community.

While computer vision (CV) is a field of study that focus on the problem of helping computer to see and understand the content of digital images such as photograph and video. Generally, image processing may be required for a computer vision system to pre-processing the raw input, typically simplifying or enhancing the content such as crop the bounds of the image, reduce digital noise from the image and enhance intensity of the image. It can be considered as a type of digital signal processing and not dealing with understanding the content of the photography. Therefore, the application of artificial intelligence (AI) or hand-crafted statistical method will be used in computer vision problem, to train a computer to perform humanlike tasks, such as detecting an object, identifying the item carried by an individual on the basis of the level of danger, recognizing the human action and making predictions.

In recent years, the field of human activity recognition (HAR) were growing rapidly, reflecting its importance in many societal applications including intelligent surveillance system, video search and retrieval, and robot perception. Basically, Human Action Recognition aims to understand human behaviours and assign a label to each action. With the emergence of human action recognition, the ability to predict and forecast the complex human activity is made possible, and so the vandalism detection.

### **1.6 Report Organisation**

The report was organized in six chapters. In the sequence of Introduction, Literature Review, System Methodology/Approach, System Design, System Implementation, System Evaluation and Discussion, and Conclusion and Recommendation. The first chapter covers the project's overview, which includes problem statement and motivation, project scope, project objectives, contribution, background information and report organisation. Next, Chapter 2 Literature Review will cover the comprehensive researches of the state of art image processing method and the review on some existing Automated Surveillance System. Furthermore, the third chapter is the System Methodology/Approach where the methodologies and general work procedures will be discussed. On the other hand, the system design with overview and detailed sub-topics will be provided in the fourth chapter. The fifth chapter, System Implementation is regarding the details on how to implement the design of the system. After that, the sixth chapter reports the experimental result and respective performance. Lastly, the conclusion and recommendation will be given at the final chapter.

## Chapter 2 : Literature Review

In the 21st century, the growth of technology is developing at a rapid stride while anomaly detection in computer vision is still one of the most long-standing and challenging problem as the ambiguities in detection do not only come from the difficulty on defining the motion of body parts, but also come from many other challenges such as blurry video, unclear vision, occlusion based on bad weather condition, camera quality or crowd situation.

There are two sections in this Chapter. In section 2.1 Comprehensive Review, we discuss the research papers compiled with several proposed methods such as image processing methods and deep learning-based methods from previous studies for general knowledge. While for section 2.2 Automated Surveillance System, we discuss the strengths and weaknesses of some methods from the reviewed paper. A fundamental knowledge on video anomaly detection will be developed in this chapter before proceeding to the next stage.

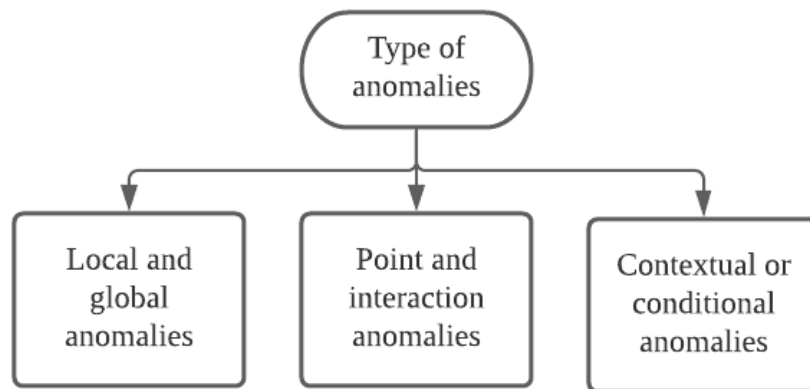
### 2.1 Comprehensive Review

#### 2.1.1 Anomalies/Outliers

Before proceeding to the state-of-the-art image processing method and the popular deep learning methods, the aspects such as *classification on anomalies, training and learning frameworks and approaches for video anomaly detection* were essential for basic understanding which treated as a key task for AI that would influence the later selection of various techniques on video anomaly detection.

#### Classification of anomalies

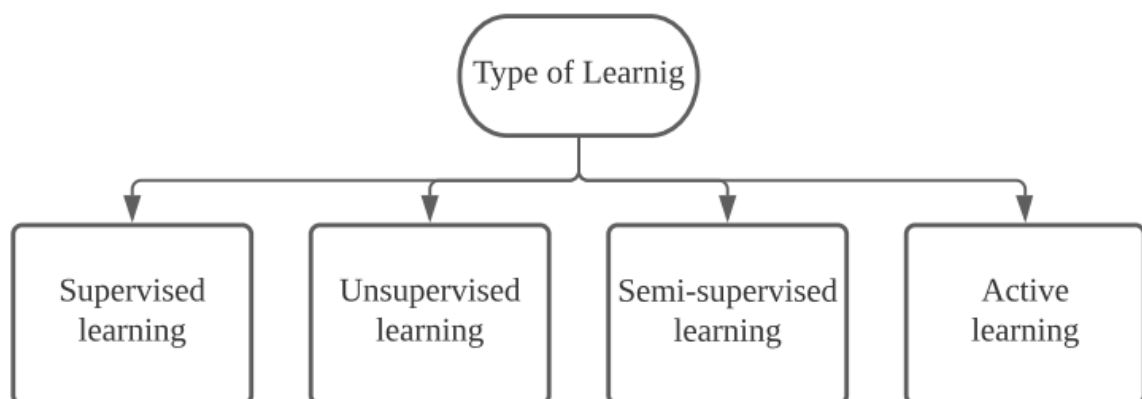
Anomalies differ decisively in their occurrences. Basically, anomalies could be divided into three types as shown in Figure 2.1. Local anomalies were those instances significantly deviates from its neighbourhood density. Inversely, global anomalies often refer to the activity that interact globally with the other but exhibit as an unusual and rare individual, global anomaly also known as collective anomalies in which collectively exhibit abnormal behaviours. Point anomalies classified as a single instance of data different from the others. Contextual or conditional anomalies were referred to the data point appeared to be anomalous only in a specific context [4].



*Figure 2.1 Classification of anomalies*

Training and learning frameworks

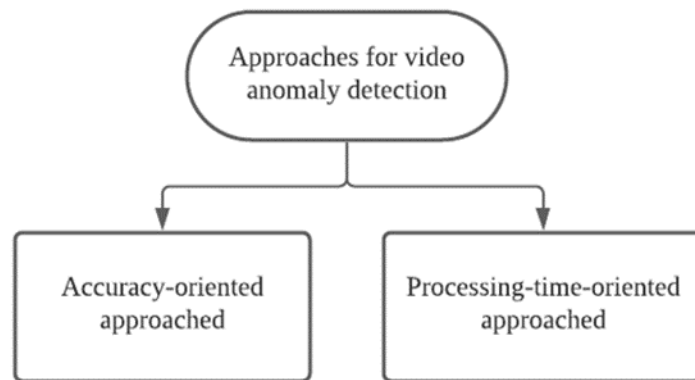
As presented in Figure 2.2 the training and learning frameworks of deep-learning methods in video anomaly detection were generally classified into four types. Supervised learning involved the training of classifier using the associated labelled datasets, while unsupervised learning detected the anomalous activity using the co-occurrence of events from unlabelled data. On the other hand, semi-supervised learning used the advantages of the supervised and unsupervised learnings whereby only weakly-labelled video were used for model training. Lastly, active learning was a technique where the domain experts (humans) were assisting in labelling the confusing sample data in order to help in minimising the ambiguity during the learning process [4].



*Figure 2.2 Classification of the training and learning frameworks*

Approaches for video anomaly detection

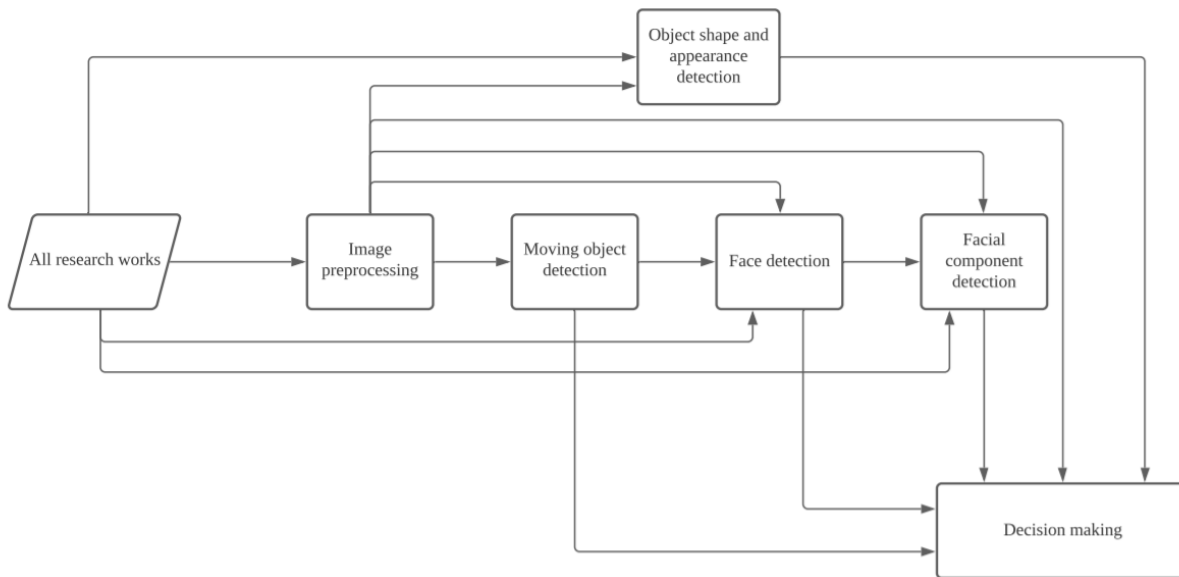
As shown in Figure 2.3, the approached for anomaly detection can be classified into two main approaches. Accuracy-oriented (AO) approach aimed at the high precision of detection and localization of video anomalies. A complex model would be trained at the expense of processing-time to achieve the desired high accuracy. Meanwhile, processing-time-oriented (PO) approach aimed at detecting and localising video anomalies with the shortest amount of time possible while maintaining a competitive degree of accuracy [4].



*Figure 2.3 Two main approaches for the video anomaly detection*

**2.1.2 Conventional Integrated Image Processing Methods**

In the systematic review of integrated video surveillance and image processing for criminal detection [5], they attempted to compile previous studies and assist future researchers in developing dynamic and versatile surveillance algorithms. As illustrated in Figure 2.4, there was the fundamental image processing steps that generally implemented in the methodology and system framework. The most common steps included, but not limited to:



**Figure 2.4 Fundamental image processing steps [5]**

- Moving Object Detection

Table 2.1 summarizes the object detection in terms of moving object features and method of extraction. Frame difference (FD) are the method commonly used in motion edge or foreground extraction. FD will obtain the pixel-based difference between the background and the current foreground motion frame, while the Canny operator will smooth the image using a Gaussian filter by eliminating the noise. Another approach was to combine the FD image's distance transformation (DT) which compute the exact distances from inexact feature and the edge detection in order to extract the foreground.

Moving object feature	Method used	Author, year
Moving edge, moving object region	Canny operator, edge differencing, particle noise removal morphological bridging, hole filling, morphological closing and opening	Dong et al., 2006
	Frame difference, Sobel operator, dilation, erosion, straight line fitting method, bounding box's aspect ratio	Lin et al., 2006
Foreign pixel	Background subtraction or Frame differencing	Sako et al., 2007
Foreground	Frame variance, histogram	Zhang et al., 2014
	Rolling average background subtraction technique, morphological operation	Goswami et al., 2015
	Distance transformation, edge detection	Zhang et al., 2018)

**Table 2.1 Studies on object detection, adapted from [5]**



- Face Detection

Table 2.2 summarizes the face detection step in terms of face features and detection methods. In face detection, there were few studies considered the ellipse as face shape and exploited the ellipse-fitting technique, while colour space could be combined with the techniques as for more accurate result. In the case of occlusion and changing view, the other technique Colour histogram matching (CHM) through searching for skin-coloured pixels well performed in face detection as it was the stable object representations. Some other studies used trained face detectors such as MCT-based AdaBoost face detector and CascadeObjectDetector which used the Viola–Jones Algorithm. However, certain parameters of this function had to be introduced by the investigator in order to improve the performance for a more robust detection against external conditions.

Feature extracted	Method used	Author, year
Face	Colour histogram matching, CHM	Sako et al., 2004
	MCT-based AdaBoost face detector	Choi et al., 2010
	Ellipse fitting method	Kim et al., 2010
Face region	HSV colour space, hole filling, morphological closing and opening	Dong et al., 2006
Face location	Elliptical approximate algorithm	Lin et al., 2006
Skin colour, Boundary shape	YCbCr elliptical model, mathematical morphology	Hongxing et al., 2013
Face like object	Vision. CascadeObjectDetector, Viola-Jones algorithm, bounding box, threshold merging, size parameter range	Ray et al., 2015

*Table 2.2 Studies on face detection, adapted from [5]*

- Facial Component Detection

Table 2.3 summarizes the facial component detection such as eye, mouth and nose. The facial components can be identified as Gabor features, this method could be improved by an additional dimensionality reduction scheme such as principal component analysis (PCA) due to the high dimensional of Gabor feature that led to the high computational complexity. In addition, the Template Matching (TM) could be used to identify eye components by comparing the regions of high correlation with eyes shape templates, because eye was considered as the only unique and common in every human face. While, the geometric and algebraic measures could be the alternative approach for the extraction of facial

characteristics, because people's face features differ from each other, it acted as a consistent measurement for each individual.

Facial component feature	Method used	Author, year
Eye	Template matching, TM	Sako et al., 2004
Eye, mouth	Gabor filtering	Wen et al., 2005
	Gabor wavelet, Principal Component Analysis, PCA	Min et al., 2014
	MCT-based AdaBoost eye and mouth detector	Choi et al., 2010
	Viola jones facial component detector	Suhr et al., 2012
Eye, mouth, nose	Principal component analysis, PCA	Yoon et al., 2002
		Kim et al., 2005
	Geometric measures (Euclidean distance, curvature, angle), Algebraic measure (matrix characteristic vectors)	Min et al., 2011
	Viola-Jones algorithm	Ray et al., 2015

*Table 2.3 Studies on facial component detection, adapted from [5]*

- Object Shape and Appearance Detection

Table 2.4 summarizes the object shape and appearance detection in terms of feature and its respective method. The histogram of oriented gradients (HOG) was applied to extract the features from the distribution of intensity gradients and edge direction. Since it operated on local cells, HOG can be extremely effective at describing changes appear in large spatial regions except for object orientation. On the other hand, the modified Hough transform (MHT) [6] could be used to detect the circular arcs in the image, normally the helmet shape in crime scenes.

Object and appearance feature	Method used	Author, year
Large spatial regions	Histogram of oriented gradients, HOG	Tripathi et al., 2016
Circular arcs	Modified Hough transform method, MHT	Wen et al., 2003

*Table 2.4 Studies on object shape and appearance detection, adapted from [5]*

- Decision Making

After all, decision making will be made through either trained classifier-based system or untrained classifier-based system. Table 2.5 summarized the decision-making and methods used. For trained classifier-based system, the machine learning classifiers had been widely applied in image classification, such as SVM, Nearest neighbour (NN) classifier, random forest classifier and etc by developing fast and efficient algorithms to analyse vast volumes of data.

For untrained classifier-based system, threshold value comparison was greatly used for abnormality analysis due to its simplicity, the threshold values could be calculated using a formula or estimated based on a set of conditions. On the other hand, decision could also be made through the relative location, distance and geometric analysis.

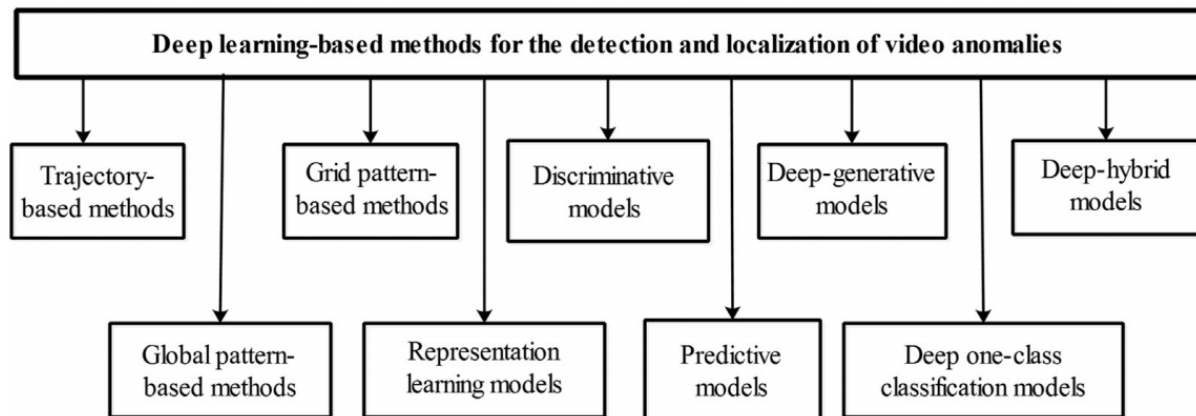
Decision making	Method used	Author, year
trained classifier-based system	Support vector machine, SVM	Yoon et al., 2002
		Kim et al., 2005
		Suhr et al., 2012
	Nearest neighbour (NN) classifier	Min et al., 2014
	Random forest classifier	Tripathi et al., 2016
	Threshold value comparison	Zhang et al., 2018
untrained classifier-based system	Threshold value comparison	Wen et al., 2003
		Sako et al., 2007
		Kim et al., 2010
		Min et al., 2011
		Goswami et al., 2015
	Geometric analysis approaches	Wen et al., 2005
	Relative location of facial feature determination	Dong et al., 2006

*Table 2.5 Studies on decision making, adapted from [5]*

In conclusion, the findings of this study show that detection method in criminal activity was diverse. Thus, it should be chosen carefully depending on reliability and suitability for corresponding objectives. However, the system framework must be followed step by step, starting from the image pre-processing to the decision making in order to introduce a complete and robust criminal detection surveillance system.

### 2.1.3 Deep-Learning Based Methods for Video Anomaly Detection

Since deep learning was introduced, it might be easier and faster to create and deploy than prior types of method, such as machine learning. Thus, this section would be limited to the video anomaly detection approach based on deep learning. The state-of-the-art of the deep-learning-based methods used for video anomaly detection could be categorized as shown in Figure 2.5.



*Figure 2.5 Classification of deep learning-based methods for the detection and localization of video anomalies [7]*

- **Trajectory-based methods**

In the trajectory-based changes detection, the object of interest was first recognized and tracked over the frames in order to create the trajectory. The accuracy of this method is highly dependent on item recognition and tracking, it was better suited for the sparsely populated environment since the tracking performances would be affected by occlusion, low video resolution, and crowd density.

- **Global pattern-based methods**

For the global pattern-based approaches, computational features such as optical flow, spatial temporal gradients, kinetic energy, and others in the frame sequences would be analysed as a whole entity. This method was found to be helpful in anomaly detection as it was suitable. However, the localization of the anomalies would be a tedious job using global pattern-based method. As a result, due to the lack of object detection and tracking, it was better suited to moderately populated as well as heavily congested environments.

- **Grid pattern-based methods**

In the other way round, only limited features from a fixed spatiotemporal region (grid or sub-region) were extracted in this method in order to reduce the processing time. For instances, anomalies in each grid, cell, or sub-region were examined independently, with the links between the items being ignored. Furthermore, rather than interpreting the frames as a single entity, patterns would be retrieved from the splitting blocks of each frame.

- **Representation learning models**

In the case of representation learning models, useful features (representations) of input video data were extracted and learned by the machine itself. The following methods were based on representation learning:

- **Sparse coding inspired deep neural networks**

The sparse coding inspired deep neural networks first learnt a dictionary from the unlabelled datasets containing only normal events and eventually differentiate the abnormal events that could not be reconstructed by the atoms of the learned dictionary.

- **Reconstruction models**

A training video dataset comprising only normal events would be used for reconstruction models to learn the normal features and behaviours. Subsequently, the anomalous event would cause high reconstruction error as they do not comply with the learned model. In this way, the anomalous activities could be detected through the reconstruction error.

- **Slow feature analysis**

Slow feature analysis was an unsupervised learning method to extract the smoothest (slowest) underlying features or representations of the rapid varying high dimensional input. Due to its low computational complexity, this method could be suitable for system which aimed with fast processing-time.

- **Discriminative models**

In the case of discriminative models, the discriminant features of the anomalous and normal event would be learnt in order to model the decision boundaries between the classes [8]. Thus, the discriminative models were found to be suitable for supervised learning technique as it provided well labelled and balanced dataset.

- **Deep-generative Predictive models**

Generative model was similar to the discriminative model but the discriminative model focused on modelling the decision boundary between the classes, whereas the generative model focused on modelling the actual distribution of each class [8]. In addition, this model was able to address the data imbalance problem unlike discriminative model which was not designed for unlabelled data. Thus, it was widely used to detect video anomalies.

- **Predictive models**

The objectives of predictive models were to practice statistical technique and model the conditional distribution for predicting the future outcomes. These were the solutions in terms of data mining technology with high computational complexity to analyse past and recent data which provided a certain pattern and produce a model to identify the future behaviour.

- **Deep one-class classification models**

One-class classification (OCC) involved fitting a model on the normal event due to the confusing nature of abnormalities and insufficiency of the respective ground truth data. Generally, the Deep OCC model combined the hierarchical feature representation capability of the DNN with the OCC objective such as hyperplane or hypersphere, the deviations from this description were then deemed to be anomalies. However, a huge training time are required for these models, especially for the high dimensional data visualization such as video stream [4].

- **Deep-hybrid models**

Since there was no a single solution could handle the complete task perfectly for video anomaly detection, multiple methods were then be comprised as for taking advantages from each method. Recently, the state-of-the-art for the video anomalies detection using Deep-hybrid models has been proposed with high performance and accuracy. However, the hybrid model approach was suboptimal due to the disadvantages of complicated architecture [4].

### 2.1.4 Comparative Analysis of Video Anomaly Detection Method

Any of the handcrafted/ machine learning/deep neural network-based video anomaly detection method had its own advantages and disadvantages. However, the selection of the most appropriate technique for a particular application was still a crucial aspect. The following Table 2.6 summarize the selection of the techniques based on various aspect.

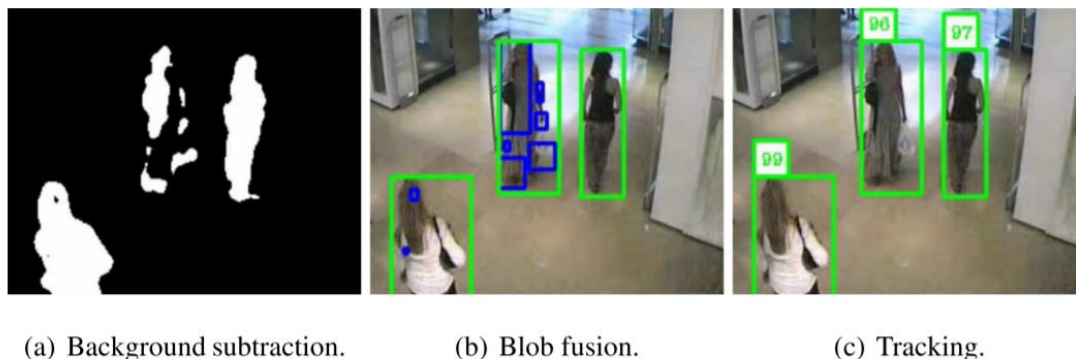
Learning	<p>Training and learning framework should be chosen based on the number and quality of the datasets available.</p> <p>Supervised learning: Labelled data</p> <p>Unsupervised learning: Unlabelled data</p> <p>Semi-supervised learning: Weakly-labelling data</p>
Approach	<p>An appropriate approach was chosen based on the required precision and application</p> <p>AO approach: Offline applications</p> <p>PO approach: Online applications</p>
Method	<p>The selection of the ML/DNN-based video anomaly detection techniques on the available datasets, targeted applications, time complexity and expected performances.</p>

*Table 2.6 Selection of video anomaly detection methods*

## 2.2 Automated Surveillance System

### 2.2.1 Real-Time Detection of Suspicious Behaviours In Shopping Malls [9]

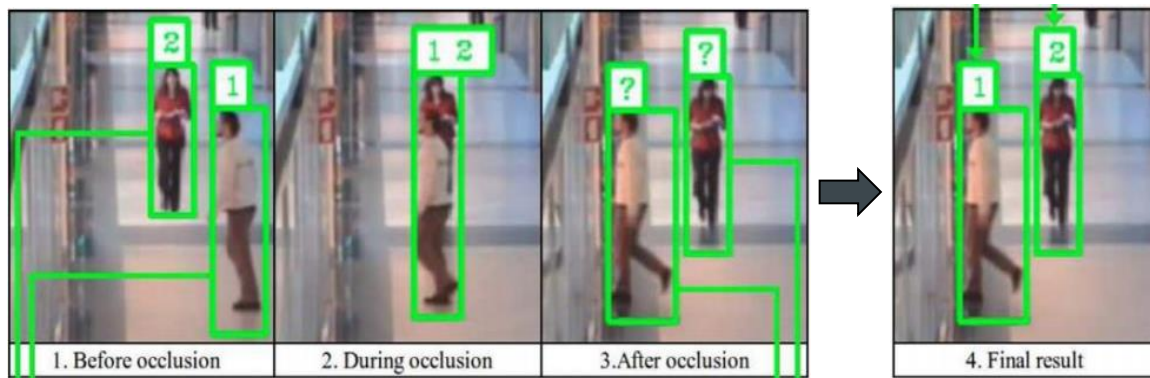
Arroyo's work centred on a specific use of video surveillance: detecting potentially suspect human behaviour in shopping malls such as running away, loitering, unattended cash desk without people nearing. The methodology implemented in the video pre-processing and human tracking were firstly *Background subtraction*, followed by *Blob fusion* to correct the imperfection of recognition in foreground object that caused by segmentation error, thirdly *LSAP association* [10] and Kalman filtering [11] were employed for object tracking, as shown in Figure 2.6.



**Figure 2.6 Video pre-processing and human tracking (Arroyo et al., 2015)**

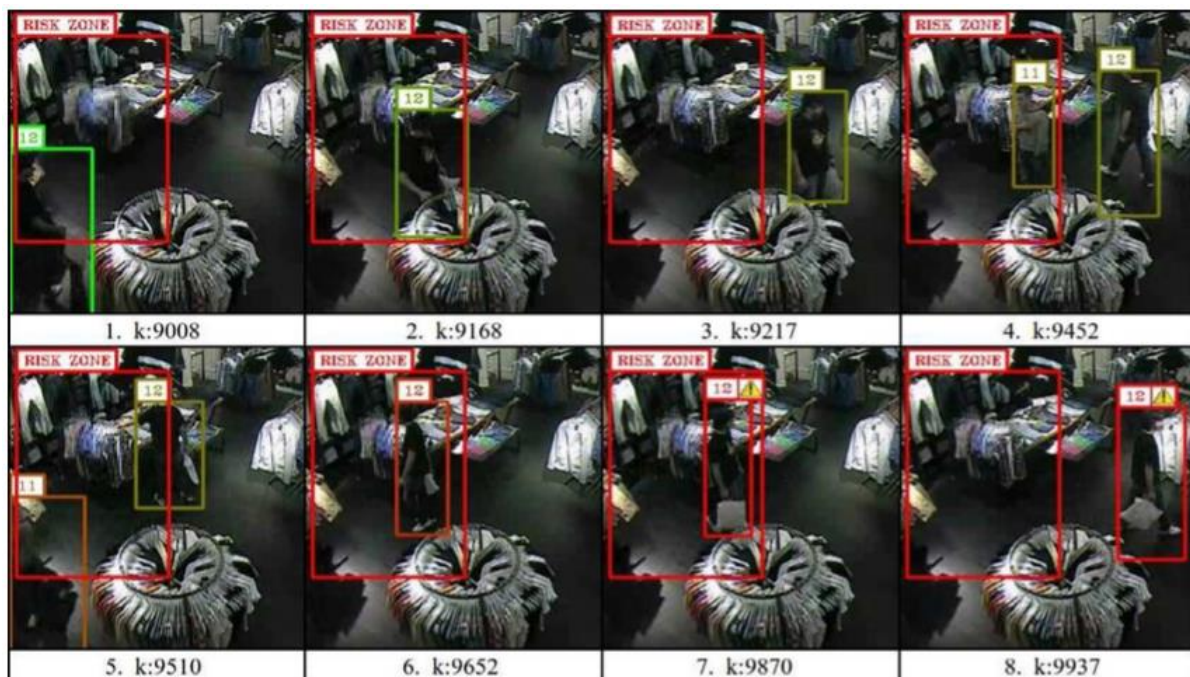
When a tracked object was in an occlusion situation, the object could not be re-identified with the technique of the proposed tracking method based on Kalman predictions and LSAP association. Thus, *Occlusion management* algorithm based on visual appearance would complement the LSAP solution and the combination gave a powerful result on object tracking. Furthermore, three image descriptors: GCH, LBP, HOG were used for investigating the human visual appearance and some SVM kernel functions were also implemented with the goal of comparing the person's appearance features before and after an occlusion. While, SVM was one of the discriminative classifiers used for classification [12]. The example shown in Figure 2.7. visualized the result of the proposed method by a simple two-people occlusion.





*Figure 2.7 Occlusion management [9]*

In order to detect the suspicious behaviours such as loitering, the human security officers would predefine a specific risk zone by marking a rectangle box, and system started to examine the tracked individuals' trajectories to see whether any of them have been loitering in the high-risk zone for an extended period of time. Finally, the level of suspicion was represented by a colour gradient between green and red as shown in Figure 2.8.



*Figure 2.8 Loitering detection [9]*

In summary, the tracking method with handling occlusion outperformed the related works on their research due to the addition of visual appearance information in occlusions management. However, the use of stereo cameras rather than monocular vision could be an attractive

enhancement for improving tri-dimensional data extraction. In addition, to trigger the alarm when suspicious behaviour was recognised, their system still required the supervision of a human operator because there were no concrete related works that consider the precise type of suspicious behaviours in stores. As a suggestion to the future main upgrade, a powerful machine learning algorithm should be applied to completely automated the entire surveillance jobs.

### **2.2.2 Fast and Robust Occluded Face Detection in ATM Surveillance [13]**

In ATM related crime, suspects often partially occluded their face which become an obstacle for many approaches, while this action also considered the abnormal activities happen at ATM. Therefore, the researchers targeted the occlusion management in scenes, and presented a reliable face occlusion detection algorithm to determine if an observed face was occluded.

To combat the suspects with covered facial component, the researcher recommended using the Omega shape as human's head and shoulder for accurate head location. Because most of the previous methods relied on the presence of facial components, which were unavailable due to extreme blockage, a new strategy was needed. A new energy function for elliptical fitting models was created.

For head tracking., the shape cues and gradient were used in a Bayesian framework. The method was then implemented via Monte-Carlo simulation which may be approximated by a finite set of correctly weighted particle samples, because the probability was unable to be represented directly through the Bayesian model [14].

Finally, to determine whether or not a face was occluded, the researcher proposed to combine a skin colour classifier using three colour spaces, i.e., RGB, YCbCr, and HSI colour space and a face template matching classifier using the AdaBoost algorithm. This aided in addressing issues such as skin colour variance due to different people. The face occlusion detection results were shown in Figure 2.9.



*Figure 2.9 Examples of face occlusion detection results [13]*

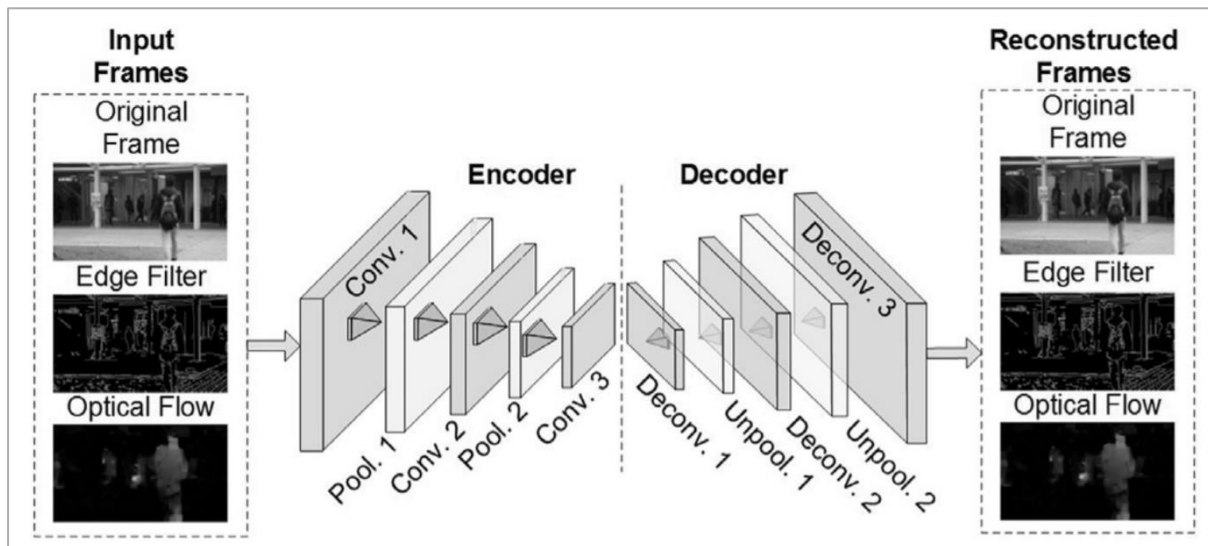
In summary, a simple but reliable system with head tracking and face occlusion detection were constructed to identify suspicious criminals who wear helmet, mask or scarf to hide their face in order to avoid being recognised by video monitoring systems. However, the proposed method should be integrated with some other supporting feature such as action recognition i.e., peeping password, beating machine to ensure a fully functional crime detecting system.

### **2.2.3 A Study of Deep Convolutional Auto-Encoders for Anomaly Detection in Videos [15]**

The Convolutional Auto-Encoders (CAEs) had been popularly used in reconstruction model of the deep learning approach when there was no general rule for the definition of anomalous event. In this case, their anomaly detection model might be viewed as a one-class classification issue, with human defining the normal classes in a high number of samples. Then, the classes that do not conform to the model would be considered as the anomaly class.

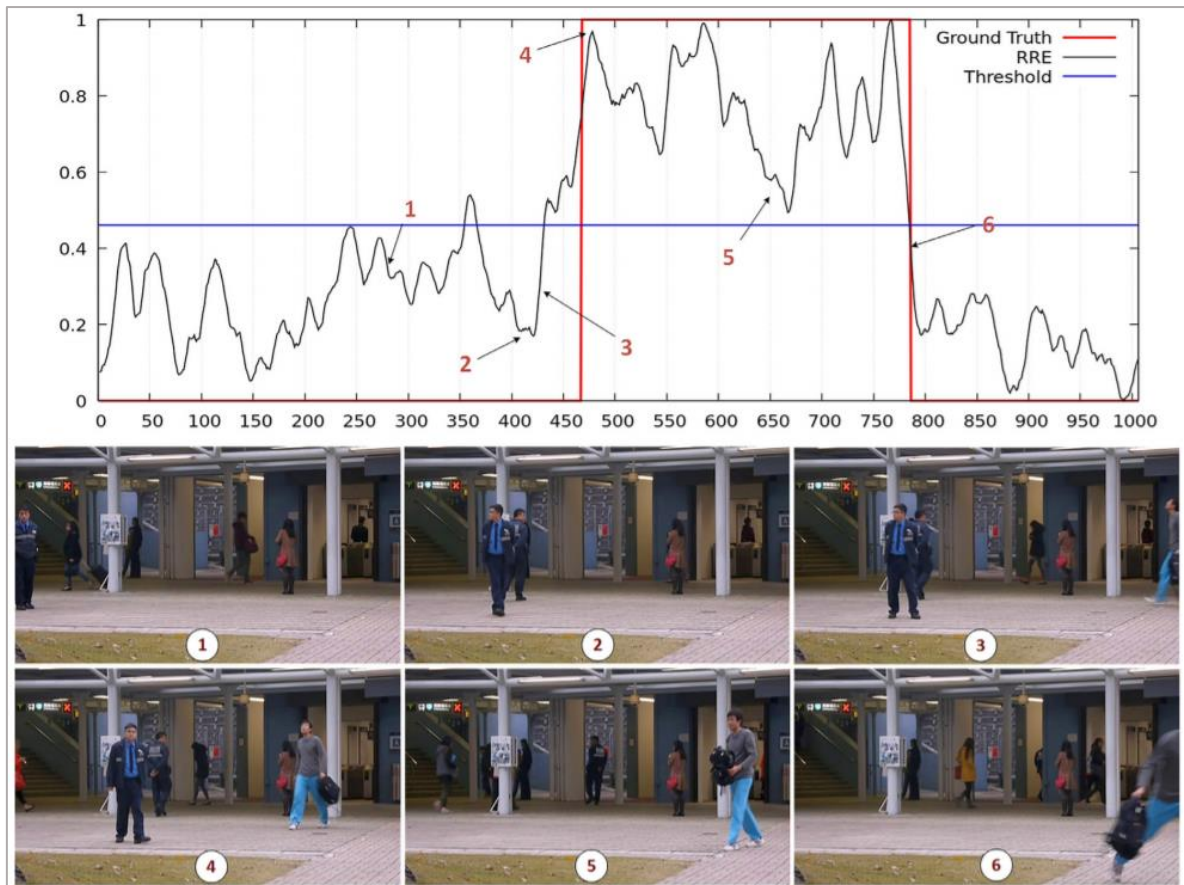
In the year of 2011, Masci et al. proposed the CAE architecture, which generally applied in the task of image reconstruction. As the Figure 2.10 shown, there was two parts in the autoencoder: the encoder and the decoder. The encoder learnt to extract the information and

compress it to an internal representation in what is called the Convolution layer, so that it can be used to recreate the same input in the decoding part. In an end-to-end training framework, a deep CAE was trained to keep the spatial information of the video sequence during encoding dynamics.



**Figure 2.10 Architecture of the proposed CAE [15]**

Therefore, CAE combined with high-level spatial and temporal information was proven to be effective in detecting contextual video anomalies. In the other way round, the Canny edge detector [16] was also applied to extract the appearance features, while the state-of-the-art optical flow algorithm was used to extract motion between two consecutive frames. In the data preparation stage, the video would be extracted with both respective appearance and motion features. These features and frames were then merged to create several scenarios (case studies) that characterize the input data to the CAE, the CAE subsequently learned the signature of the normal event in the training phase. While the experimental result was shown in Figure 2.11.



*Figure 2.11 Reconstruction error on detecting anomalies [15]*

*1-2. normal event, 3. Unusual movement detected (man entering scene), 4-5. The man throws up a backpack (behaviour no presented in the training set), 6. The man vanishes from the scene*

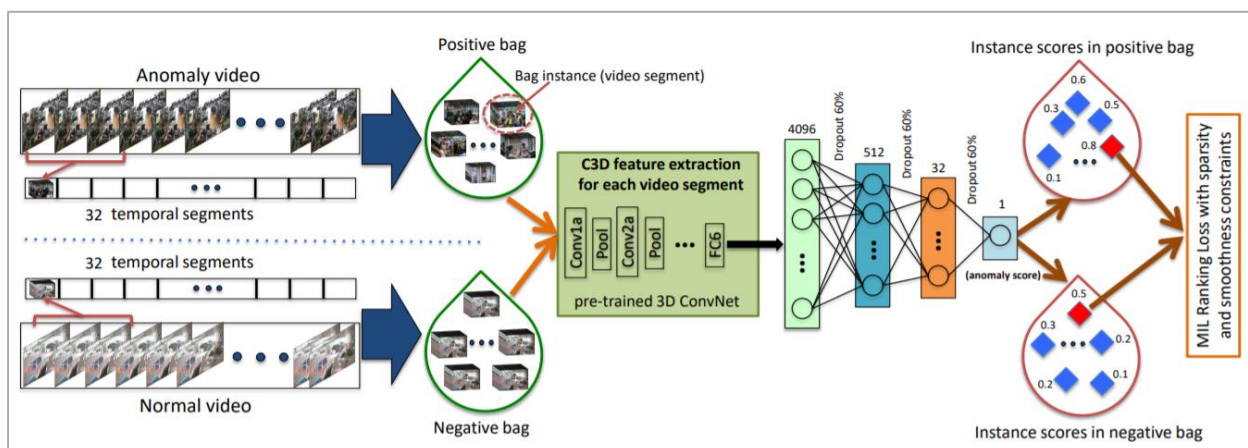
In summary, the proposed model successfully detected the ‘anomalous events’ which were not included in the training data. However, the nature of anomaly were still an ambiguous field, thus resulting in more false alarm in a particular subject issue. Thus, a suggestion for future improvement were to work with more real-world datasets and construct a more formal and comprehensive definition of normality and anomaly in order to support and improve the proposed method.

## **2.2.4 Real-world Anomaly Detection in Surveillance Videos [3]**

This research was one of the popular recent studies utilized predictive model in deep learning approach. Sultani et al (2018) proposed to learn anomalies such as fighting, shoplifting, crimes and illegal activities by exploiting both normal and anomalous videos. A deep learning

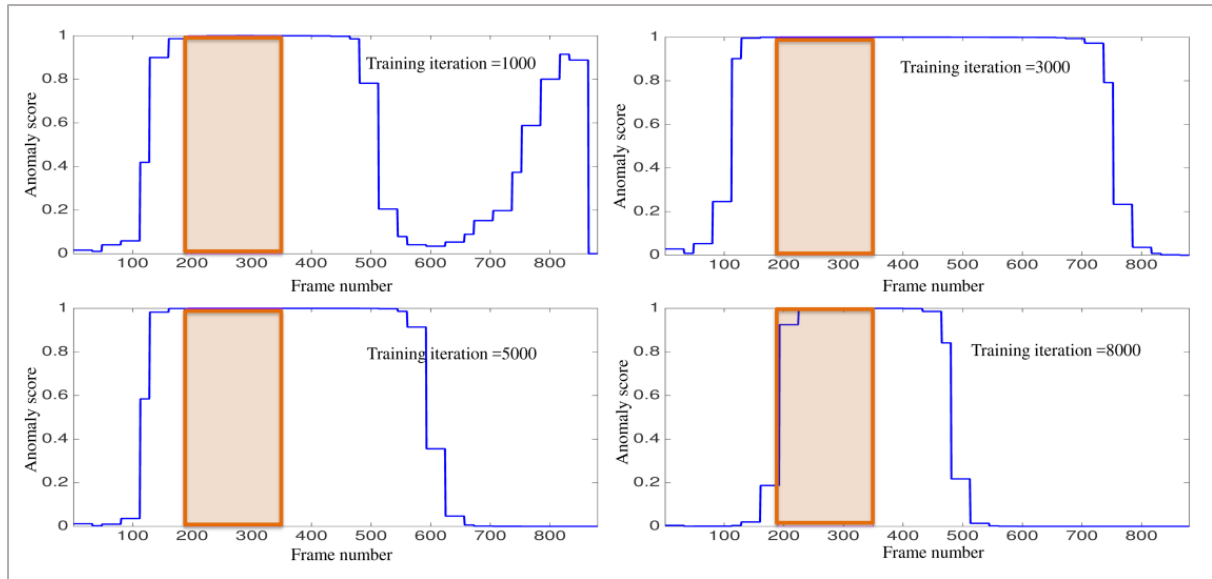
approach was introduced to predict and identify real-world surveillance video abnormalities which aim to increase the public safety.

By referring to the Figure 2.12, the proposed approach started with breaking the surveillance footage into a set number of segments during the training stage. In multiple instance learning (MIL), training sample came in bags; a positive bag (anomalous) and a negative bag (normal) whereby the different temporal segments/clips of each videos represented an individual instance in the bag. After extracting C3D features for each video segment, a fully connected neural network was trained by using a novel ranking loss function which computed the ranking loss between the highest scored instances in the positive bag and negative bag (shown in red). The highest anomaly score in the positive bag represented the most likely anomalous segments (true positive), while the highest anomaly score in the negative bag represented the most likely anomalous segments but actually a normal segment (false positive).



**Figure 2.12** Flow diagram of the proposed anomaly detection approach [3]

The network was then learned from numerous positive and negative bags. Figure 2.13 depicted the progression of anomaly score for a training anomalous sample over the iterations, given blue line indicating the anomaly score predicted by the network and coloured window representing the actual anomalous region in the video. At the beginning, the network generated high anomaly scores for both anomalous and normal video segments. After thousands of iterations, the result became desirable whereby the network able to keep the anomalous segments on high score and produce low score to non-anomalous segments. The network was expected to learn from more instances and localize anomaly precisely as the number of iterations increased.



**Figure 2.13** Evolution of score on training video over iterations [3]

Their experimental result had proved that their MIL solution achieved significant improvement on anomaly detection performance when compared to the state-of-the-art approaches. Apparently, their approach got a substantially lower rate of false alarms compared to the other approaches, and this validated that the deep learning model was able to learn more general normal patterns by employing both anomalous and normal videos for model training. In addition, their dataset for anomalous activity recognition was special and challenging, due to the long and untrimmed videos with intra-class differences, thus they provided results of baseline methods, C3D [17] and TCNN [18] on recognizing thirteen different anomalous activities to overcome the labour-intensive temporal annotations on videos. However, there is still a room for improvement since their method failed to detect the anomalous event due to the occlusion problem and darkness of the captured scene. Intensity enhancement and occlusion management were suggested to be implemented in order to develop a complete and robust real-time anomaly detection in surveillance system.

## 2.2.5 Summary of the Reviewed System

	<b>Arroyo et al. (2015)</b>	<b>Zhang et al. (2018)</b>	<b>Ribeiro et al. (2017)</b>	<b>Sultani et al. (2018)</b>
<b>Human detection &amp; Tracking</b>	✓	✓	✓	✗
<b>Occlusion handling in human tracking</b>	✓	✗	✗	✗
<b>Occluded face detection</b>	✗	✓	✗	✗
<b>Illegal object detection</b>	✗	✗	✗	✗
<b>Specific area warning</b>	✓	✗	✗	✗
<b>Suspicious behaviour detection</b>	✓	✗	✓	✓
<b>Machine learning applied</b>	✓	✓	✗	✗
<b>Deep learning applied</b>	✗	✗	✓	✓

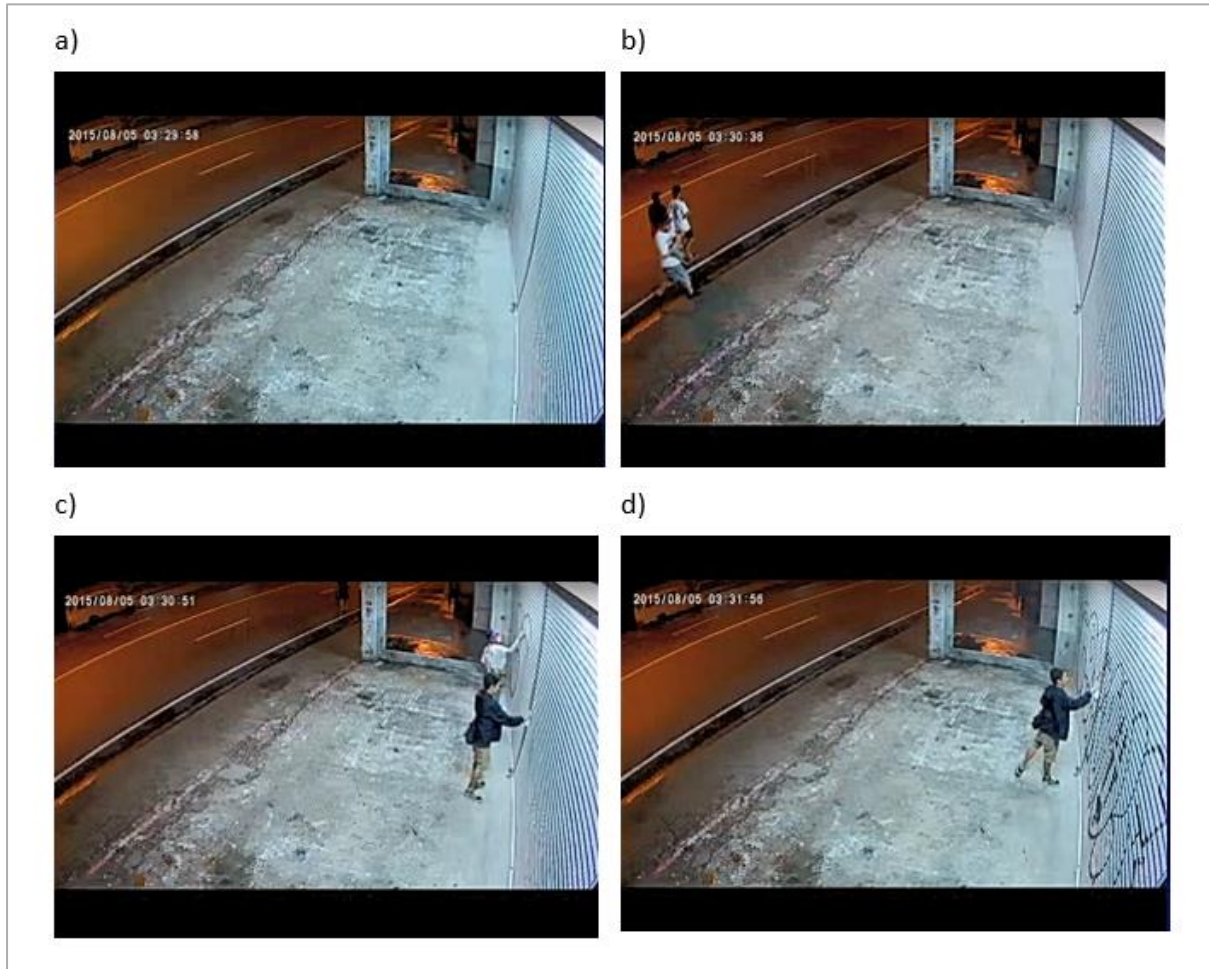
*Table 2.7 Table of Comparison*



## Chapter 3 : System Methodology/Approach

### 3.1 Real Case Scenario

#### 3.1.1 Scenario A

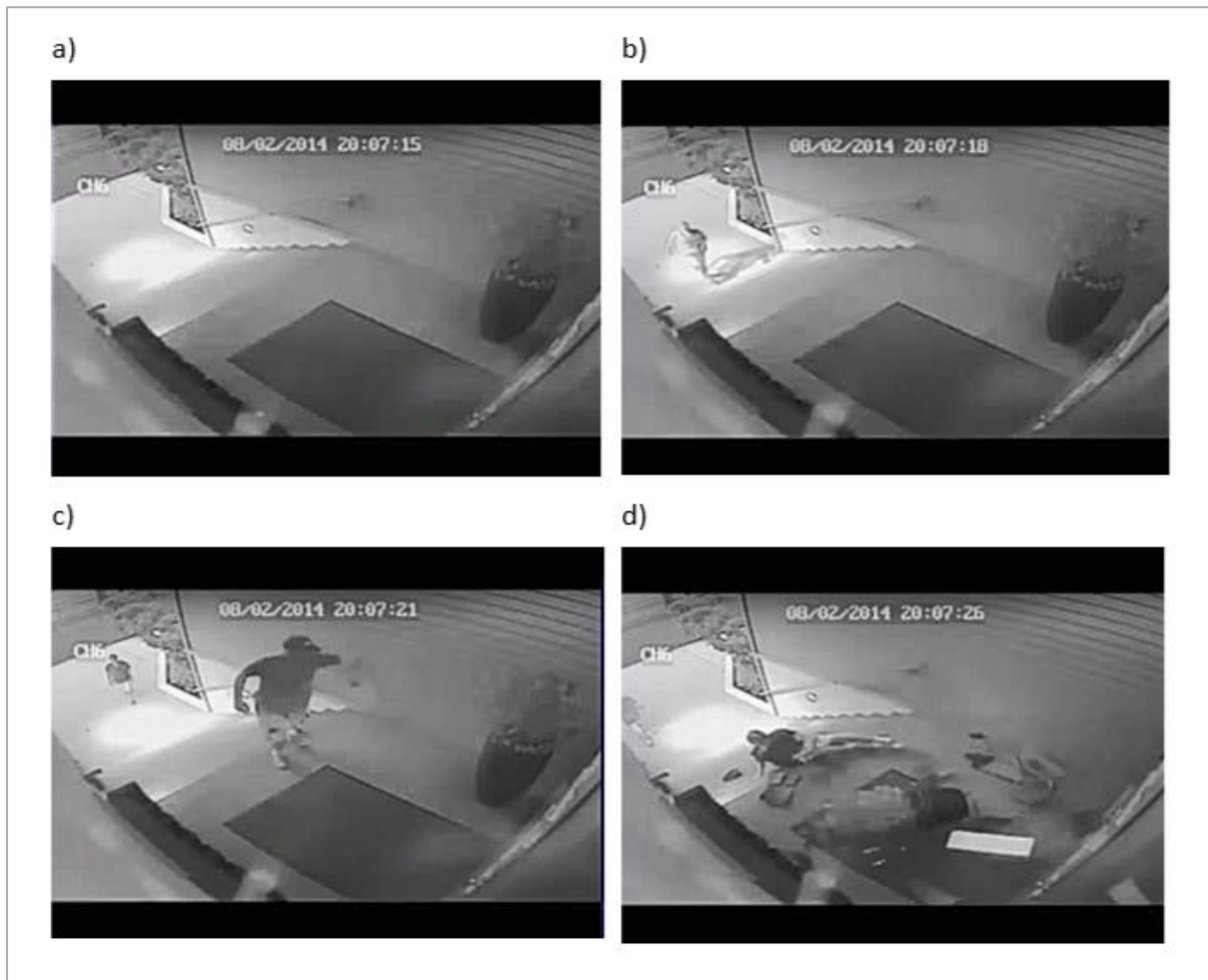


*Figure 3.1 Video sequence of scenario A*

*Category: Graffiti*

- a) Normal event
- b) Some pedestrians walk into the scene.
- c) Two men approach to the store gate and start painting illegally.
- d) A large graffiti art is leave on the store gate.

3.1.2 Scenario B

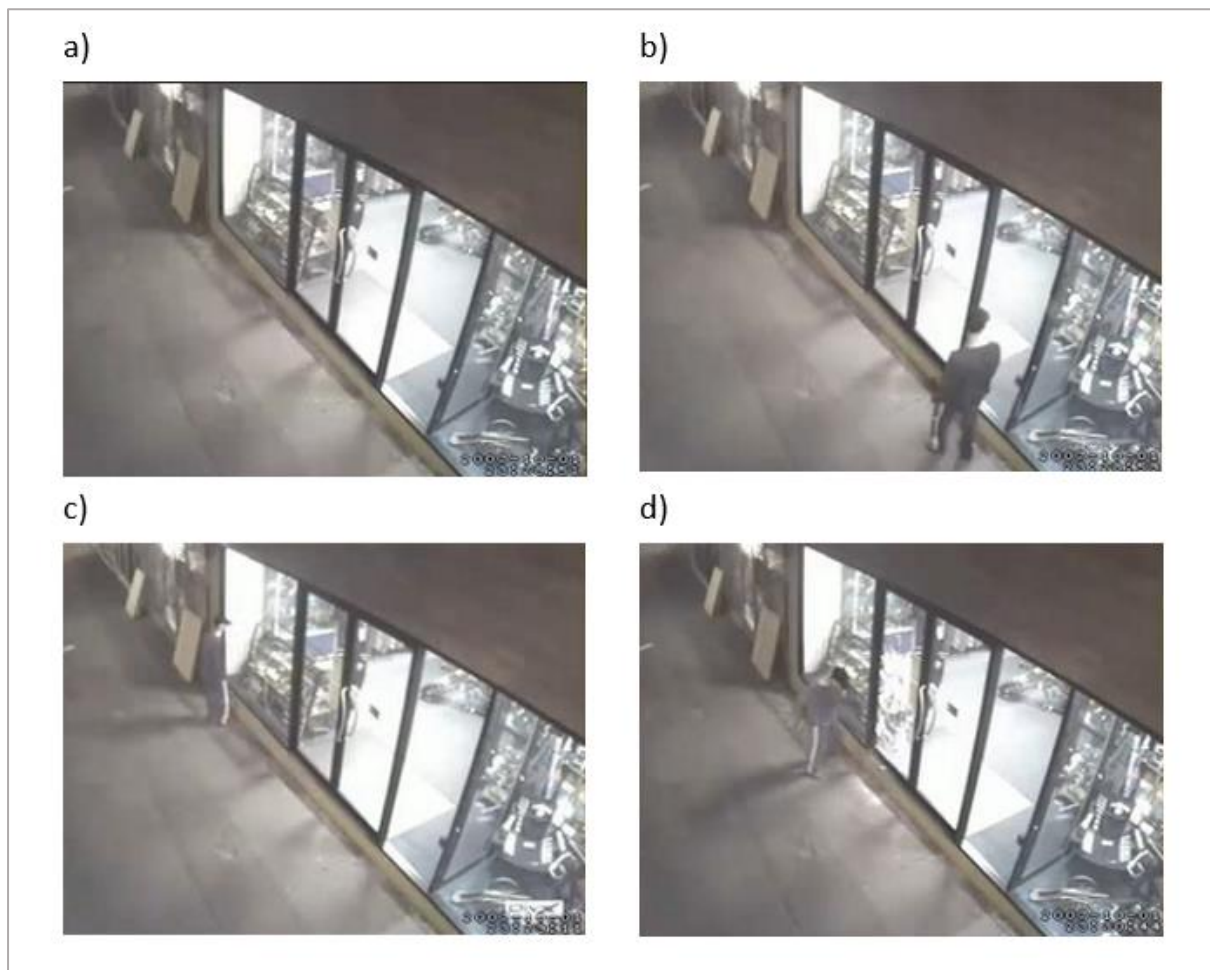


*Figure 3.2 Video sequence of scenario B*

*Category: Defacing of property*

- a) Normal event
- b) A man runs into the scenes.
- c) The man stays at the doorway for a period of time.
- d) The man kicks the vase and make a huge destruction.

### 3.1.3 Scenario C



*Figure 3.3 Video sequence of scenario C*

*Category: Glass breakage*

- a) Normal event
- b) A man walks into the scenes.
- c) The man loiters in front of the store for a period of time.
- d) The man smashes the window and the shattered glasses are captured by the scenes.

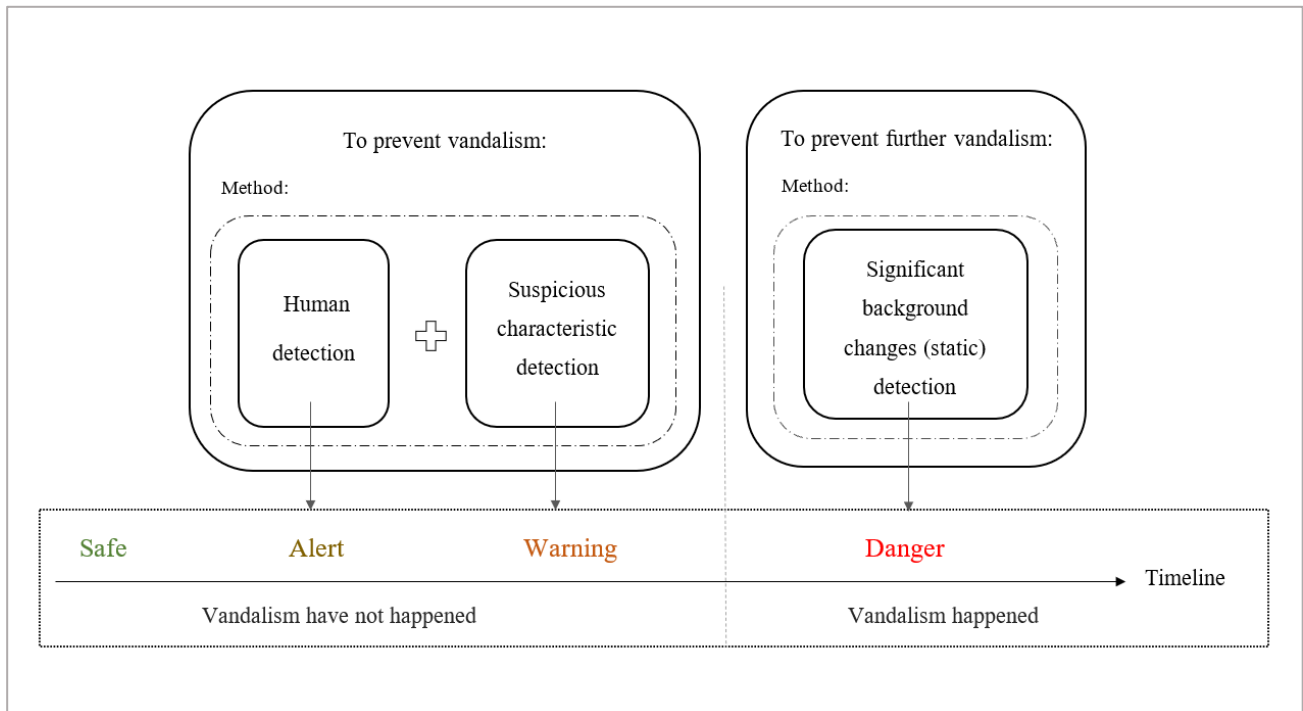
### 3.1.4 Discussion

Based on the three scenarios:

- The acts of vandalism are generally made by human.
- Some vandals would loiter at the targeted area with intent of committing a crime.
- The vandalism will mostly happen in non-crowded scenes such that the others would not notice the vandalism event.
- The vandalism act will cause a consistent background change in the scene which imply the destruction of property that is supposedly unchanged in a normal event.
- The vandalism event happens fast. Thus, fast approach with high detection rate is needed to give an instant alert to the security centre or owner.

## 3.2 Design Specifications

### 3.2.1 Methodologies and General Work Procedures



*Figure 3.4 Methodology on vandalism detection*

The proposed methodology of the project is illustrated in Figure 3.4. Most of the vandalism take place in isolation or at low traffic moments, and we assume that the acts of vandalism are generally made by human. Therefore, the **human detection** is required likely as an off-the shelf “motion sensor light” to alert any intruder that they are under monitoring. In addition, it will be complemented with **the suspicious characteristic detection** such as loitering detection and give out early warning to notify the security centre or the owner.

On the other hand, when the **significant background changes** happen in the scenes, indicating a permanent damage, an alarm will be given off to prevent further vandalism in order to cease the destruction of the properties.

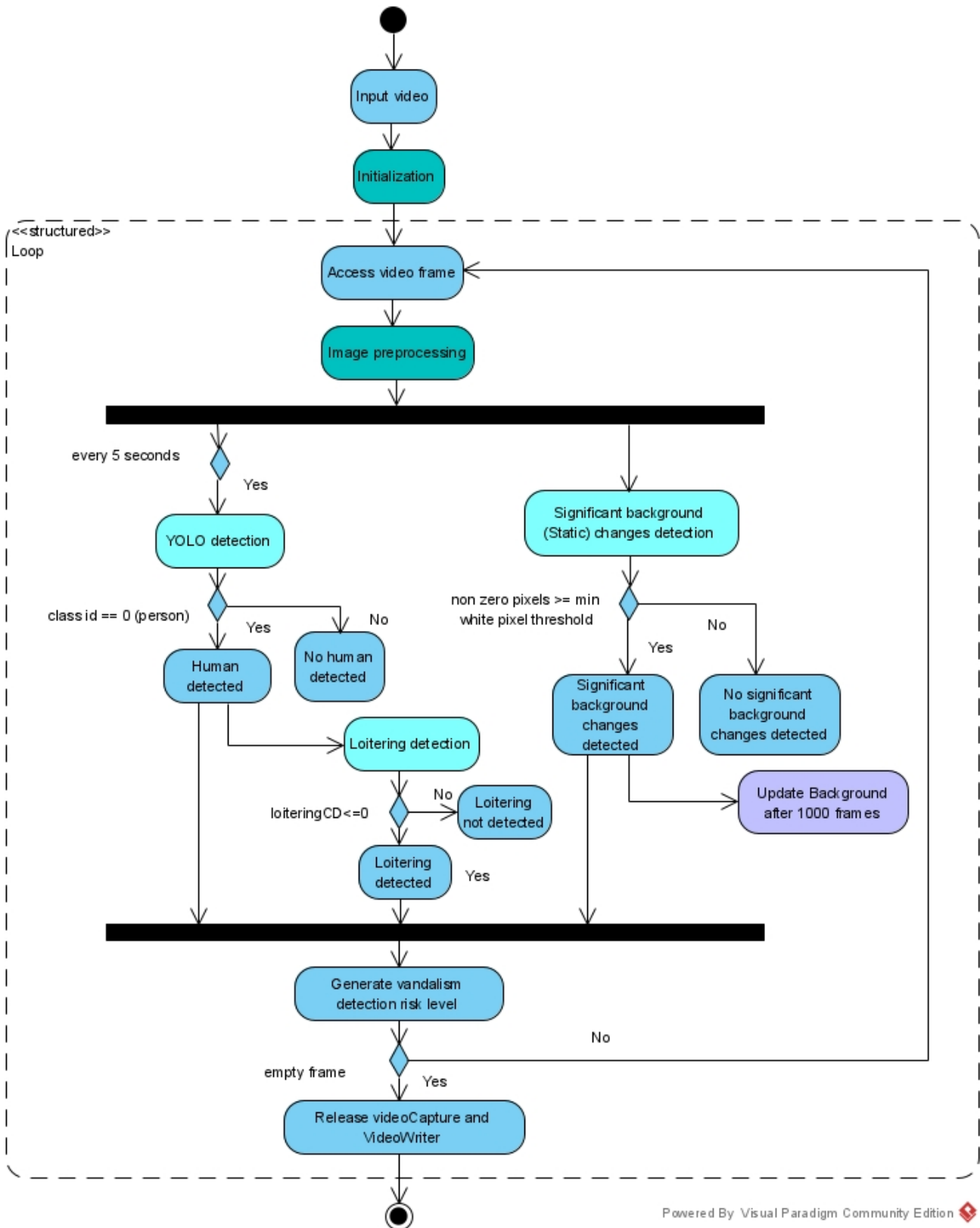
For a better understanding, the system is described as a two-level defence system, the first level is to detect the potential vandals in order to prevent the vandalism event from happening. On the other hand, the second level defence is to detect the vandalism event by observing any permanent damages captured by the scenes, such that early warning can be given off to prevent further vandalism.

### **3.2.2 Assumptions**

- I. The camera is assumed to be static and non-moving.
- II. The vandals must be captured into the scene so that the human and suspicious characteristic detector could work.
- III. The vandalism-prone objects must be captured into the scene so that the significant background changes can be detected.
- IV. The captured damages/destruction of vandalism-prone object must be significant enough to indicate a vandalism event.

Chapter 4 : System Design

4.1 System Design / Overview



Powered By Visual Paradigm Community Edition

Figure 4.1 System diagram

Figure 4.1 shows the System diagram which processes the vandalism video analysis. The system starts with capturing the video and proceed to the initialization. Then, each of the frame will undergo the image pre-processing.

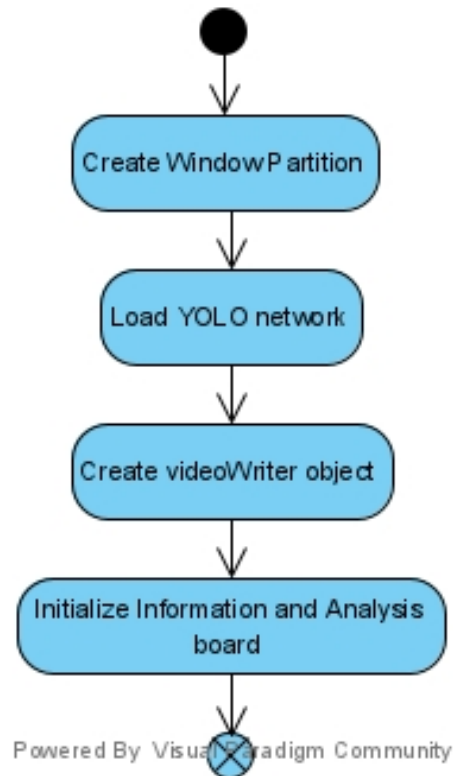
The image sequence will go through three modules such as **YOLO detection** to detect human, **Loitering detection** to detect suspicious behaviour i.e., loitering and most importantly the **Significant background (static) changes detection** to detect significant background changes (static) which may represent the permanent damage caused by vandals in the scene. Lastly, the risk level for detected suspicious vandal and vandalized event will be generated at the analysis board.

In the following sub chapters, the Initialization and Pre-processing, YOLO detection, Loitering detection, Significant background (static) changes detection, and Background estimation will be explained in detail.



## 4.2 System Components Specifications

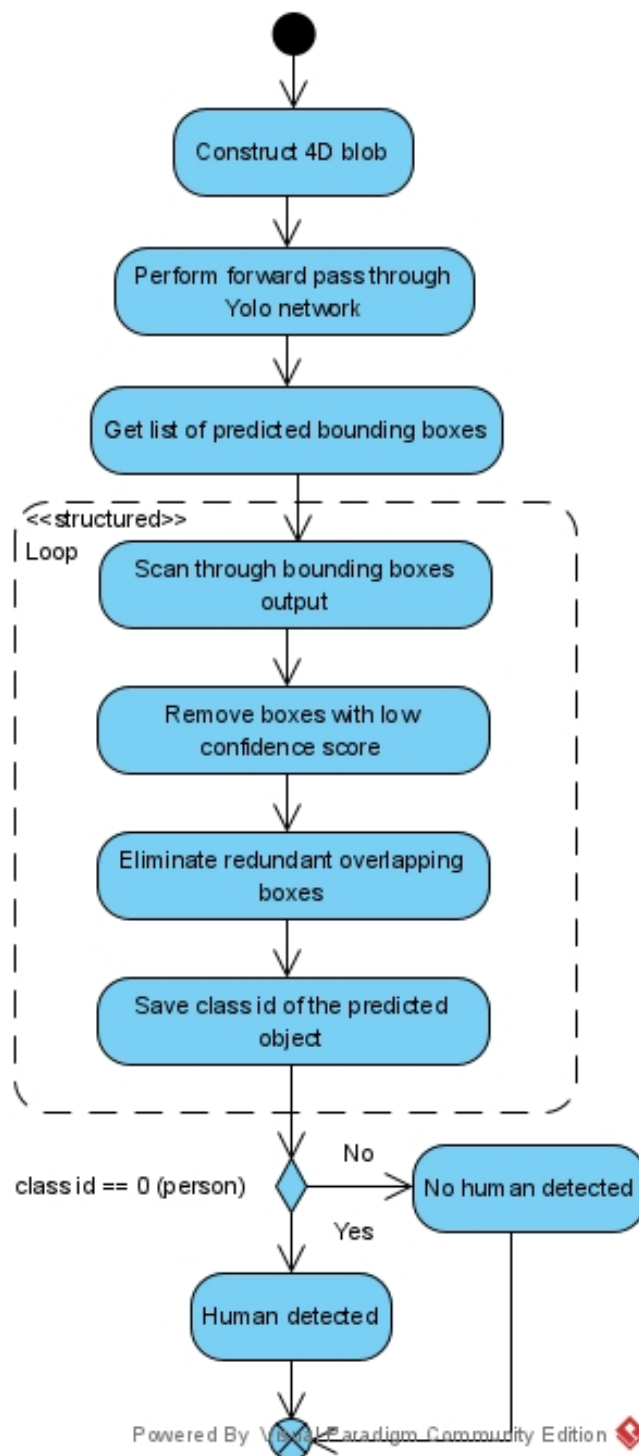
### 4.2.1 Initialization and Pre-processing



*Figure 4.2 Initialization*

The initialization included the creation of window partition, network loading for YOLO, creation of videoWriter object, and information and analysis board initialization. In addition, each of the frame will undergo the image pre-processing, such as converting image to grayscale, HSV, and resizing image. The system was built to deal with low-resolution images whereby the highest resolution on either width or height will be limited to 410 pixels in order to improve the processing time.

## 4.2.2 YOLO Detection [19]



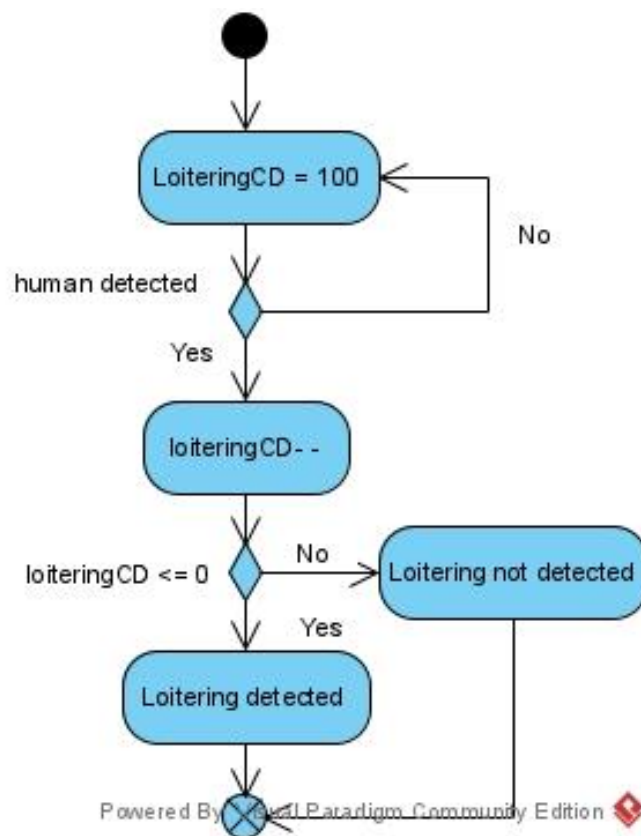
**Figure 4.3** YOLO detection

You Only Look Once (YOLO) that achieve near state-of-the-art results with a single end-to-end deep learning model will be chosen to detect the human in the scenes, as it performs a comparative fast detection compare to other method. The published YOLOv3 with pre-trained

weights will be used in this project[20]. Still, the algorithm required high computational resources. Therefore, YOLO detection will be configured to run at every 5 seconds to detect human.

The input frame to the neural network needs to be in a certain format called a blob, thus, the `blobFromImage` function will be used to convert the image to an output frame. When the blob is prepared, it will then pass into the network, and a forward pass is performed to get a list of predicted bounding boxes as the network's output. After that, the output boxes will be scan within a loop in order to eliminate redundant overlapping boxes with low confidence scores. Last but not least, the class ID of the detection with high confidence will be extracted for subsequent decision making, i.e., class ID of object "person" is 0.

### 4.2.3 Loitering Detection



**Figure 4.4 Loitering detection**

Once the human is detected, the system will start to countdown from 100. For every frame, the variable will be reduced by one until it reaches zero and once it is less than zero then the flag for loitering will be set from “alert” to “warning”. However, if the subsequent frame shows that there is no human detected, the variable will be reset back to 100, and remove the flag for loitering behaviour.

4.2.4 Significant Background Changes Detection

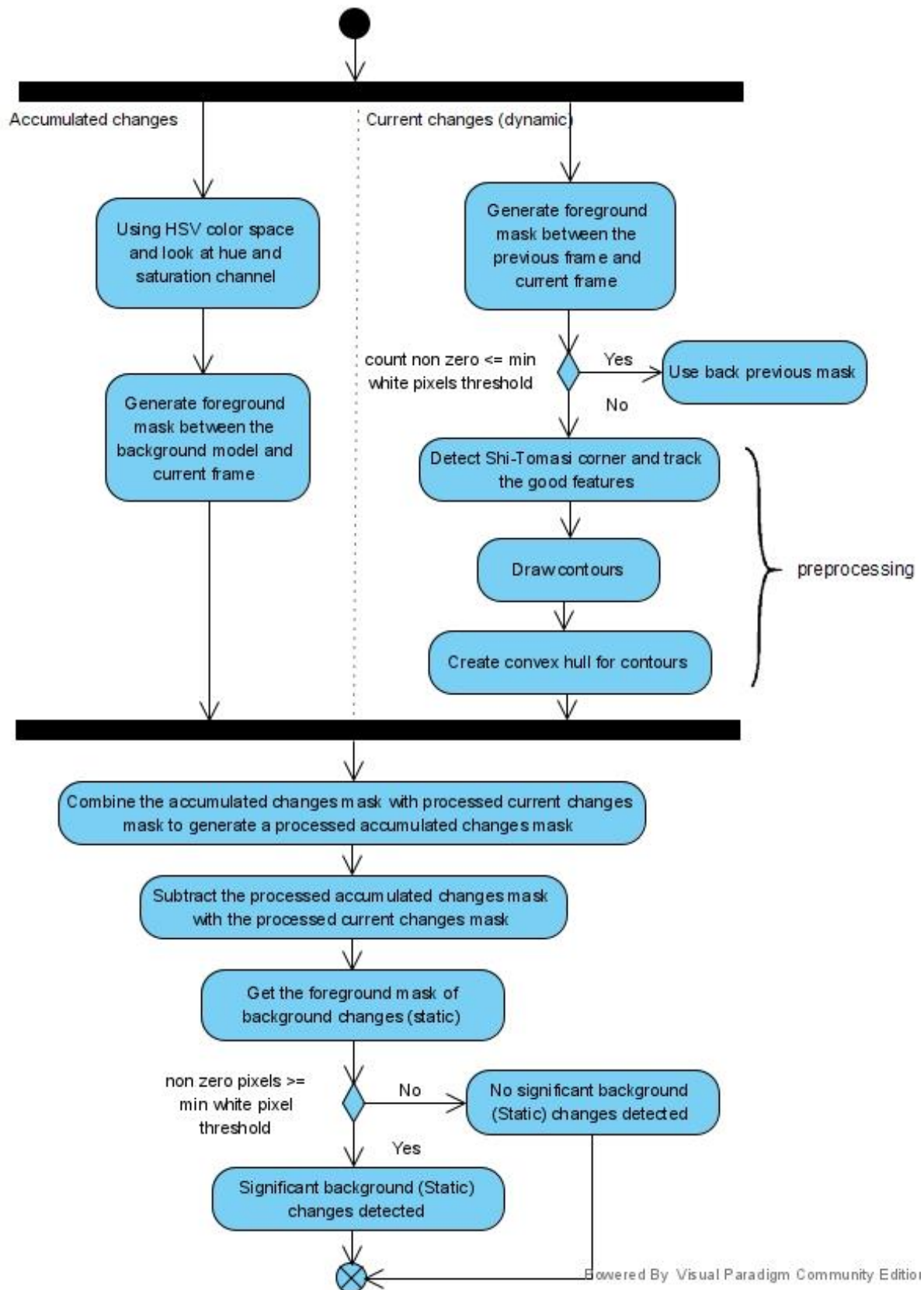


Figure 4.5 Significant background changes detection

The accumulated changes and the current changes (dynamic) in the scenes will be extracted in order to compute the background changes later:

- i) accumulated changes and the current changes (dynamic) can be calculated as:

$$\text{background model} - \text{current frame} = \text{accumulated changes}$$

$$\text{previous frame} - \text{current frame} = \text{current changes}$$

For extracting the accumulated changes, the approach is to use colour space that separates luminance from hue and saturation, i.e., using HSV colour space [21]. Since, the real-world videos might prone to noise based on bad weather condition and bad lighting. Thus, using just hue and saturation channel can certainly avoid influence of value variations (light). Then, the foreground mask (accumulated changes) can be generated between the background model and current frame.

On the other hand, the foreground mask of current changes can be extracted from the subtraction between previous frame and current frame. In order to generate a better and accurate result for the subsequent operation, the extracted foreground mask of current changes must go through some improvements (pre-processing) and it will be complemented with Lukas-Kanade algorithm and Shi-Tomasi Corner Detector to tracks the human joints in the frames [22]. The current changes (motion mask) will then be further dilated, contoured and bounded using convexHull function [23].

- ii) Background changes (dynamic) can be calculated as:

$$\text{accumulated changes} - \text{current changes} = \text{background changes}$$

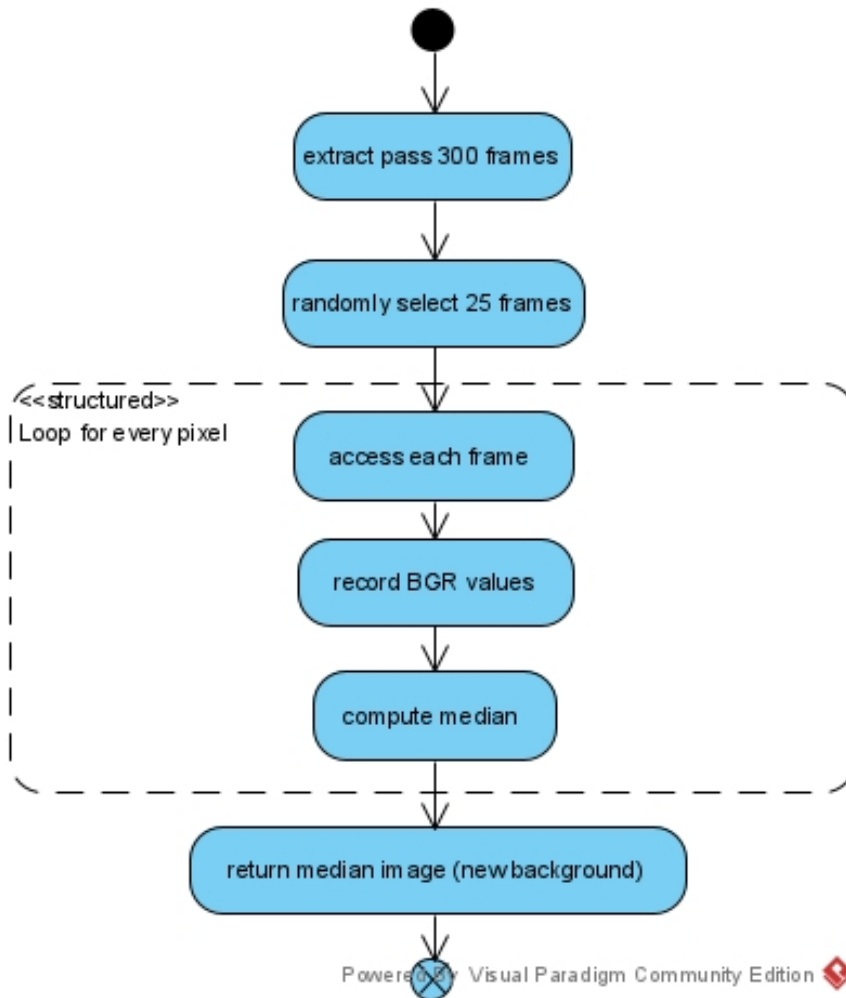
After the processing, the background changes can now be extracted from the processed accumulated changes and processed current changes.

The decision to trigger the warning for **significant** static changes is based on the threshold value comparison, once the static changes (non-zero pixel) exceed the threshold define in  $\text{minW}_{\text{pixel}}$  (based on the multiplication of the resolution of video frame and percentage), the risk level “danger” will be flagged and give off alarm to prevent further vandalism event.

#### **4.2.5 Background Updating**

Nevertheless, the background might be altered due to many aspects, such as viewing positions and angles, lighting condition (different time during the day or night and changes in season, weather), and the environment, those changes will lead to the occurrence of false alarm. Thus, action to ignored the scene changes over the time is a must. Moreover, a security alarm should be turn off automatically after some period of time, such that it would not keep ringing and disturb the others. In these cases, a background updating will be required to adapt the new environment.

However, the background image, so called background model must be a representation of scene with no object in motion as a basic standard [24]. Thus, the temporal median filter can be used to provide an adequate background model instead of extracting one of the frames from the video sequence that might captured moving object.



**Figure 4.6 Background Estimation**

In the proposed method, the background estimation will be implied after the vandalised event happened for a period of time. As Figure 4.6 shown, 25 samples in the past recorded 300 frames will be randomly selected and calculate the median of every pixel in BGR value over these 25 frames. Then, a newly generated background model (median image) will be updated to the system.



## Chapter 5 : System Implementation

### 5.1 Hardware Setup



The following table shows the hardware tools specification to develop the proposed system.

System	Information
Computer Model	ASUS K551L Laptop
Operating system	Windows 10 Home
Processor	Intel(R) Core(TM) i5-4210U
Graphic Card	NVIDIA® GeForce® 840M
CPU	1.70GHz
Memory (Ram)	8.00GB
System Type	64-bit operating system, x64-based processor

*Table 5.1 Hardware Tools for Development*

### 5.2 Software Setup

The following table shows the software tools and descriptions of each function to develop the proposed system.

Software	Description
	Microsoft Visual Studio is an integrated development environment (IDE) made by Microsoft. The platform is used to develop the computer vision program in this practice.
	OpenCV is an open-source library for Microsoft Visual Studio to provide programming functions targeted at image processing, machine learning, and computer vision.

*Table 5.2 Software Tools for Development*

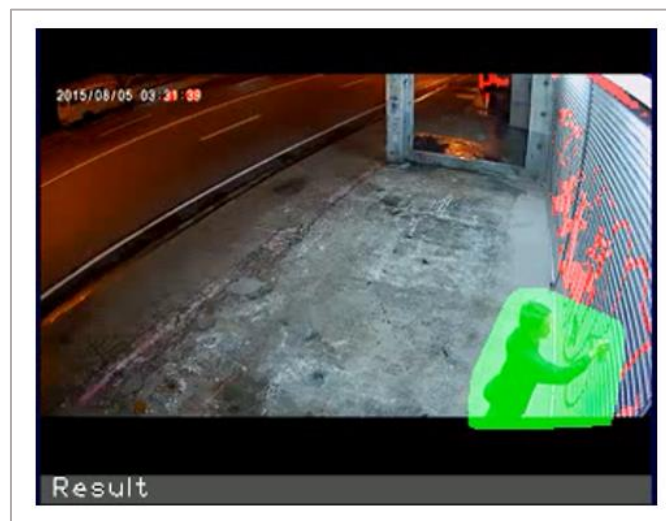
### 5.3 System Operation

#### 5.3.1 Vandalism Video Analysis



*Figure 5.1 Vandalism video analytic interface*

The surveillance analytic interface as shown in Figure 5.1 depicted the result generated from the methodology of vandalism detection. The result and analysis board will be further explained below.



*Figure 5.2 Result plane*

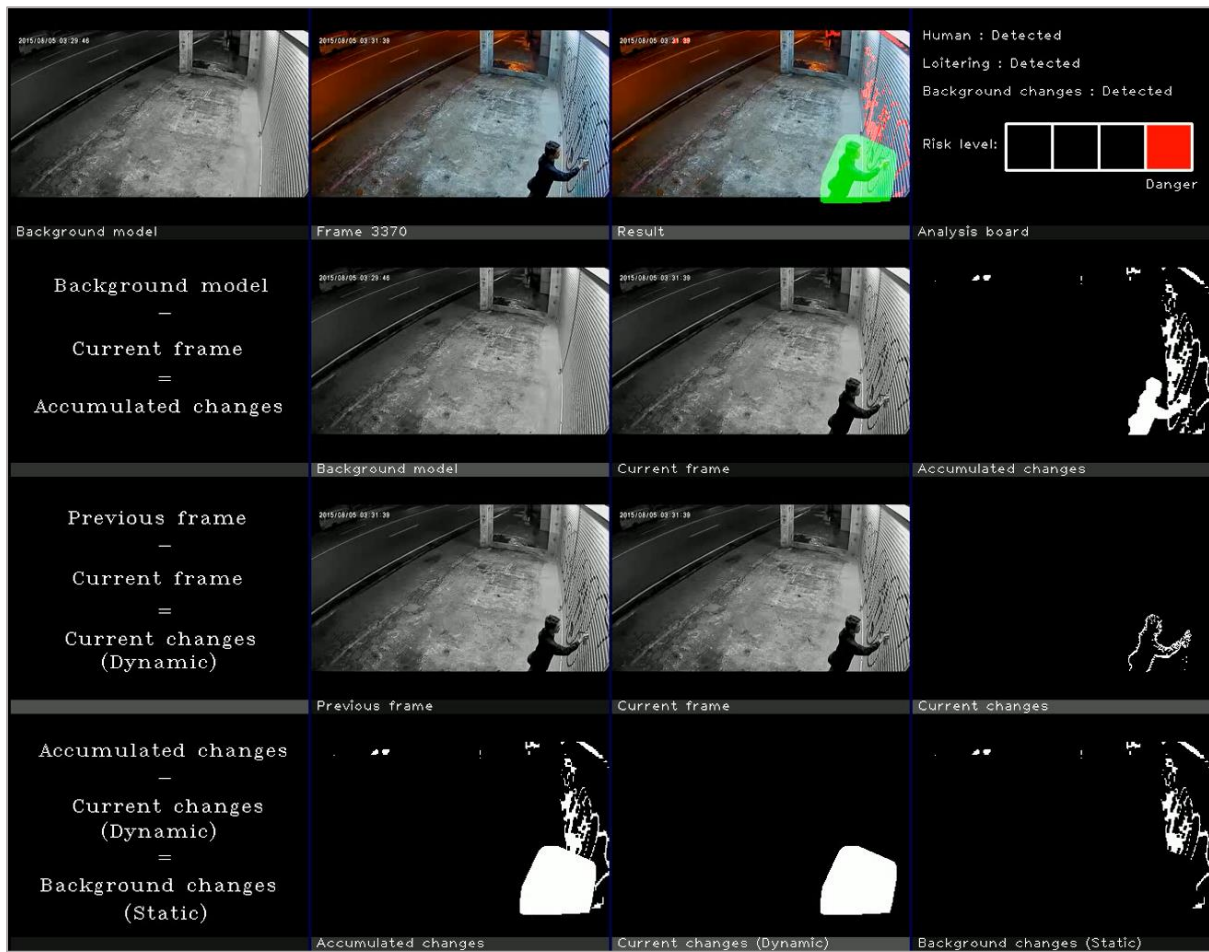
Noticed that the result plane is the outcome of significant background changes detection. The pixels painted in green in the result plane are the current changes (dynamic) that represent motion. While the pixels painted in red is the extracted background changes (static) that represent the permanent damages from vandalism, which can be easily classify by the human security officer/ user.



*Figure 5.3 Analysis board*

The analyzed risk level can be classified into four phases: Safe, Alert, Warning, and Danger. The risk level will stay at safe phase when there is no human/vandal detected and no significant destruction caused by vandalism. The risk level will rise to alert phase when the YOLO algorithm detected human in the scenes. If the person staying in the scene for too long or loitering at the particular area, the warning phase is entered as such behavior is considered suspicious and can be recognized as potential vandal. Ultimately, assuming that the vandal start the vandalism of public or private property, the system will enter the danger phase when the destruction of property made by the vandal is significant enough to be captured into the scenes and trigger the vandalism flag.

### 5.3.2 Detailed Window (Significant background changes detection)



*Figure 5.4 Detailed Window for Significant background changes detection*

The large window as shown in Figure 5.4 is dedicated for the Significant background changes detection approach which have been explained in Chapter 4.2.4, this window meticulously shows each step on how to extract the background changes (static) which represent the visible damage made from the vandalism act.

## Chapter 6 : System Evaluation and Discussion

### 6.1 System Testing and Result

#### 6.1.1 Scenario A

*Category: Graffiti*

a) normal event



b) two men entering the scene



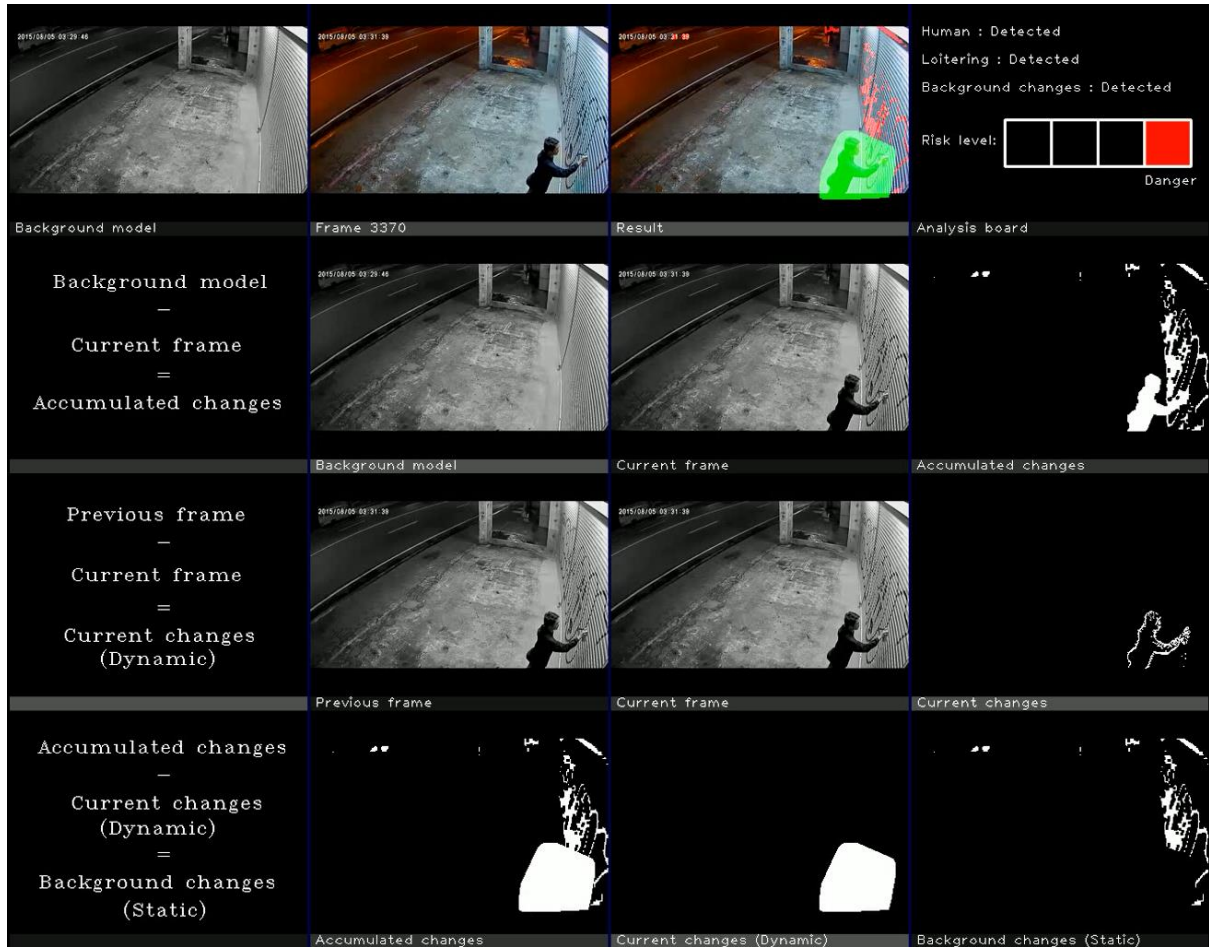
c) men approaching the metal gate and staying in the scene for a period



d) the metal gate is vandalised with illegal paint

*Result Plane:*

*[Green highlight: Current changes (dynamic); Red highlight: Background changes (Static)]*



e) background updating (background estimation)



**Figure 6.1 Scenario A Result**

From Figure 6.1 a), the system extracts the 1<sup>st</sup> frame as background model and there is no human and vandalism happen, thus the risk level remains “safe”. Then, the **YOLO algorithm** run on every 5 seconds and a person were detected, then the risk level rises to “alert” as shown in Figure 6.1 b). Since human was detected, the system starts the countdown on **loitering detection**. The risk level then raised to “warning” as shown in in Figure 6.1 c) when those men staying in the scene for too long. Finally, they started to spray paint on the metal gate. By utilizing the proposed method: **significant background changes detection**, the permanent damages of the property were successfully detected, eventually trigger the “danger” alarm. In addition, the red highlighted region in Figure 6.1 d)’s result plane represented the vandalised region, while green highlighted region represented the moving object i.e., the vandal.

Assuming the owner already noticed the vandalism event, the background required an updating after that. Thus, the system will estimate the background, the output of **background estimation** is generated and update to the background model as shown in Figure 6.1 e). Once the new background frame is updated, the system will stop the triggered alarm and reset the risk level to “Safe” as there is no frame difference between the new background (background model) and foreground (current frame) now.

### 6.1.2 Scenario B

Category: Defacing of property

a) normal event



b) man running into the scene



c) man staying in the scene for a period of time

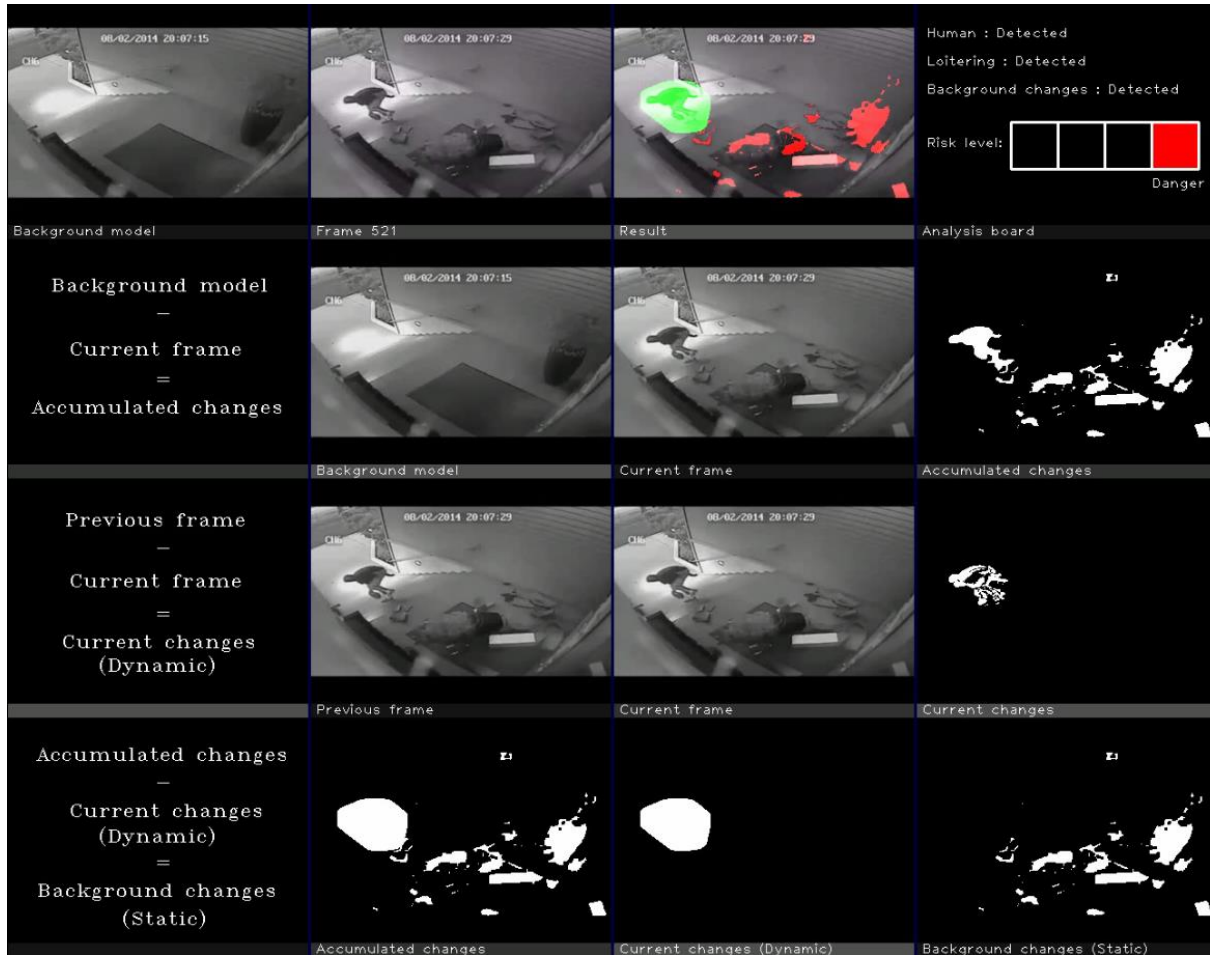




d) man kicked the vase and caused massive destruction

*Result Plane:*

*[Green highlight: Current changes (dynamic); Red highlight: Background changes (Static)]*



e) background updating (background estimation)



**Figure 6.2 Scenario B Result**

In Figure 6.2 a), the risk level detected as “safe” for normal event. Then, the risk level rises to “alert” as shown in Figure 6.2 b) when the man running into the scene, proven that the **human detection** is working. The **loitering** status is flagged as shown in Figure 6.2 c) after the man staying in the scene for a period of time. Finally, the man kicks the vase and causes it fell on the ground, the broken vase spread into fragments and the massive destruction are captured into the scene. Based on the result from Figure 6.2 d), the system is able to extract and detect the **significant background changes** i.e., the vase fragments and instantly trigger the “danger” alarm.

The system then **updated the background model** to stop the “danger” alarm after the end of vandalism event and reset the risk level back to “Safe”.

### 6.1.3 Scenario C

Category: Glass breakage

a) normal event



b) a young man entering the scene



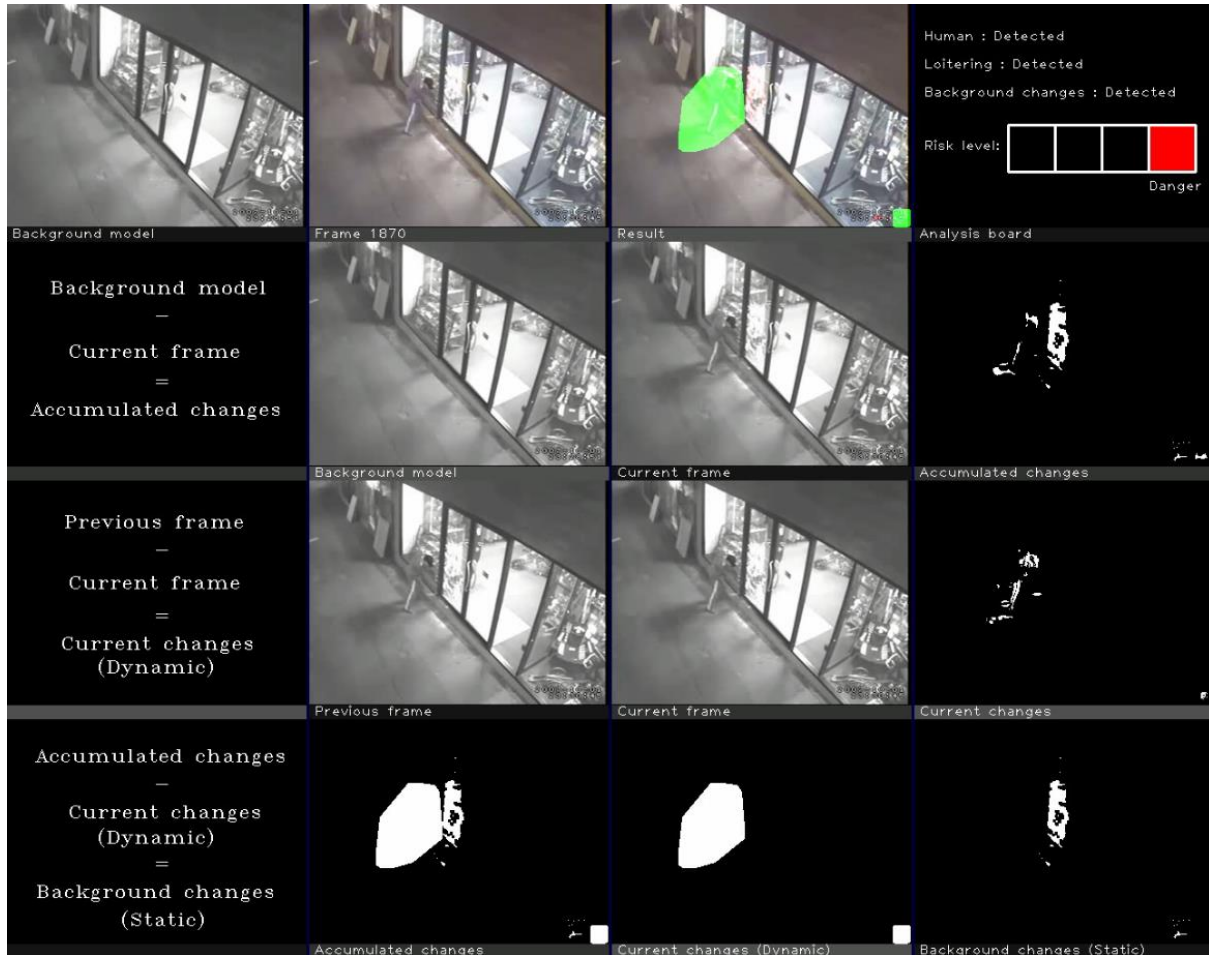
c) the young man loiters in front of the store



d) the young man breaks the window

*Result Plane:*

*[Green highlight: Current changes (dynamic); Red highlight: Background changes (Static)]*



**Figure 6.3 Scenario C Result**

In Figure 6.3 a), the risk level detected as “safe” for normal event. Then the risk level rises to “alert” for **human detection** when the young man walks into the scene as shown in Figure 6.3 b). After staying in the scene for a period of time, the **loitering** status is flagged and risk level rises to “warning” as shown in Figure 6.3 c). He was then started to smash the shop window. By utilizing the proposed method, the shattered glass captured in the scene are detected as **significant background changes (static)** and eventually triggered the “danger” alarm as shown in Figure 6.3 d).

### 6.1.4 Scenario D

Category: *Paint splasher*

a) normal event



b) a man entering the scene and approaching to a car



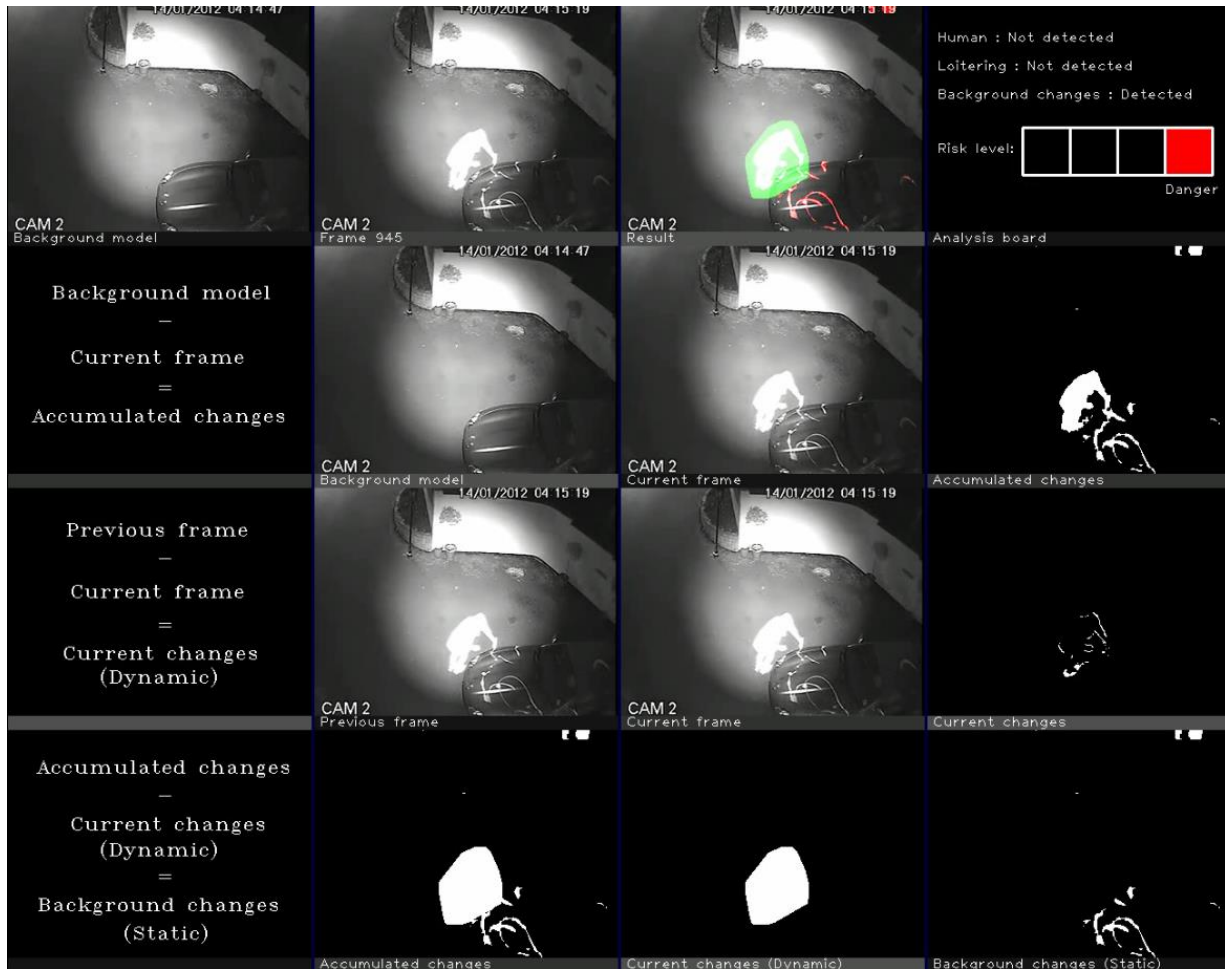
c) the man poured some paint on the car and go behind the car



d) the man continue to splash paint on the car

*Result Plane:*

*[Green highlight: Current changes (dynamic); Red highlight: Background changes (Static)]*



**Figure 6.4 Scenario D Result**

In Figure 6.4 a), the risk level detected as “safe” for normal event. Noticed that in Figure 6.4 b), **human detection** and **loitering detection** was failed due to video quality and poor illumination. Moreover, the man disappears from the scenes and could not be captured by the camera as shown in Figure 6.4 c), this is because of the bad placement of surveillance camera.

However, in the second level defence: the **significant background changes (static) detection** successfully detected the permanent damages made by the vandal and instantly triggered the “danger” warning as shown in Figure 6.4 d). In this way, even though the system fails in the first level defence due to the stated problems, the system is still able to stop the further vandalism using significant background changes (static) detection.

### 6.1.5 Scenario E

Category: Keyed car

a) normal event



b) a man entering the scene

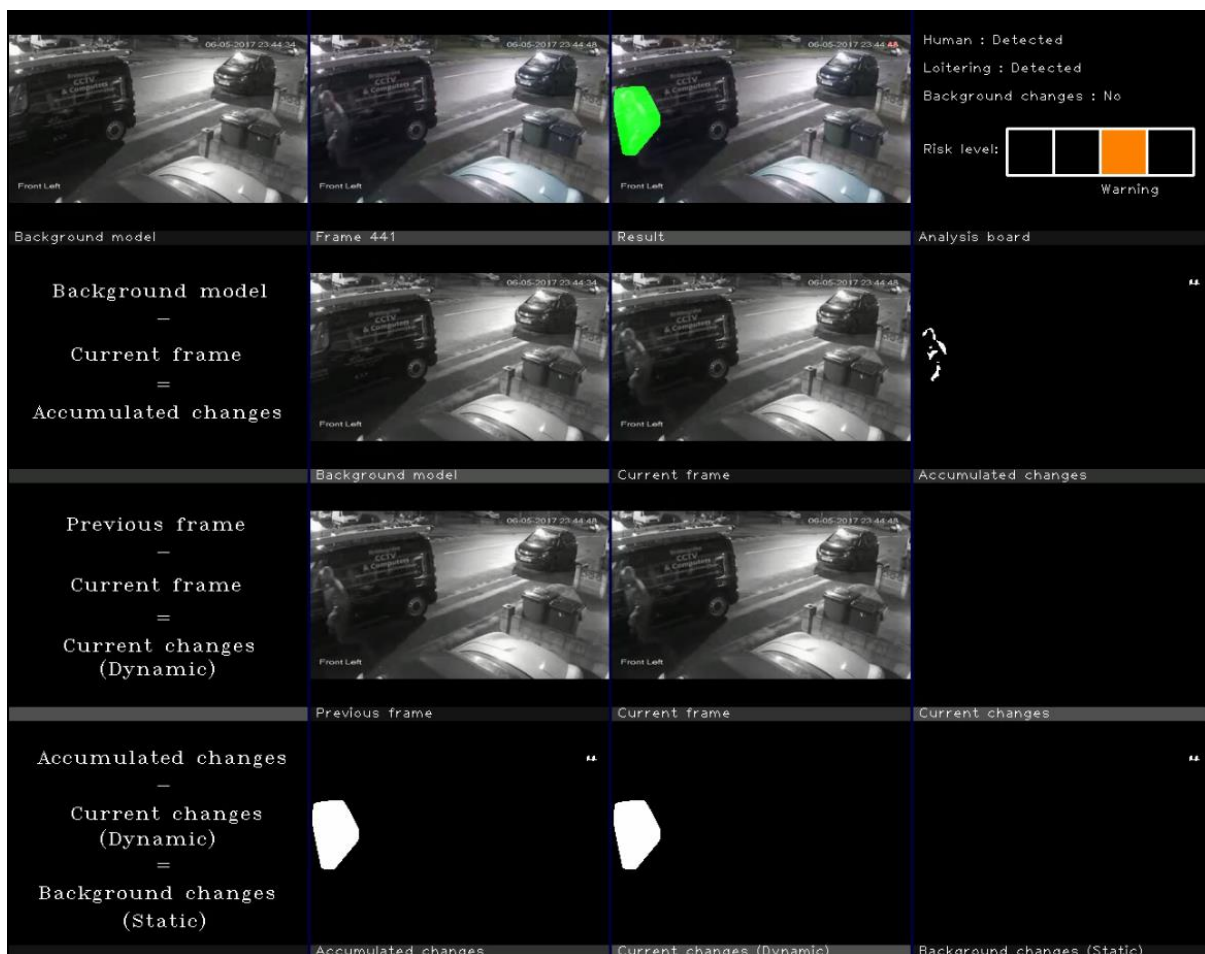


c) the man stays close to the van



d) the man keyed the van and quickly leave crime scene

[Green highlight: Current changes (dynamic); Red highlight: Background changes (Static)]



**Figure 6.5 Scenario E Result**

In Figure 6.5 a), the risk level detected as “safe” for normal event. Then the risk level rises to “alert” for **human detection** when the man walked into the scene as shown in Figure 6.5 b). After staying in the scene for a period of time, the **loitering** status is flagged and risk level rises to “warning” as shown in Figure 6.5 c). He was then scratched the van using a key/sharp object. However, the damages were so inconspicuous and not significant enough to be captured into the scene. In this case, the proposed **significant background changes (static) detection** is unable to detect the damages as shown in Figure 6.5 d).



## 6.2 Project Challenges

Most of the vandalism acts take place in isolation and specially during night time. Despite the issue of intensity, the real-world videos are also prone with other challenges such as blurry video, unclear vision, occlusion, noise based on bad weather condition, camera quality or camera placement. When dealing with varied variables from the unrestricted real world, image processing and detection jobs will subsequently be more challenging.

Besides, a 3D real life environment can only be described in 2D form in which the image processing can retrieve the information, and thus causing the lack of spatial and contextual information. Moreover, the vandalizing event can be tricky to track when 1) the vandal does not pose in manner worthy of attention and 2) there is no obvious damage on the property. For instances, scratching a car – is hard to catch and noticed through the CCTV even with human supervision.

### 6.3 Objectives Evaluation

✓ **A real-time automated vandalism detection surveillance system is developed**

The proposed solution operated at a frame rate of around 13 frames per second and responded to the vandalism event immediately by sending out early warning. Besides, the system is fully automated whereby users are not required to setup or predefine a region of interest (ROI) / risk zone in the scenes manually.

✓ **Vandalism detection system is developed by accomplishing the following sub-objectives:**

- **The potential vandals can be detected by identifying characteristics.**  
“Alert” and “Warning” signal are given off when human was detected when he/she loitered and aroused suspicion.
- **The significant background changes (damages) can be detected.**  
“Danger” alarm is given off when the unauthorized changes in the scene is significant enough to declare a vandalism event.

## **Chapter 7 : Conclusion and Recommendation**

### **7.1 Conclusion**

As the problem statement mentioned earlier, the physical vandalism is becoming an increasingly crime to be discussed and considered in our community as the problem give a negative impact on evolving social life and morality. Thus, the invention of monitoring system is found to be a key tool in reducing the crime rate. However, the problem come one after another because such system requires the supervision of human operators and their continuous attention to the captured scenes which is a time consuming and exhausted work. Hence, the intelligent video surveillance system can definitely solve these problems by analysing the vandalism event using computer vision techniques.

This project studied the UCF\_Crimes dataset and analyse the vandalism scenes to explore the solutions to detect the vandalism. The proposed vandalism detection system is made possible to prevent the vandalism by giving off early warning. For better understanding, the system is said to be a two-level defence system whereby it detects both the potential vandals and the vandalised activities, which prevent the vandalism before it happens and the latter try to stop the further vandalising action in order to minimize the destruction of the property. In the case of predicting suspected vandals, the techniques included YOLO detection, suspicious characteristic detection such as loitering to recognise human and their trajectory in the scene, while in the case of defining the vandalism event, the damages detection were applied to identify the significant static changes, representing damage in the scene. While, the experimental result had proved the successful of proposed vandalism detection system on detecting vandalism.

## **7.2 Recommendation**

More innovative ideas can be implemented with the vandalism video analysis and solutions to develop a stronger and reliable system. In future works, the human activity recognition (HAR) [25] of vandal is recommended to improve the proposed system such that the action of breaking, smashing and defacing of property can be detected. On the other hand, the proposed system can also be improved by detecting suspicious object shape and appearance. For example, using the modified Hough transform (MHT) [6] to detect the circular arcs in the image, i.e., helmet shape since some intensive vandals would wear helmet to hide their face from the security camera.

## Bibliography






















- [1] M. Barker and C. Bridgeman, Preventing vandalism: What works?, Police research group, 1994, pp. 1-3.
- [2] Y. Cannix, "HK\$65 million bill for repairs on public facilities wrecked by radicals," *South China Morning Post*, 8 Jan 2020. [Online]. Available: <https://www.scmp.com/news/hong-kong/transport/article/3045180/hk65-million-bill-repairs-public-facilities-vandalised>. [Accessed 1 Mar 2021].
- [3] W. Sultani, C. Chen and M. Shah, "Real-world anomaly detection in surveillance videos," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6479-6488, 2018.
- [4] R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," *arXiv preprint arXiv:1901.03407*, 2019.
- [5] T. Sikandar, K. H. Ghazali and M. F. Rabbi, "ATM crime detection using Image Processing Integrated Video Surveillance: A systematic review," *Multimedia Systems*, vol. 25, no. 3, p. 229–251, 2018.
- [6] C. Y. Wen, S. H. Chiu, J. J. Liaw and C. P. Lu, "The safety helmet detection for ATM's surveillance system via the modified Hough transform," *IEEE 37th Annual 2003 International Carnahan Conference on Security Technology*, pp. 364-369, 2003.
- [7] R. Nayak, U. C. Pati and S. K. Das, "A comprehensive review on Deep Learning-based methods for video anomaly detection," *Image and Vision Computing*, vol. 106, pp. 1-16, 2020.
- [8] P. M. Joshi, "Generative vs discriminative models," *Medium*, 1 Sep 2018. [Online]. Available: <https://medium.com/@mlengineer/generative-and-discriminative-models-af5637a66a3>. [Accessed 1 Jun 2021].

- [9] R. Arroyo, J. J. Yebes, L. M. Bergasa, I. G. Daza and J. Almazán, "Expert video-surveillance system for real-time detection of suspicious behaviors in shopping malls," *Expert Systems with Applications*, vol. 42, no. 21, pp. 7991-8005, 2015.
- [10] T. E. Easterfield, "A combinatorial algorithm," *Journal of the London Mathematical Society*, Vols. s1-21, no. 3, pp. 219-226, 1946.
- [11] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME—Journal of Basic Engineering*, vol. 82, pp. 35-45, 1960.
- [12] V. Vapnik, S. E. Golowich and A. Smola, "On discriminative vs generative classifiers: a comparison of logistics regression and naïve bayes," *Advances in Neural Information Processing Systems*, vol. 9, pp. 841-848, 1996.
- [13] T. Zhang, J. Li, W. Jia, J. Sun and H. Yang, "Fast and robust occluded face detection in ATM surveillance," *Pattern Recognition Letters*, vol. 107, p. 33–40, 2018.
- [14] C. Liu, "A bayesian discriminating features method for face detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 725-740, 2003.
- [15] M. Ribeiro, E. Lazzaretti and H. S. Lopes, "A study of deep convolutional auto-encoders for anomaly detection in videos," *Pattern Recognition Letters*, vol. 105, pp. 13-22, 2018.
- [16] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 6, pp. 679-698, 1986.
- [17] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, "Learning spatiotemporal features with 3D convolutional networks," " *IEEE International Conference on Computer Vision*, pp. 4489-4497, 2015.
- [18] R. Hou, C. Chen and M. Shah, "Tube Convolutional Neural Network (T-CNN) for action detection in videos," *IEEE International Conference on Computer Vision*, pp. 5822-5831, 2017.

- [19] A. Rosebrock, "YOLO object detection with OpenCV," PyImageSearch, 12 Nov 2018. [Online]. Available: <https://pyimagesearch.com/2018/11/12/yolo-object-detection-with-opencv/>. [Accessed 4 Apr 2021].
- [20] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.
- [21] S. Sural, S. Pramanik and Gang Qian, "Segmentation and histogram generation using the HSV color space for image retrieval," *Proceedings. International Conference on Image Processing*, vol. 2, 2002.
- [22] M. Kuklin, "Optical flow in opencv (C++/Python)," LearnOpenCV, 16 Jul 2021. [Online]. Available: <https://learnopencv.com/optical-flow-in-opencv/>. [Accessed 17 Aug 2021].
- [23] J. Sklansky, "Finding the convex hull of a simple polygon," *Pattern Recognition Letters*, vol. 1, no. 2, p. 79–83, 1982.
- [24] M. Piccardi, "Background subtraction techniques: A Review," *IEEE International Conference on Systems, Man and Cybernetics*, vol. 4, pp. 3099-3104, 2004.
- [25] R. K. Tripathi, A. S. Jalal and S. C. Agrawal, "Suspicious human activity recognition: a review," *Artif Intell Rev*, no. 50, p. 283–339, 2018.
- [26] J. Masci, U. Meier, D. Ciresan and J. Schmidhuber, "Stacked convolutional auto-encoders for hierarchical feature extraction," *International conference on artificial neural networks*, pp. 52-59, 2011.

# Appendices

## A.1 Experimental Result

 <p>Background model</p>	 <p>Frame 1</p>	 <p>Result</p>	<p>Human : Not detected Loitering : Not detected Background changes : No</p> <p>Risk level: <span style="display: inline-block; width: 20px; height: 15px; background-color: green; border: 1px solid black;"></span> <span style="display: inline-block; width: 20px; height: 15px; background-color: black; border: 1px solid black;"></span> <span style="display: inline-block; width: 20px; height: 15px; background-color: black; border: 1px solid black;"></span> <span style="display: inline-block; width: 20px; height: 15px; background-color: black; border: 1px solid black;"></span></p> <p>Safe</p>
 <p>Background model</p>	 <p>Frame 177</p>	 <p>Result</p>	<p>Human : Detected Loitering : Not detected Background changes : No</p> <p>Risk level: <span style="display: inline-block; width: 20px; height: 15px; background-color: black; border: 1px solid black;"></span> <span style="display: inline-block; width: 20px; height: 15px; background-color: yellow; border: 1px solid black;"></span> <span style="display: inline-block; width: 20px; height: 15px; background-color: black; border: 1px solid black;"></span> <span style="display: inline-block; width: 20px; height: 15px; background-color: black; border: 1px solid black;"></span></p> <p>Alert</p>
 <p>Background model</p>	 <p>Frame 294</p>	 <p>Result</p>	<p>Human : Detected Loitering : Detected Background changes : No</p> <p>Risk level: <span style="display: inline-block; width: 20px; height: 15px; background-color: black; border: 1px solid black;"></span> <span style="display: inline-block; width: 20px; height: 15px; background-color: black; border: 1px solid black;"></span> <span style="display: inline-block; width: 20px; height: 15px; background-color: orange; border: 1px solid black;"></span> <span style="display: inline-block; width: 20px; height: 15px; background-color: black; border: 1px solid black;"></span></p> <p>Warning</p>
 <p>Background model</p>	 <p>Frame 513</p>	 <p>Result</p>	<p>Human : Detected Loitering : Detected Background changes : Detected</p> <p>Risk level: <span style="display: inline-block; width: 20px; height: 15px; background-color: black; border: 1px solid black;"></span> <span style="display: inline-block; width: 20px; height: 15px; background-color: black; border: 1px solid black;"></span> <span style="display: inline-block; width: 20px; height: 15px; background-color: black; border: 1px solid black;"></span> <span style="display: inline-block; width: 20px; height: 15px; background-color: red; border: 1px solid black;"></span></p> <p>Danger</p>
<p>Background model - Current frame = Accumulated changes</p>	 <p>Background model</p>	 <p>Current frame</p>	 <p>Accumulated changes</p>
<p>Previous frame - Current frame = Current changes (Dynamic)</p>	 <p>Previous frame</p>	 <p>Current frame</p>	 <p>Current changes</p>
<p>Accumulated changes - Current changes (Dynamic) = Background changes (Static)</p>	 <p>Accumulated changes</p>	 <p>Current changes (Dynamic)</p>	 <p>Background changes (Static)</p>



			Human : Not detected Loitering : Not detected Background changes : No Risk level:  Safe
Background model	Frame 11	Result	Analysis board

			Human : Detected Loitering : Not detected Background changes : No Risk level:  Alert
Background model	Frame 124	Result	Analysis board

			Human : Detected Loitering : Detected Background changes : No Risk level:  Warning
Background model	Frame 187	Result	Analysis board

			Human : Detected Loitering : Not detected Background changes : Detected Risk level:  Danger
Background model	Frame 546	Result	Analysis board

$\begin{aligned} &\text{Background model} \\ &- \\ &\text{Current frame} \\ &= \\ &\text{Accumulated changes} \end{aligned}$			
	Background model	Current frame	Accumulated changes
$\begin{aligned} &\text{Previous frame} \\ &- \\ &\text{Current frame} \\ &= \\ &\text{Current changes} \\ &\text{(Dynamic)} \end{aligned}$			
	Previous frame	Current frame	Current changes
$\begin{aligned} &\text{Accumulated changes} \\ &- \\ &\text{Current changes} \\ &\text{(Dynamic)} \\ &= \\ &\text{Background changes} \\ &\text{(Static)} \end{aligned}$			
	Accumulated changes	Current changes (Dynamic)	Background changes (Static)

## A.2 Final Year Project 2 Weekly Report

# FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Y3T3	Study week no.:1
Student Name & ID: Britney Muk Yuen Kuan, 1803450	
Supervisor: Prof. Dr Leung Kar Hang	
Project Title: Vandalism Video Analysis Employing Computer Vision Technique	

### 1. WORK DONE

Report of work done in FYP1.  
Structure the project plan.

### 2. WORK TO BE DONE

Prepare to submit FYP draft report around week 6/10  
Complete chapter 1 by week 2.  
Send video demo for evaluation.

### 3. PROBLEMS ENCOUNTERED

-

### 4. SELF EVALUATION OF THE PROGRESS

-



Supervisor's signature

25 Jan 2022



Student's signature

# FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

<b>Trimester, Year: Y3T3</b>	<b>Study week no.: 3</b>
<b>Student Name &amp; ID: Britney Muk Yuen Kuan, 1803450</b>	
<b>Supervisor: Prof. Dr Leung Kar Hang</b>	
<b>Project Title: Vandalism Video Analysis Employing Computer Vision Technique</b>	

## 1. WORK DONE

Tidy up FYP report (Chapter 1 & 2, IEEE citation, etc.)

Classified vandalism into different group.

Structure the project plan.

## 2. WORK TO BE DONE

Correcting some wording within the report as discussed.

Changing the terms in the Static Changes Detection and

Improve the result interface (demo) for better understanding.

## 3. PROBLEMS ENCOUNTERED

-

## 4. SELF EVALUATION OF THE PROGRESS

-



Supervisor's signature

8 Feb 2022



Student's signature

# FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

<b>Trimester, Year: Y3T3</b>	<b>Study week no.: 4</b>
<b>Student Name &amp; ID: Britney Muk Yuen Kuan, 1803450</b>	
<b>Supervisor: Prof. Dr Leung Kar Hang</b>	
<b>Project Title: Vandalism Video Analysis Employing Computer Vision Technique</b>	

## 1. WORK DONE

Corrected some wording within the report as discussed.

Changed the terms in the Static/Background Changes Detection for better understanding.

## 2. WORK TO BE DONE

Improve the system and the result interface (demo).

## 3. PROBLEMS ENCOUNTERED

-

## 4. SELF EVALUATION OF THE PROGRESS

-



Supervisor's signature

15 Feb 2022



Student's signature

# FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

<b>Trimester, Year: Y3T3</b>	<b>Study week no.: 5</b>
<b>Student Name &amp; ID: Britney Muk Yuen Kuan, 1803450</b>	
<b>Supervisor: Prof. Dr Leung Kar Hang</b>	
<b>Project Title: Vandalism Video Analysis Employing Computer Vision Technique</b>	

## 1. WORK DONE

Improved the result interface for clearer understanding.

## 2. WORK TO BE DONE

Improve the system and the result interface (demo).

## 3. PROBLEMS ENCOUNTERED

-

## 4. SELF EVALUATION OF THE PROGRESS

-



Supervisor's signature

22 Feb 2022



Student's signature

# FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

<b>Trimester, Year: Y3T3</b>	<b>Study week no.: 6</b>
<b>Student Name &amp; ID: Britney Muk Yuen Kuan, 1803450</b>	
<b>Supervisor: Prof. Dr Leung Kar Hang</b>	
<b>Project Title: Vandalism Video Analysis Employing Computer Vision Technique</b>	

## 1. WORK DONE

Improved the result interface for clearer understanding.

## 2. WORK TO BE DONE

Improve the system.

Go beyond the current group and start to tackle another type of vandalism scene.

## 3. PROBLEMS ENCOUNTERED

-

## 4. SELF EVALUATION OF THE PROGRESS

-



Supervisor's signature

1 Mar 2022



Student's signature

# FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

<b>Trimester, Year: Y3T3</b>	<b>Study week no.: 7</b>
<b>Student Name &amp; ID: Britney Muk Yuen Kuan, 1803450</b>	
<b>Supervisor: Prof. Dr Leung Kar Hang</b>	
<b>Project Title: Vandalism Video Analysis Employing Computer Vision Technique</b>	

## 1. WORK DONE

Identified the problem of false alarm and ideal solution.

## 2. WORK TO BE DONE

Handle the false alarm due to light intensity.

## 3. PROBLEMS ENCOUNTERED

Intensity changing due to weather, car light. False detected as background changes (static)

## 4. SELF EVALUATION OF THE PROGRESS

-



Supervisor's signature

8 Mar 2022



Student's signature

# FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

<b>Trimester, Year: Y3T3</b>	<b>Study week no.: 8</b>
<b>Student Name &amp; ID: Britney Muk Yuen Kuan, 1803450</b>	
<b>Supervisor: Prof. Dr Leung Kar Hang</b>	
<b>Project Title: Vandalism Video Analysis Employing Computer Vision Technique</b>	

## 1. WORK DONE

Identified the problem of false alarm and ideal solution.

Handle the false alarm due to light intensity (e.g. car light) using colorspace transformation and thresholding the V("value") channel.

## 2. WORK TO BE DONE

Handle the false alarm due to sudden light change.

## 3. PROBLEMS ENCOUNTERED

-

## 4. SELF EVALUATION OF THE PROGRESS

-



Supervisor's signature

15 Mar 2022



Student's signature



# FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

<b>Trimester, Year: Y3T3</b>	<b>Study week no.: 9</b>
<b>Student Name &amp; ID: Britney Muk Yuen Kuan, 1803450</b>	
<b>Supervisor: Prof. Dr Leung Kar Hang</b>	
<b>Project Title: Vandalism Video Analysis Employing Computer Vision Technique</b>	

## 1. WORK DONE

Improve/stabilize the system using histogram equalization to eliminate the problem of sudden light/illumination changes.

## 2. WORK TO BE DONE

Continue to improve the system and tackle the other type of vandalism scene.

## 3. PROBLEMS ENCOUNTERED

-

## 4. SELF EVALUATION OF THE PROGRESS

-



Supervisor's signature

22 Mar 2022



Student's signature

# FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

<b>Trimester, Year: Y3T3</b>	<b>Study week no.: 10</b>
<b>Student Name &amp; ID: Britney Muk Yuen Kuan, 1803450</b>	
<b>Supervisor: Prof. Dr Leung Kar Hang</b>	
<b>Project Title: Vandalism Video Analysis Employing Computer Vision Technique</b>	

## 1. WORK DONE

Generate result demo for vandalism analysis.

## 2. WORK TO BE DONE

Start to write and complete the FYP report.

## 3. PROBLEMS ENCOUNTERED

-

## 4. SELF EVALUATION OF THE PROGRESS

-



Supervisor's signature

29 Mar 2022



Student's signature

# FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

<b>Trimester, Year: Y3T3</b>	<b>Study week no.: 11</b>
<b>Student Name &amp; ID: Britney Muk Yuen Kuan, 1803450</b>	
<b>Supervisor: Prof. Dr Leung Kar Hang</b>	
<b>Project Title: Vandalism Video Analysis Employing Computer Vision Technique</b>	

## 1. WORK DONE

Start to write and complete the FYP report.

## 2. WORK TO BE DONE

Complete remaining chapters.

## 3. PROBLEMS ENCOUNTERED

-

## 4. SELF EVALUATION OF THE PROGRESS

-




Supervisor's signature

6 Apr 2022



Student's signature

### A.3 Poster



**UTAR**  
UNIVERSITI TUNKU ABDUL RAHMAN

University Tunku Abdul Rahman  
Faculty of Information and Communication  
Technology  
Bachelor of Computer Science (Honours)

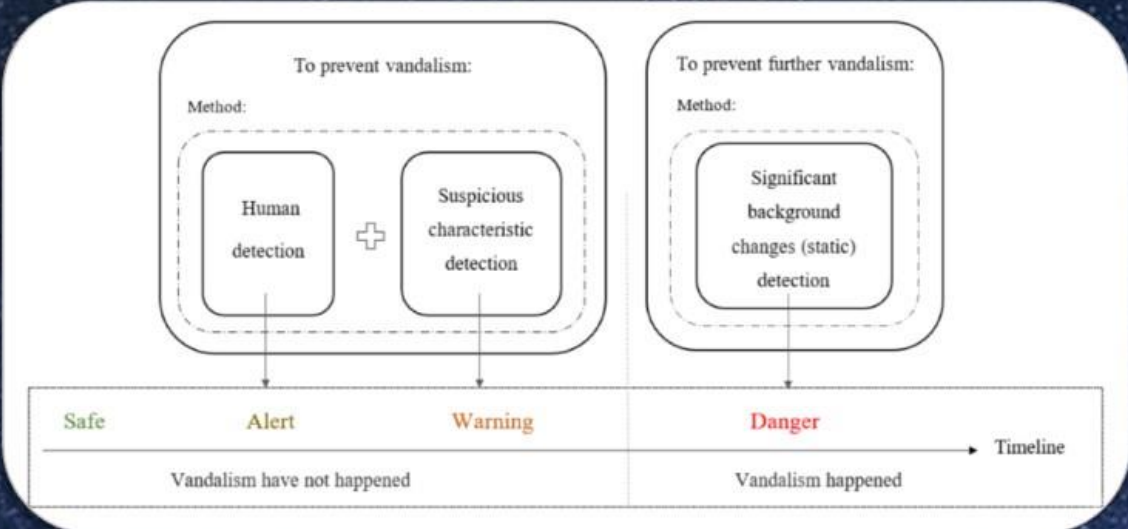
Student Name: Britney Muk Yuen Kuan  
ID: 18ACB03450  
Supervisor: Prof. Maylor Leung Kar Hang

---

## Vandalism Video Analysis Employing Computer Vision Technique

Computer can be trained to perform humanlike tasks !  
To prevent the vandalism, the proposed intelligent surveillance system can detect vandalism events and protect your property by giving out early warning!


**Methodology:**



The methodology is divided into two main phases:




- To prevent vandalism:** This phase uses a method combining **Human detection** and **Suspicious characteristic detection**. It results in a progression from **Safe** to **Alert** to **Warning** as vandalism has not yet occurred.
- To prevent further vandalism:** This phase uses **Significant background changes (static) detection**. It results in a progression from **Warning** to **Danger** once vandalism has happened.

**System Implementation:**

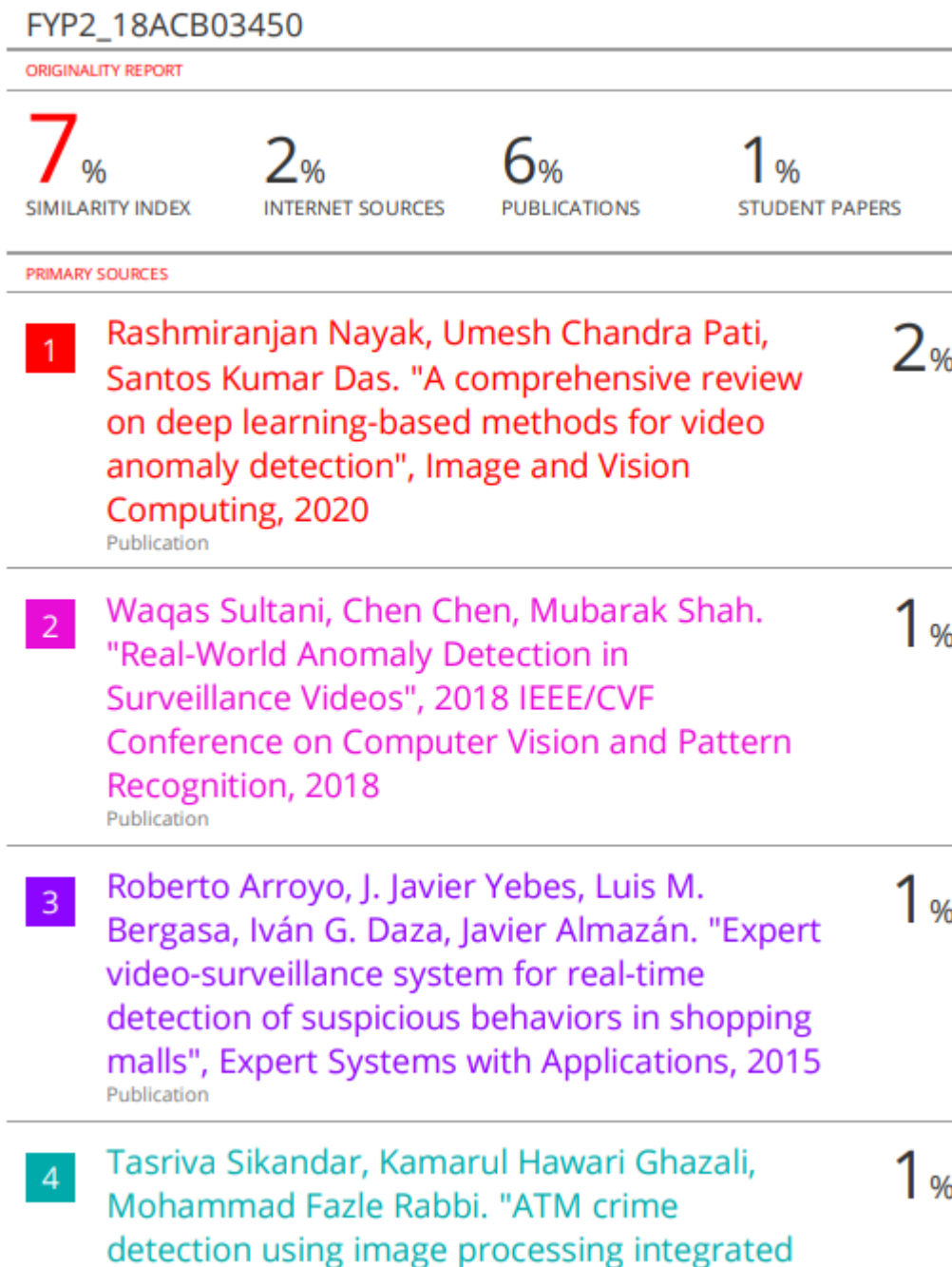


The implementation shows two examples of the system's output:

- Frame 1882:** Shows a person in a hallway. The analysis board indicates: Human : Detected, Loitering : Detected, Background changes : No. The risk level is **Warning** (represented by an orange bar).
- Frame 3370:** Shows a person writing on a wall. The analysis board indicates: Human : Detected, Loitering : Detected, Background changes : Detected. The risk level is **Danger** (represented by a red bar).

Automatic  **YES!**    Run in real-time  **YES!**    High detection rate  **YES!**

## A.4 Plagiarism check result



---

video surveillance: a systematic review",  
Multimedia Systems, 2018

Publication

- 
- |    |  |     |
|----|--|-----|
| 5  | Manassés Ribeiro, André Eugênio Lazzaretti, Heitor Silvério Lopes. "A study of deep convolutional auto-encoders for anomaly detection in videos", Pattern Recognition Letters, 2018<br>Publication | 1%  |
| 6  | Tao Zhang, Jingjing Li, Wenjing Jia, Jun Sun, Huihua Yang. "Fast and robust occluded face detection in ATM surveillance", Pattern Recognition Letters, 2017<br>Publication                         | <1% |
| 7  | worldwidescience.org<br>Internet Source  | <1% |
| 8  | Submitted to Liverpool John Moores University<br>Student Paper   | <1% |
| 9  | Mohammed Ghazal, Carlos Vázquez, Aishy Amer. "Real-time vandalism detection by monitoring object activities", Multimedia Tools and Applications, 2011<br>Publication                               | <1% |
| 10 | Submitted to Volunteer State Community College<br>Student Paper  | <1% |
-

11	trap.ncirl.ie Internet Source	<1 %
12	zenodo.org Internet Source	<1 %
13	inmeta.medium.com Internet Source	<1 %
14	www.irjet.net Internet Source	<1 %
15	norma.ncirl.ie Internet Source	<1 %
16	Submitted to Staffordshire University Student Paper	<1 %
17	gene.kias.re.kr Internet Source	<1 %

Exclude quotes Off  
Exclude bibliography On

Exclude matches Off

<b>Universiti Tunku Abdul Rahman</b>			
<b>Form Title : Supervisor's Comments on Originality Report Generated by Turnitin for Submission of Final Year Project Report (for Undergraduate Programmes)</b>			
Form Number: FM-IAD-005	Rev No.: 0	Effective Date: 01/10/2013	Page No.: 1 of 1



**FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY**

<b>Full Name(s) of Candidate(s)</b>	Britney Muk Yuen Kuan
<b>ID Number(s)</b>	18ACB03450
<b>Programme / Course</b>	Bachelor of Computer Science (Honours)
<b>Title of Final Year Project</b>	Vandalism Video Analysis Employing Computer Vision Technique

<b>Similarity</b>	<b>Supervisor's Comments (Compulsory if parameters of originality exceeds the limits approved by UTAR)</b>
<b>Overall similarity index: 7</b>  <b>% Similarity by source</b> Internet Sources: 2 % Publications: 6 % Student Papers: 1 %	
<b>Number of individual sources listed of more than 3% similarity: None</b>	
<b>Parameters of originality required and limits approved by UTAR are as Follows:</b> (i) Overall similarity index is 20% and below, and (ii) Matching of individual sources listed must be less than 3% each, and (iii) Matching texts in continuous block must not exceed 8 words <i>Note: Parameters (i) – (ii) shall exclude quotes, bibliography and text matches which are less than 8 words.</i>	

Note Supervisor/Candidate(s) is/are required to provide softcopy of full set of the originality report to Faculty/Institute

***Based on the above results, I hereby declare that I am satisfied with the originality of the Final Year Project Report submitted by my student(s) as named above.***

Signature of Supervisor

Name: Leung Kar Hang

Date: 18 Apr 2022

Signature of Co-Supervisor

Name: \_\_\_\_\_

Date: \_\_\_\_\_



## A.5 FYP 2 Checklist



**UNIVERSITI TUNKU ABDUL RAHMAN**  
**FACULTY OF INFORMATION & COMMUNICATION TECHNOLOGY**  
**(KAMPAR CAMPUS)**

**CHECKLIST FOR FYP2 THESIS SUBMISSION**

Student Id	18ACB03450
Student Name	Britney Muk Yuen Kuan
Supervisor Name	Prof. Leung Kar Hang

TICK (✓)	DOCUMENT ITEMS
	Your report must include all the items below. Put a tick on the left column after you have checked your report with respect to the corresponding item.
	Front Plastic Cover (for hardcopy)
✓	Title Page
✓	Signed Report Status Declaration Form
✓	Signed FYP Thesis Submission Form
✓	Signed form of the Declaration of Originality
✓	Acknowledgement
✓	Abstract
✓	Table of Contents
✓	List of Figures (if applicable)
✓	List of Tables (if applicable)
	List of Symbols (if applicable)
✓	List of Abbreviations (if applicable)
✓	Chapters / Content
✓	Bibliography (or References)
✓	All references in bibliography are cited in the thesis, especially in the chapter of literature review
✓	Appendices (if applicable)
✓	Weekly Log
✓	Poster
✓	Signed Turnitin Report (Plagiarism Check Result - Form Number: FM-IAD-005)
✓	I agree 5 marks will be deducted due to incorrect format, declare wrongly the ticked of these items, and/or any dispute happening for these items in this report.

\*Include this form (checklist) in the thesis (Bind together as the last page)

I, the author, have checked and confirmed all the items listed in the table are included in my report.

(Signature of Student)

Date: 17/4/2022

