

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

AI/ML FOR ANOMALOUS SURVEILLANCE:

A DETECTION AND ALERT SYSTEM

An Undergraduate Thesis

Presented to the Faculty of the
College of Information and Communications Technology

West Visayas State University

La Paz, Iloilo City

In Partial Fulfillment

of the Requirements for the Degree
Bachelor of Science in Computer Science

by

Jesmar Troy A. Adolacion

Nezel C. Leagogo

Francine Anne N. Millares

Francis Rey M. Soldevilla

June 2022

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Iloilo City, Philippines

Approval Sheet

AI/ML FOR ANOMALOUS SURVEILLANCE:

A DETECTION AND ALERT SYSTEM

An Undergraduate Thesis for the Degree
Bachelor of Science in Computer Science

by

Jesmar Troy A. Adolacion

Nezel C. Leagogo

Francine Anne N. Millares

Francis Rey M. Soldevilla

Approved:

DR. EVANS B. SANSOLIS
Adviser

DR. MA. LUCHE P. SABAYLE
Chair, Computer Science Department

DR. MA. BETH S. CONCEPCION
Dean, CICT

June 2022

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Ilog, Iloilo City, Philippines

Acknowledgment

The researchers would like to express their deepest appreciation to the following persons, who in one way or another have made this work possible:

First of all, we would like to thank our parents for their invaluable financial and moral support - and for believing that we can make this study happen.

Secondly, we would like to extend our heartfelt appreciation to our Thesis Adviser, Dr. Evans Sansolis, for his guidance and for sharing his expertise in this field of study. We really appreciate his untiring motivation to make this study into reality.

Next, we would also like to thank Alyssa Danielle Magallanes and family for assisting us in the needed resources and providing us a place to conduct our study. We really appreciate the generosity in helping us in finishing this paper.

We would also like to express our gratitude to our alumni who lent us a hand with their previous studies. We are grateful for the gesture and guidance with our queries. We are also grateful to our classmates and friends who helped us be clarified in what we do not understand with

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Ilog, Iloilo City, Philippines

regards to our Thesis paper. Their belief in our capacity scaffolded our own to finish this study.

We are also grateful for the respondents who took the time to answer our surveys and shared their insights about our study to further enhance the system in the near future.

In addition, we also would like to extend our appreciation to the panelists who mold us to think outside the box and motivates us to do better.

We would like to express our gratitude to the editors, Dr. Frank Elijorde, Professor Marievic Violeta, Dr. Ma. Luche Sabayle, and Dr. Regin Cabacas for allowing us to grow and pushing us to improve and become a better version of ourselves through tasks. Additionally, we would like to give our warmest heartfelt gratitude to Sir John Cristopher Mateo and Sir Mark Joseph Solidarios for guiding us and coaching us to be able to make this happen.

And lastly, we would like to thank our Almighty Father for the strength and wisdom that He bestowed upon us. We offer everything that we have done to Him. Without Him, we

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

may not have been able to come up and finish this Thesis paper.

Francis Rey M. Soldevilla

Francine Anne N. Millares

Jesmar Troy A. Adolacion

Nezel C. Leagogo

June 2022

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Iloilo City, Philippines

Soldevilla, Francis Rey M.; Millare, Francine Anne N.; Adolacion, Jesmar Troy A.; Leagogo, Nezel C.; "AI/ML for Anomalous Surveillance: A Detection and Alert System". Unpublished Undergraduate Thesis, Bachelor of Science in Computer Science, West Visayas State University, Iloilo City, Philippines, June 2022.

Abstract

Due to the pandemic and work shifts family members feel more concerned with older folks being on their own at home. There is a constant worry for the elderly loved ones left alone since they are more fragile and prone to accidents. When their family members are deployed to work away from them, they can be left unsupervised or placed under caregivers who are hired to attend to their needs. In this study, the researchers proposed an automated system that detects anomalies among elderly people with the use of algorithms, namely, LSTM, CNN, and YOLOv3. The anomalies were identified as drowsiness, falling, weapons, and violence. An alarm system was also installed to notify users of detection that occurred. A total of 40 respondents were asked to evaluate the system where an overall mean of

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

4.394 was garnered, indicating that the system was "Very Good." Therefore, the system was successfully able to detect anomalous behavior presented from live feedback coming from the camera installed in the observation environment where the subjects were located.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Iga, Iloilo City, Philippines

TABLE OF CONTENTS

	Page
Title Page	i
Approval Sheet	ii
Acknowledgment	iii
Abstract	vi
Table of Contents	viii
List of Figures	xi
List of Equations	xiii
List of Tables	xiv
List of Appendices	xv
Chapter	
1 INTRODUCTION TO THE STUDY	
Background of the Study	1
Theoretical Framework	7
Objectives of the Study	27
Significance of the Study	28
Definition of Terms	29
Delimitation of the Study	35
2 REVIEW OF RELATED STUDIES	
Review of Related Studies	36

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Iga, Iloilo City, Philippines

3 RESEARCH DESIGN AND METHODOLOGY	
Description of the Proposed Study	58
Methods and Proposed Enhancements	
Assumptions and Preconditions	59
Methods	59
Components and Design	
System Architecture	76
Procedural and Object-Oriented Design	80
Process Design	86
Methodology	
System Development Life Cycle	
for Research Methodology	88
4 RESULTS AND DISCUSSION	
Implementation	90
System Inputs and Outputs	103
System Evaluation Results	113
5 SUMMARY, CONCLUSIONS, and RECOMMENDATIONS	
Summary of the Proposed Study and Implementation	117
Summary of Findings	118
Conclusions	121
Recommendations	122
References	124

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Iga, Iloilo City, Philippines

Appendices

Appendix A - Letter to the Adviser	136
Appendix B - Letter to the Technical Editor	137
Appendix C - Letter to the English Editor	138
Appendix D - Letter to the Thesis Format Editor	139
Appendix E - Letter to the Undergraduate Research Coordinator	140
Appendix F - Gantt Chart	141
Appendix G - Data Dictionary	142
Appendix H - Sample Program Codes	148
Appendix I - Software Quality Survey Form	171
Appendix J - Disclaimer	175

List of Figures

	Page
Figure 1: Functioning of a Neuron	9
Figure 2: A Schematic Diagram of the LSTM unit with forget gates	14
Figure 3: Human Body Motion Dynamic System	15
Figure 4: Proposed object Detection Framework	23
Figure 5: LSTM Neural Network Architecture	61
Figure 6: YOLOv3 Network Architecture	64
Figure 7: Convolutional Neural Network Model Architecture	65
Figure 8: Architecture of Fall Detection System	76
Figure 9: Architecture of Violence Detection System	77
Figure 10: Architecture of Weapon Detection System	78
Figure 11: Architecture of Drowsiness Detection System	79
Figure 12: Flowchart of Fall Detection System	80
Figure 13: Flowchart of Violence Detection System	81
Figure 14: Flowchart of Weapon Detection System	82

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Iga, Iloilo City, Philippines

Figure 15: Flowchart of Drowsiness Detection System	83
Figure 16: Use Case Diagram of the Proposed System	84
Figure 17: Data Flow Diagram of the Proposed System	86
Figure 18: Rapid Application Development Methodology	87
Figure 19: System GUI	104
Figure 20: No Fall Scenario for Fall Detection System	105
Figure 21: Fall Scenario for Fall Detection System	106
Figure 22: Non-Violence Scenario for Violence Detection System	107
Figure 23: Violence Scenario for Violence Detection System	108
Figure 24: No Weapon Scenario for Weapon Detection System	109
Figure 25: Weapon Scenario for Weapon Detection System	110
Figure 26: Open Eyes Scenario for Drowsiness Detection System	111
Figure 27: Close Eyes Scenario for Drowsiness Detection System	112
Figure 28: Survey Results in Google Sheets	120

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Iga, Iloilo City, Philippines

List of Equations

	Page
Equation 1	8
Equation 2	10
Equation 3	18
Equation 4	19
Equation 5	20
Equation 6	98

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Iga, Iloilo City, Philippines

List of Tables

	Page
Table 1: Possible Outputs from a Fall Detection System	18
Table 2: Deep Learning-Based Methods	26
Table 3: Results with the Roman Literals Defining the Features Used for Classification	92
Table 4: Activities of UP Fall Detection Dataset	93
Table 5: The comparison between the previous methods and the proposed method	96
Table 6: Experimental Result for trained deep learning models	99
Table 7: Comparison of classifier types in terms of Accuracy level	102
Table 8: Survey Results from Survey Evaluation	115

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Iga, Iloilo City, Philippines

List of Appendices

	Page
Appendix A - Letter to the Adviser	136
Appendix B - Letter to the Technical Editor	137
Appendix C - Letter to the English Editor	138
Appendix D - Letter to the Thesis Format Editor	139
Appendix E - Letter to the Undergraduate	
Research Coordinator	140
Appendix F - Gantt Chart	141
Appendix G - Data Dictionary	142
Appendix H - Sample Program Codes	148
Appendix I - Software Quality Survey Form	171
Appendix J - Disclaimer	175

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

CHAPTER 1 INTRODUCTION TO THE STUDY

Background of the Study

The global hit of the pandemic has taken a toll on people's everyday lives which caused them to shift to working from home, remote work, or being relocated for their jobs thereby creating a gap in giving their full attention to their loved ones, most especially the elderly. The elderly are considered at high risk when they get infected with the disease. One challenge that modern society is currently facing is the lack of development in the field of health care, especially for those older people who are in need of support for their daily activities as cited by Hamm, Money, Atwal, and Paraskevopoulos (2016).

In an article written by Goyala (2019), by culture, asian people tend to have a strong family oriented system, most especially Filipinos. We pay huge respect to our elders and have provided them with tender loving care. This statement is further supported by an article written by Shiho (2017) that "the idea that caring for older people is the responsibility of their children is rooted as firmly in

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Filipino society." The author also explained that this may be the reason why families still choose to care and live in the same household as their elders.

Many Filipinos traditionally live in a multi-generational household which emphasizes the idea of the aged-in-place, wherein many Filipino elders live under the same roof as their families as it is common for them to help raise their grandchildren (Growing old or just aging, 2012). They also added that "the continual support and even the later decisions that the family makes towards the health care of the elderly have a great impact, not only on the aging adult but on the entire family".

Hence, the researchers considered various ways to provide safety measures on this issue by identifying anomalies that can be found in day to day routines in caring for the elderly. These anomalies are the presence of weapons, violence, falling, and drowsiness. Paraskevopoulos et al (2016) defined that falling can be described as an unintentional or a sudden change of position of the body from an upright, sitting or lying position to a lower inclining position. As people age, it results in physical degradation,

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

hence, the ratios of elderly people suffering from falls are high and frequent, and are causing negative consequences to their health. As linked to falling, one common cause is drowsiness as some elderly take sleep medication every night and are more likely to fall and experience fall-related anxiety. Pharmacists' increased involvement in the community and pharmacist-led complete medication reviews are two initiatives that may minimize sleep medicine use and lead to fewer falls among the elderly (Nguyen & Watanabe, 2020).

Drowsiness is common in the elderly as their number of sleeping hours changes overtime. In the senior population, there is a special need to investigate the relationship between sleep duration and health consequences. Sleep physiology changes dramatically during life, and sleep length distributions shift with age (da Silva, de Mello, Schaan, Fuchs, F., Redline, & Fuchs, S., 2016).

Similarly, as worded in the study of Chen, Chiu, H.T., and Chiu, H. Y. (2016), falls were nearly twice more common in older persons with dementia than in those without dementia. Daytime sleepiness, daytime naps, and difficulties breathing during sleep were all higher in older persons with dementia

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

than in those without dementia. Despite correcting for putative risk variables, daytime sleepiness was the sole sleep feature that was substantially associated with an elevated risk of falls in older persons with dementia. In addition, Excessive Daily Sleepiness (EDS) made women more likely to report a fall, and the fall was more likely to occur outside. Following adjustment for the use of a walking aid, incidences of *Nocturia*, and antidepressant drug use, EDS was also linked to an increased risk of falling (Hayley, Williams, Kennedy, Holloway, Berk, Olsen, & Pasco, 2015).

On this note, it is also common for some people to have other people care for their loved ones, however, it is very risky for some if they encounter careless caregivers. In relation to this, was the article published by City News Service (2020) where a caregiver was intoxicated and accidentally shot the elderly woman he was caring for, where the latter was rushed to a hospital but later died due to the injuries sustained. Another similar article by Hutchinson (2021) which reported the death of an older man incurred by his caregiver through a blow to his head and stab wounds on his stomach and chest. As related to these incidents, the

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

researchers considered weapon detection as anomaly to be detected in the proposed system as there are many places where crime rate caused by guns or knives is very high, especially in places where there are no gun control laws. The early detection of violent crime involving weapons is of paramount importance for citizens' security (Kumar, Akshita, Arjun, & Geetha, 2021). In the system proposed by Kumar et al (2021), the yolov3 algorithm detects weapons in surveillance systems faster than the popular CNN, R-CNN, and faster CNN algorithms. Object detection has become one of the most fascinating fields in this age of automation. When it comes to surveillance system object detection, speed is crucial for swiftly detecting an object and alerting authorities. This effort attempted to attain the same goal, and it was able to produce a faster result than earlier systems.

However, anomaly for the elderly does not reflect on weapons alone, as they can physically be assaulted. In the study conducted by Hazrati, Mashayekh, Sharifi, & Motalebi (2020), the results of their study showed "a considerable rate of domestic abuse against elderly people, causing a serious risk for their health and security" as the results

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

indicated that 52.5% of the participants were female and 51.8% aged 60-69 years old. A total of 159 cases (39.8%) reflected at least one form of elder abuse or neglect. According to the findings of their study, 21% of the participants were mistreated by their own children. The most common type of abuse was care neglect, followed by psychological abuse, emotional neglect, and financial abuse. Motion constraints were the most prevalent type of neglect, followed by oral issues. The findings also revealed a link between domestic elder abuse and one's financial level. Locally, an article published by Coconuts Manila (2020) mentioned the risks of elderly abuse were high especially at the height of the pandemic as warned by the Commission on Human Rights. In the same article, the words of Commissioner Karen Gomez-Dumpit stated that "Around the world, there is emerging evidence that violence, abuse, and neglect of older persons increased due to the COVID-19 outbreak. In the country, we are still unaware of their actual situation within households, institutions, and communities." In addition, due to underreporting and insufficient research on the subject, data on elder abuse is scarce. According to her, a survey performed by the CHR in urban disadvantaged neighborhoods in 2004 found

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

that over 40% of older residents were subjected to various sorts of abuse, ranging from physical to verbal (Cabico, 2020).

In relation to the findings of these studies, the researchers proposed to create a system that detects anomalies, namely falling, weapons, violence, and drowsiness as security for elderly care which is much needed during the pandemic. In the words of family members, they need an extra eye on their elderly as they are unable to take care of them. They also need to observe their caretakers for security and assurance of family members that their elderly's welfare is taken care of.

Theoretical Framework

Neural Networks

A machine learning model inspired by the functioning and structure of a biological brain named neural network(NN), contains simple computational units called nodes or neurons. To produce an output, a neuron obtains inputs along the incoming edges, multiplies the inputs by corresponding edge weights, and then applies an *activation*

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

function which is classified as a nonlinear function to the weighted sum as illustrated in Figure 1.

Figure 1 is the working of a neuron and can be mathematically displayed by the vector equation 1, where x, w, b, f, y represent input vector, weight vector, neuron bias, element-wise multiplication, activation function, and neuron output respectively (Johansson, 2015).

$$y(x) = f(wx + b)$$

Equation 1

The equation 1 is meant to compute for neural network gradients in reference to the study of Johansson (2015).

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

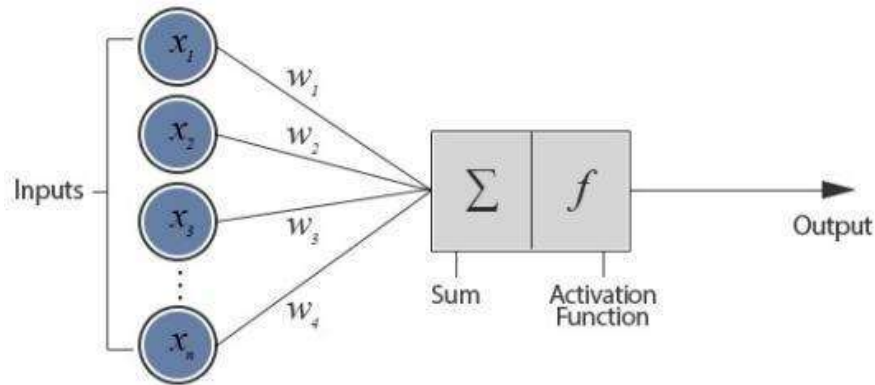


Figure 1. Functioning of a Neuron

The figure above is taken from the article of Jacobson (2013), which visualizes the functions of a neuron.

Training NNs

To minimize the loss function by tuning the parameters of the NN, the learning problem is transformed into an optimization or error minimization exercise. Gradient descent is the optimization algorithm used to train NNs, which includes calculating the gradients of the loss function with respect to the network parameters, that is, weights and biases. To calculate the gradients, back-propagation is used and is based on the chain rule of derivatives. According to equation 1, the learning rate, which is a scalar value, is used to update the parameters (θ) in the opposite direction of the gradient. The process is done through making

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

several passes over the training data. After every epoch, which is a pass over training data, the parameter advances closer to their optimum values which minimizes the loss function.

$$\theta = \theta - \gamma \frac{\partial L(\theta)}{\partial \theta}$$

Equation 2

The equation above discusses the formula used to train the neurons with regards to the study of Jacobson (2013).

Deep Learning and Deep Neural Networks

Feature engineering is a vital component of traditional machine learning whose deliberately designed features are requisite in conventional machine learning algorithms and do not work well with raw data. Nevertheless, feature engineering needs considerable domain expertise and is not straight forward. The ability to automatically learn high level representations pertinent for the task at hand is one of the main reasons for the accomplishment of the deep learning models. Deep neural networks (DNNs) are NNs with multiple hidden layers stacked together, where each layer is

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

a non-linear module that obtains the output of its previous layer.

Hence, a DNN is similar to a processing pipeline where each layer does part of the task and passes its output to the next layer. Deep learning techniques have imposed cutting edge results in a variety of domains from computer vision to language translation. Many different factors such as availability of large labeled datasets, advances made in computer engineering, distributed systems, and computational power including GPUs, contributed to this success.

Long-Short Term Memory (LSTM)

LSTM is able to grasp dependencies varying over arbitrary long time intervals. By replacing an ordinary neuron through a complex architecture called the LSTM unit or block, LSTM subdues the vanishing gradients problem. The LSTM unit is composed of simpler nodes connected in a particular way. In the study of Hochreiter and Schmidhuber (2010), the main components of the LSTM architecture are:

Constant error carousel (CEC). This is a central unit possessing a recurrent connection with a unit weight. A feedback loop with a time step equal to 1 is what recurrent

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

connection depicts. The internal state which is portrayed as the memory for past information defines the CEC's activation.

Input Gate. This is a multiplicative unit that secures the information kept in CEC from disturbance by unnecessary inputs; and

Output Gate. This is a multiplicative unit which secures other units from interference by information kept in CEC.

LSTM with Forget Gates

The LSTM unit's architecture with forget gates is shown in figure 1.2 as written by Hochreiter & Schmidhuber (1997). The main components of the LSTM unit are:

Input. The current input vector taken by the LSTM unit is denoted by x_t and the output from the previous time step denoted by h_{t-1} . Through tanh activation, the weighted inputs are summed and proceed, resulting in z_t (Hochreiter & Schmidhuber, 1997).

Input gate. x_t and h_{t-1} are read by the input gate, calculates the weighted sum, and utilizes sigmoid activation. To give the input flowing into the memory cell, z_t is multiplied with the result (Hochreiter & Schmidhuber, 1997).

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Forget gate. When memory contents become old and are no longer necessary, an LSTM learns to reset memory contents through the forget gate mechanism. x_t and h_{t-1} are read by forget gates and use a sigmoid activation to weighted inputs. The result, f_t is multiplied by the cell state at the previous time step, that is, s_{t-1} which allows for *forgetting* the memory contents which are no longer needed (Hochreiter & Schmidhuber, 1997).

Memory cell. This contains the CEC which has a recurrent edge with unit weight. By disregarding unnecessary information from the previous time step and receiving relevant information from the current input, the current cell state is calculated (Hochreiter & Schmidhuber, 1997).

Output gate. This utilizes the weighted sum of x_t and h_{t-1} and implements sigmoid activation to manage what information would flow out of the LSTM unit (Hochreiter & Schmidhuber, 1997).

Output. The h_t output of the LSTM unit is calculated through passing the cell state s_t through a tanh and multiplying it with the output gate as shown in Figure 2, o_t (Hochreiter & Schmidhuber, 1997).

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

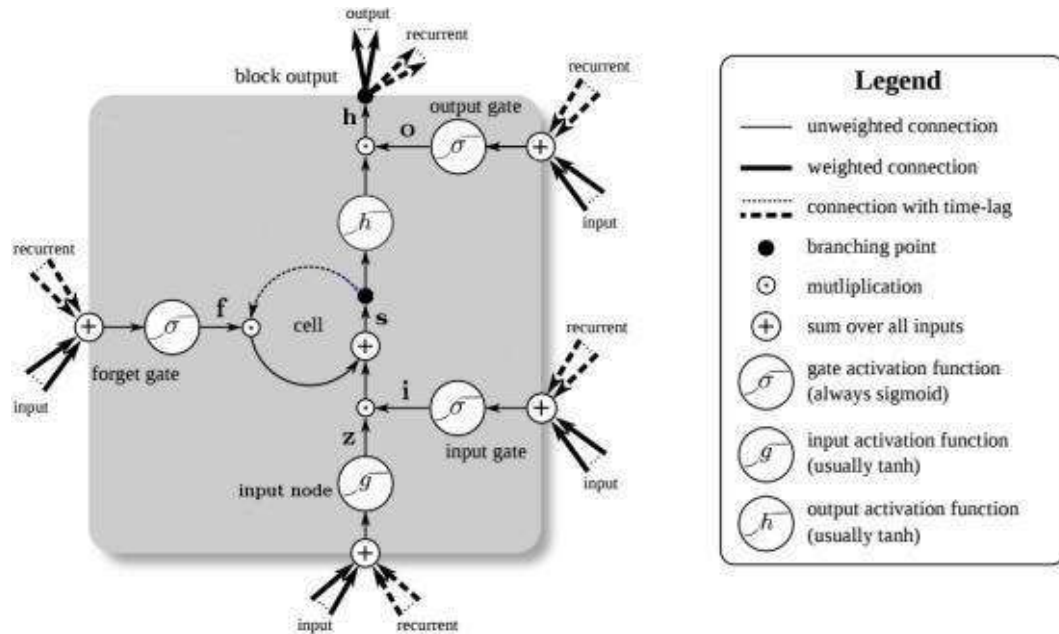


Figure 2. A Schematic Diagram of the LSTM Unit with Forget Gates.

Fall Detection

Human Motion Dynamics. Human Motion itself is complex and dynamics can hardly be characterized as motion is usually non-linear and time variant, given that day to day fitness varies over time. Studying human motion can take place through motion sensing cameras. Frequency components and the amplitude of the signals from sensors would be analyzed and requirements on monitoring of human motion could be acquired.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

The coordinate system of the human body could be further defined in Figure 3 (Johansson, 2015).

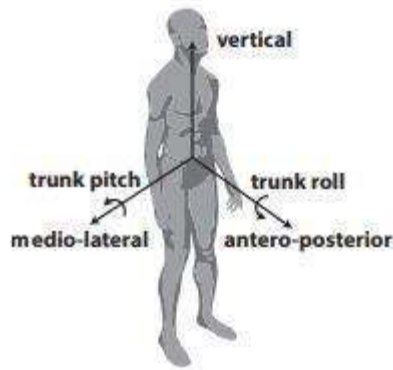


Figure 3. Human Body Motion Dynamic System

The figure above represents the indicators in the human body to detect points (Johansson, 2015).

Human Fall Properties. A human fall is described as a situation when a person unknowingly comes to recline on the floor or other lower level. The fall could be classified into 4 phases. The pre-fall which signifies the beginning of the fall. Next, the critical phase which indicates the impact of the pre-fall. Then, the post-fall which is when the person remains idle reclining on the floor after the impact. Lastly, the recovery phase which shows the person independently or by assistance of another regain footing. It was stated in the survey of the fall detection system in the previous years

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

that there were three important fall scenarios for the elderly. Fall from an upright position, chair and bed. During these falls, the critical phase time differs from 1-4 seconds. Moreover, in the previous years, fall detection was used to develop a mobile human airbag system. The system must execute the whole process of fall detection, mechanical triggering and airbag inflation within 0.9s to be able to protect the user during the fall. Their algorithm, with the worst case measure of triggering and inflating must detect the fall 330ms precedent to impact (Johansson, 2015).

Fall Detections Systems and Motion Monitoring Systems. This study made use of the parameters used in evaluating the system performance while discussing its disadvantages as well as its advantages. It would also cover pre-fall detection, the focus would which will be more on its monitoring system as used in detection motion, specifically falls as an anomaly (Johansson, 2015).

As cited from Howcroft et al. (2013) the working principle of the said system is data collection of human movement, which would be analyzed further based on specific features to extract and evaluate. Since a fall detection

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

system falls under motion monitoring, its extracted feature is to detect fall, which relies on the application. One example of this is analyzing the amount of movement, which can also be referred to as energy expenditure analysis. This would be useful in monitoring overall physical activity. Motion monitoring is perceived as useful in assessing the risk of falls (Johansson, 2015).

In the words of Pannurat et al (2014), automatic fall detection systems studies have been done since the 1990s and the first ever prototype was built in 1998. Human movement is analyzed by a fall detection system thoroughly to monitor if a fall event has occurred or is currently occurring. The purpose of this is to reduce severe effects of the impact through a notification system that would call the attention of medical care or hopefully reduce the force of the impact. Through this, the detection can be divided into two parts: pre-impact and post-impact detection, the former is during the critical phase and the latter is the recovery phase. Two different approaches were presented in the field: Threshold Based Methods (TBM) and Machine Learning Method (MLM). This was further supplemented and based on the proposal by Noury

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

et al (2008) where the two possibilities are presented in Table. 1.

Event	A fall occurs	A fall does not occur
A fall is detected	True positive, TP	False positive, FP
A fall is not detected	False negative, FN	True negative, TN

Table 1. Possible outputs from a fall detection system

To further evaluate the performance, the indexes are presented and defined with the following:

Sensitivity, which analyzes the capacity in order to detect a fall with the formula in which the number of detected falls is divided by total number of falls.

$$SE = \frac{TP}{TP + FN}$$

Equation 3

Specificity, which analyzes the capacity to detect a fall only during occurrence. This is done to lessen false alarms.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

$$SP = \frac{TN}{TN + FP}$$

Equation 4

False Positive Rate, which will be labeled as FPR, is another parameter that would describe the rate of false positivity for the dataset provided.

Similarly, as taken from Zhao et al (2012), the performance of fall detection systems shows a comparison between sensitivity and specificity. However, the presence of live constraints would affect the performance of the fall detection system's pre-impact analysis as the evaluation is time lapsed between the impact and the detection. This performance parameter could also be referred to as Pre-impact Lead Time (PLT). Additionally, Wolk Company (2015) further strengthened the study of Zhao et al (2012) as they mentioned that acceptable performance could be determined through sensibility, pre- impact lead time, and specificity in their research. They have identified at least two solutions that could aid in the development of fall protection systems, yet

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

there were no available sources and information available with regards to how fall is detected.

$$T_{PLT} = T_{impact} - T_{detect}$$

Equation 5

Weapon Detection Security

No place in the world is absolved from the repercussions of firearm violence. The problems related with this violence covers the entire range of human security, from domestic and street violence with severe economic and social consequences for the society at large, to large-scale armed conflicts in which these arms enable widespread violence and account for the majority of deaths.

Weapon Detection in Images and Videos

Iain Darker and his team in 2007 were first to study the concept of firearm detection with surveillance cameras. Same team later in 2008 attempted to apply automated surveillance methods for the first time. Darker proposed a SIFT based firearm detection algorithm using motion segmentation method of ROI estimation. After determining ROI, SIFT algorithm was

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

applied within the ROI to detect the firearm (pistol). Another category of weapons along with firearms that can be assessed for detection is knives. Knife detection research is also being accomplished in recent years, however there is a limited number of papers published in the field of knife detection so far.

Two different algorithms for knife and gun detection were described by the authors. Edge histograms descriptors were used for knife detection. Gun detection was done on image, detecting humans in the image first, then the gun in the person's hands. An algorithm that used edge histogram feature extraction and SVM based classification for knife detection was proposed by Grega et al. Four fold cross-validation with 2627 samples in each was used to validate the results of this algorithm. Therefore, a total of 10508 images from the dataset containing 12899 images were used. Results were recorded in terms of specificity and sensitivity, where 81% is sensitivity and 93% specificity were achieved.

Machine Learning - Neural Network for Object Detection

IVS (intelligent Video Surveillance) systems need to detect threats and consider appropriate measure(s) with

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

comparatively minimal delays to be effective. Lee et al proposed a machine learning approach using convolutional neural networks (CNNs) in a hierarchical feature model (HFM). CNN is defined as the feed-forward neural network that comprised a group of receptive fields, emulating the manner in which brain neurons are organized in a visual context. A neural network comprises several nodes which each carry a different weight.

N nodes in a neural network in its classical form, are used to categorize an object with N features. A CNN is a more efficient variation that lessens the amount of nodes, hence the computational time. To reduce the number of the requisite codes, a series of convolution and max- pooling layers are used. Convolution is a process that aims to extract features that precisely represent the responses of the nodes. Consequently, items that are given as input to the network will produce the same classification outcome. Max-pooling, a down-sampling method that lessens the number of nodes through keeping the maximum values of neighboring nodes. The maximum weights, hence providing more to the distinction between classes, were utilized to constitute their respective areas

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

and the nodes with lesser weights were dropped from the generated model.

The CNN generated HFM (hierarchical feature model) was used to create a hierarchical classifier ensemble (HCE), which eventually performed the object detection. The object detection framework can be visualized in the Figure 4 below:

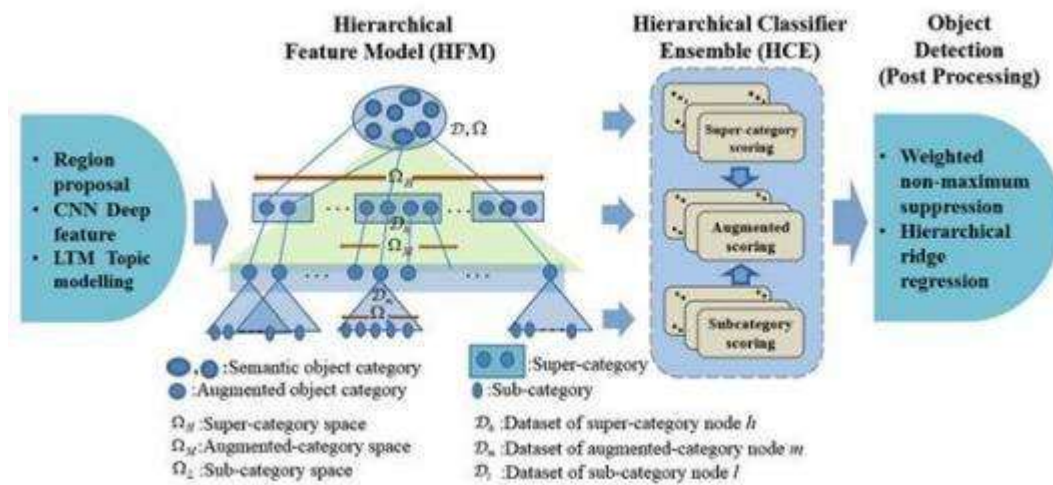


Figure 4. Proposed Object Detection Framework Proposed

A visual diagram by Lee Et Al., Based On the Hierarchical Feature Model (Hfm), And Hierarchical Classifier Ensemble (Hce). The authors identified three advantages of their proposed framework:

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

- They included an augmented object category that resolves issues with inter-class ambiguity and intra-class variation;
- HFM which had shown to be more effective coupled with the HCE, because HFM's clustering facilitates building the HCE; and
- Confusing data samples that were clustered properly to sub-categories and overall detection accuracy which can be improved.

Violence Detection

Deep Learning Based Methods for Violence Detection. In terms of over trajectory-based and non-object centric methods for video activity detection, deep learning methods obtained high accuracy. Deep learning models handle feature selection and classification as a single module where using feature extractors or descriptors separately was not necessary. According to Zhang et al, Deep learning techniques have obtained more recognition and popularity among other techniques to solve the challenges in violence detection. Unsupervised learning techniques were used as deep learning methods, like deep belief networks, recurrent neural

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

networks, CNNs, and long short-term memory (LSTM), for activity recognition.

A RNN in the study of Traoré and Akhloufi was used in the deep learning framework for recognizing abnormal human actions and attained an accuracy of 91.43%. Another deep learning framework in the study of S. Sudhakaran and O. Lanz, used LSTM to recognize violent activity. 94% accuracy was achieved from the three standard data sets, which were the hockey fight dataset, violent flow dataset, and the movie dataset. A simple deep neural network framework was proposed using the Weber local descriptor to extract optical flow by Mondal et al to detect violent activities. Using a crowded violence dataset, an accuracy of 90% was achieved. Additional deep learning framework by R. Halder and R. Chatterjee to detect violent activities from videos used CNN Bi-LSTM. Using three widely used datasets of hockey fights, movies, and violent flow the authors obtained an accuracy of 94%.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Classification and Feature Selection	Conv Layers /FC layers	Acc uracy
Recurrent Neural Network (RNN)	3/2	91%
Long short-term memory (LSTM)	5/3	94%
Simple Deep Neural Network	4/1	90%
CNN Bi LSTM	3/2	9%

Table 2. Deep Learning-Based Methods

Table 2 lists the different trajectory-based approaches for violence detection.

Drowsiness Detection

Drowsiness Detection Using Deep Learning. Recently, deep learning has been widely used to solve large issues that cannot be rightly solved by conventional methods. It made a significant breakthrough when DL focused on the CNN for computer vision tasks including classification of images, object recognition, emotion recognition, scene segmentation

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

in general. CNN had best implemented, according to the study of Dwivedi et al. for drowsy driver detection with a 78% accuracy. Three networks for modern architecture were suggested in the study of Park et al. The first network to learn the image functionality was AlexNet composed of 5 CNNs and 3-layer FC. The second network used to extract facial features was 16-layer VGG-FaceNet. The third network extracted computation features using Flow ImageNet.

Objectives of the Study

The main objective of this study was to design and develop a program that would detect and predict anomalous human activity and sound an alarm to notify the person monitoring the elderly.

Specifically, this study aimed to:

1. Design a deep learning model for anomalous human behavior detection using Long-short Term Memory (LSTM), and Convolutional Neural Networks (CNN).

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

2. Develop a video surveillance desktop software integrated with the proposed deep learning model for anomalous human behavior detection.
3. Incorporate an Emergency Alert System to the video surveillance desktop software.
4. Implement the proposed system and perform user testing and evaluation using ISO-standard Usability Assessment Tool based on system functionality, performance, usability, reliability, and maintainability.

Significance of the Study

This study aimed to develop a surveillance system that detects anomalies, specifically for elderly people who were left alone or with caregivers. The system was designed to detect specific anomalies, which were falling accidents, weapons and violence and which would automatically notify responders for immediate medical response.

Benefitting the study were the various sectors as follows:

Patients. The study could provide security and assistance to the elderly when they are in unsupervised situations or with their caregivers.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Family of the Patients. The results of this study could provide extra precaution in terms of safeguarding their elderly loved ones and may assure the family members who may not be constantly present for them.

Hospitals. The results of this study could be an additional hospital resource to address the physical and emotional needs of the patients.

Retirement Homes. This study may greatly help with the monitoring of elderly people. It could assist in a way of sounding alarms upon detection of anomalies which can be immediately responded to and notify the family of the elderly immediately.

Future Researchers. Future researchers may use this study as a reference to research of similar design. The recommendations of this study may be used by future researchers to improve results of their findings as well.

Definition of Terms

For better understanding, the following terms were defined conceptually and operationally:

AI/ML. This is short for Artificial Intelligence (AI) and Machine Learning (ML). AI serves as a simulation of human

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

intelligence done by machines and ML is a concept to make data driven decisions (Cohen, 2020).

In this paper, AI/ML is used as a technique for fall detection.

Alert System. This is an automated method of contacting a group of people within an organization and distributing important information during an emergency situation (Crocetti, P., n. d.).

In this study, an alert system is a system that notifies people to contact about the detected fall.

Anomalous. The term is defined as inconsistency with or deviating from what is usual, normal, or expected (Merriam-Webster, 2021).

In this study, anomalousness is seen through the falling of the elderly, and other indicators described in this study.

Convolutional neural networks (CNN). This is a type of deep learning model for processing data with a grid pattern, such as images, that was inspired by the organization of animal visual cortex and designed to learn spatial hierarchies of features, from low-level to high-level patterns, automatically and adaptively (Yamashita, Nishio, Do, & Togashi, 2018).

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

In this paper, CNN is used as the algorithm to recognize and detect signs of drowsiness.

Deep Learning. This term is under Machine Learning in artificial intelligence that imitates how works and processes can be done by the human brain and create patterns that would be used in decision making (Hargrave, 2021).

In this study, deep learning is defined as the specific model used to design the system.

Drowsiness. It refers to feeling abnormally sleepy during the day. People who are drowsy may fall asleep in inappropriate situations or at inappropriate times (MedicinePlus, 2021).

In this study, drowsiness refers to one of the anomalies for detection where the subject manifests sleepiness while doing daily routines.

Drowsiness Detection. This is referred to as identifying sleepiness. This is accomplished in most versions of this feature by keeping note of how frequently you close your eyes over a given period of time (MyCarDoesWhat.org, 2015).

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

For this study, Drowsiness Detection refers to how often a subject shows signs of falling asleep, in terms of closing their eyes.

Fall. This is defined as an event which results in a person coming to rest inadvertently on the ground or floor or other lower level (Oxford Dictionary, 2021).

In this study, fall is defined as an unintentional or sudden change of position of the body to a lower inclining position, an anomalous accident that is common amongst elderly people due to health degradation caused by aging.

Fall Detection. This term refers to a system that can be an assistive device whose main objective is to alert when a fall event has occurred (Igual, Medrano, & Plaza, 2013).

For this study, fall detection is defined as the system used to observe a fall before, during and after it happened.

LSTM. This could also be referred to as Long Short Term Memory networks. This is a modified version of RNN as it can smoothly remember past data through predicting sequence problems (Mittal, 2019).

In this study, LSTM was used to predict whether the person has fallen or not.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Surveillance. This is an act that keeps close watch over someone or something (Merriam-Webster, 2021).

In this study, surveillance is the act of monitoring senior citizens in a closed location.

Violence, as defined by World Health Organization, as the intentional use of physical force or power, threatened or actual, against oneself, another person, or against a group or community, that either results in or has a high likelihood of resulting in injury, death, psychological harm, maldevelopment or deprivation (SaferSpaces, 2021).

In this study, Violence is the physical harm done to another individual specifically to the elderly.

Violence Detection. This is a technology that can be easily integrated with any security system. Its primary function is to ensure public safety through visual crowd surveillance, so any violent activity generates automatic security and police alerts (Abto Software, 2018).

In the present study, the Violence Detection in this study is incorporated into the surveillance systems to detect any form of physical harm.

Weapons. As defined by Macmillan Dictionary (2021), a weapon can be used to harm people or cause property damage.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

In this study, weapons was meant to be in the form of objects that can potentially risk a person's safety. For example, a gun, knife, pen, pepper spray, etc.

Weapon Detection refers to the use of AI and deep learning algorithms for security applications. Such vision-based systems can recognize and interpret scenes using the footage of video surveillance systems. AI vision methods were used to recognize knives and guns with the goal to reduce crimes and increase safety and security (Viso.ai, 2021).

In this study, Weapon Detection means the use of the YOLOv3 algorithm to detect objects that can be classified as weapons.

You Only Look Once, Version 3 (YOLOv3). This is a real-time object detection algorithm that recognizes specific things in films, live feeds, and photos (Meel, 2021).

In this paper, YOLOv3 is defined as the algorithm used to analyze weapons in the detection system.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Delimitation of the Study

The researchers proposed a system intended for elderly subjects and was initially planned to be implemented on homes, hospitals, or retirement homes. One problem the researchers faced was the lack of available data sets for training for each model present in the system. So instead, the researchers only used the available data set found in the internet for training the models. Additionally, due to the nature of the system which required a high-end device to train, and with the lack of said resources on the part of the researchers, they decided to train the model based on the capability their current resources can handle.

Moreover, the researchers chose respondents based on their availability to test the systems instead of implementing the system on different homes, hospitals, or retirement homes due to the restrictions caused by the Covid-19 pandemic. During the testing phase, ten (10) respondents were test subjects for the system, and were able to answer a survey. The survey questionnaire was composed of different criteria where the respondents rated the system based on their observation of the system.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

CHAPTER 2 REVIEW OF RELATED STUDIES

Review of Existing and Related Studies

Chapter 2, Review of Related Literature, contains existing and related studies that were reviewed in relation to the study entitled, "*AI/ML for Anomalous Surveillance: A Detection and Alert System.*" This chapter will further elaborate the application and comparisons of the said studies in this thesis paper, the researchers will compare the following from other related studies: (1) AI/ML, (2) Anomalous Surveillance, (3) Fall Detection System, (4) Weapon Detection System, and (5) Violence Detection System.

AI/ML

Artificial Intelligence

Since the question "Can machines think?" was raised by Alan Turing in 1950, humans have been concerned with the development of artificial intelligence (Roy, 2020). The simulation of natural intellect in robots that were intended to learn and emulate the actions of humans was what artificial intelligence is. These machines are capable of learning from their mistakes and doing jobs that

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

were similar to those performed by humans. As technologies such as artificial intelligence (AI) continue to develop, they will have a significant impact on our overall quality of life. It's only natural that everyone today wants to connect with artificial intelligence technology in some way, whether it's as an end-user or as someone who wants to pursue a career in artificial intelligence (Advani (2021)).

In 1956, John McCarthy convened a summer workshop called the Dartmouth Summer Research Project on Artificial Intelligence, which brought together researchers from a variety of disciplines, including language simulation, neuron nets, complexity theory, and others, to discuss what would eventually become the field of artificial intelligence. The workshop was the first time the term "artificial intelligence" was used. As a result of this collaboration, the researchers were able to clarify and develop the concepts around "thinking machines," which had been rather divergent up until that point. McCarthy was said to have chosen the term artificial intelligence because of its neutrality; he did not want to draw attention to any of the tracks that were

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

being pursued at the time for the field of "thinking machines," which included cybernetics, automata theory, and complex information processing, among other things (Marr, 2018).

Modern dictionary definitions of Artificial Intelligence (AI) emphasized the fact that it is a subfield of computer science and that robots can simulate human intelligence (being human-like rather than becoming human). According to the English Oxford Living Dictionary, artificial intelligence is defined as "the theory and development of computer systems capable of performing tasks that would normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages."

Machine Learning

As defined by Roy (2020), Machine Learning is a subset of Artificial Intelligence that employs statistical learning algorithms to create systems that are capable of autonomously learning from and improving on their own without the need to be explicitly programmed. Because it provides

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

businesses with a better understanding of trends in customer behavior and business operational patterns, machine learning is essential for the development of new goods. Today's major organizations, such as Facebook, Google, and Uber, place a high priority on machine learning as a critical component of their operations. As a competitive differentiation, machine learning has become more important for many businesses (Burns, 2021).

Brown (2021) discussed that chatbots and predictive text, language translation applications, the shows Netflix recommends to you, and the way your social media feeds are presented are all powered by machine learning. It fuels self-driving cars and robots that can identify medical issues based on photographs taken by the driver. Moreover, she reiterated that when corporations today deploy artificial intelligence programs, they are almost certainly utilizing machine learning techniques – so much so that the phrases are frequently used interchangeably and occasionally ambiguously in the context of artificial intelligence. Machine learning is an area of artificial intelligence that allows computers the ability to learn without having to be explicitly taught

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

in the process.

AI/ML Relationship

Artificial Intelligence (AI) and Machine Learning (ML) are two extremely popular buzzwords right now, and the terms are frequently used interchangeably in the same sentence. It is true that they are not exactly the same thing, but the perception that they are can lead to some confusion at times. When it comes to Big Data, analytics, and the broader waves of technological change that are sweeping through our world, both terms are routinely used interchangeably (Marr, 2016)

In a nutshell, the simplest response is that Artificial Intelligence is a wide notion that refers to computers that are capable of performing tasks in a manner that we would consider "clever." Machine Learning is a recent application of artificial intelligence that is founded on the premise that we should be able to simply provide machines with access to data and allow them to learn for themselves.

Parthasarathy (2019) discussed that, machine learning makes use of past experience to search for patterns that it

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

has learnt. AI learns from its experiences in order to gain information and skills, as well as how to apply that knowledge in new situations. Both artificial intelligence and machine learning have the potential to be useful in the commercial world. However, machine learning (ML) has seen a significant increase in its use in many firms to solve crucial business problems.

Anomalous Surveillance

In the article released by the PreScouter Editorial Team (2021), it was stated that Artificial intelligence and machine learning technologies have been used in a variety of industries and disciplines, including healthcare, finance and governance, sports and entertainment, mining, construction, logistics, food, and fashion, to name a few. Using automated processes and robotics technology, AI and ML technologies have been particularly successful in automating digital and physical tasks. One area in particular that has gained popularity is the application of artificial intelligence in anomalous surveillance or anomaly detection. Anomaly detection has been effectively used to improve the efficiency

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

of operations in a variety of different sectors.

Anomaly detection, as defined by Johnson (2020), is any technique that discovers the outliers in a dataset; that is, the objects that do not belong in the dataset. These anomalies could indicate unexpected network traffic, reveal a sensor that is malfunctioning, or simply identify data that needs to be cleaned before being analyzed.

Branch et al., (2020), defined anomaly detection as the process of identifying anomalous patterns in data that differ from the rest. These patterns are referred to as anomalies or outliers. An increase in the number of attributes or features complicates the detection of anomalies because the amount of data required to generalize accurately increases as the number of attributes or features increases. This leads to an increase in data sparsity, which is characterized by data points that are more scattered and isolated.

In a study conducted by Hao et al., (2020), it was stated that in recent years, there has been a great deal of research into anomaly event detection in computer vision. In videos, the majority of standard anomaly event detection systems can

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

only take use of single-modal signals, and thus are incapable of dealing with the complimentary information underlying additional modalities. In addition, detection of anomalous events in surveillance videos has been widely applied in a variety of security-related scenarios, such as traffic accident investigation, criminal or illegal activity surveillance, forensics investigation, and violence alerting. Given the rarity with which anomalous events manifest themselves in real life, abnormal patterns of behavior or appearance that differ from the usual are frequently referred to as anomalies. In the article by Choudhary (2017), he discussed that in order to discover anomalies in data, the most straightforward technique is to locate data points that deviate from typical statistical features of a distribution, such as the mean, median, mode, and quantiles, and then flag those data points. Assuming that the definition of an anomalous data point is one that deviates from the mean by a specified standard deviation is correct. Because time-series data is not static, traversing the mean over it is not a straightforward task. To compute the average over all of the data points, you would need to use a rolling window. When

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

referred to in technical terms, a rolling average or a moving average is used to smooth out short-term oscillations while drawing attention to long-term fluctuations. Alternatively, a mathematically defined "low pass filter" can be used to describe the behavior of an n-period simple moving average.

Fall Detection System

Falling is among the most damaging events elderly people may experience. With the ever-growing aging population, there is an urgent need for the development of fall detection systems. Thanks to the rapid development of sensor networks and the Internet of Things (IoT), human-computer interaction using sensor fusion has been regarded as an effective method to address the problem of fall detection (Azzopardi et al., 2020). In the research of Peng and Ren (2019), they discussed that detecting falls and preventing falls are two essential ways for dealing with the problem of elderly people falling, and they have been studied extensively over the past two decades. Researchers have spent a great deal of time into fall detection technologies in particular. While numerous types of sensors were employed in these systems, in order to

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

acquire usable signals for subsequent processing and analysis, various analytical algorithms were employed in order to process the data collected. Accelerations are often used by the majority of fall detection systems to detect the shock induced by the body impact.

Harari et al., (2021) conducted a prospective study to investigate the performance of a smartphone-based online, real-time fall detection system. According to the study's findings, the system they built comprises a machine-learning classifier for detecting falls based on data from the smartphone's accelerometer and gyroscope. In addition, the system incorporates an activity recognition model, a position tracker, a web portal for data investigation, and a real-time online notification system, which creates a timely notification of the fall and the faller's status. The system was also equipped with a fall detection system.

In their study, it was concluded that because cellphones are so common, a wide global population may be able to employ this fall detection system. If a faller is severely injured

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

or unconscious and unable to call for help, the system's capacity to automatically detect, locate, and notify them may be vital. The system's fall- related data kept in the online portal may provide new insights for future studies on fall prevention, detection, and treatment.

Shu and Shu (2021) discovered, after studying the physics of falls, that a few characteristics, such as velocity, acceleration, and the motions of the head and legs, are highly connected with the occurrence of falls. It is possible to utilize any machine learning algorithm to recognize falls by extracting these factors and using them to improve image semantics-based features, as the researchers have done. As a result of this theory, they are able to apply a parsimonious artificial intelligence model to do exceptionally efficient video processing at high frame rates, thereby overcoming the current problems in fall detection accuracy, computing efficiency, and financial cost.

Chelli and Pätzold (2019) proposed a machine learning strategy for ADL recognition and fall detection. Based on acceleration and angular velocity data, they evaluated four

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

algorithms' performance in recognizing falling, walking, going upstairs, walking downstairs, sitting, standing, and lying. They introduced new temporal and frequency domain characteristics and demonstrated their value in improving classifier accuracy and precision. The researchers also evaluated the algorithms' performance on real-world acceleration data from public databases. These algorithms' core parameters were optimized using training data. Afterwards, the trained algorithms' performance was evaluated using test data. Initially, activity recognition relied solely on acceleration data. A 66-size feature vector was produced and sent into the classification method. Their results show that the KNN, ANN, QSVM, and EBT algorithms are 81.2%, 87.8%, 93.2%, and 94.1% accurate. They then improved the performance of the four classification systems by extracting new features from the acceleration and angular velocity data. The built-in feature vector is 328 bytes. Using the proposed feature vector, we found that the KNN, ANN, QSVM, and EBT algorithms achieve overall accuracy of 85.8%, 91.88%, 96.18%, and 97.7%. The QSVM and EBT fall detection accuracy is 100% with no false alarms, which is the best possible

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

performance.

According to Becker et al., (2020), in order to develop and evaluate the suggested fall detection systems, the majority of fall detection studies used simulated falls (often done by young individuals). Such an approach leads to concerns with generalize ability and transferability of results from simulated to real-world contexts, which has been the subject of recent research articles. A simulated environment can, in fact, conceal a number of potential flaws that might otherwise become apparent in a real-world scenario. In reality, simulations are carried out in a controlled setting, which allows for the development of a standardized and clean version of the problem of fall detection. The opposite is true in an uncontrolled setting where a variety of procedural and technical difficulties must be dealt with.

The prevalence of this problem was demonstrated by Chaudhuri et al., (2015), who tested a wearable device for fall detection on 18 participants. Fall detection proved unreliable, especially when considering the sensitivity of

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

the system (percentage of falls correctly detected). In contrast to the manufacturer's declared sensitivity of between 94.1 percent and 94.4 percent (tested using simulated falls from 59 people), the sensitivity evaluated on real-world falls plummeted to a value of only 25 percent. Following their findings, the authors argue that "clinicians dealing with older persons should evaluate the availability (and accuracy) of real-world testing for any fall detection systems before recommending them to their patients."

Violence Detection System

The presence of aggressive behavior in public areas poses a major threat to human safety as well as social stability. At the moment, millions of pieces of equipment are being used in public locations, putting a tremendous amount of strain on the security personnel. The ability to automatically detect violent occurrences from enormous amounts of surveillance video data is therefore extremely important (Zhou et al., 2018). The footage captured by surveillance cameras are subjected to computer vision techniques for the detection of violence. The installation of

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

cameras and other surveillance equipment in various locations for public safety, such as educational institutions, hospitals, banks, markets, and streets, has become more common in recent years. The activities of people are monitored by these cameras and other surveillance equipment (Sultani et al., 2018).

With the rapid expansion of surveillance cameras in various spheres of life to monitor human behavior, there is an increasing demand for systems that can automatically detect violent occurrences and alert the appropriate authorities. Violent action detection is becoming a popular study area in computer vision, attracting new researchers. Indeed, a large number of academics developed a variety of strategies for detecting such behaviors from video footage. It is the primary purpose of the systematic review undertaken by Ramzan et al., (2019) to investigate the most recent advances in research in the field of violence detection systems. The systematic study provides specifics on methods for detecting violence using SVM, CNN, and classic machine learning classification-based methods, among others. These strategies are outlined in depth, and the advantages and

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

disadvantages of each are discussed. Furthermore, thorough tables list all of the datasets and video attributes that were employed in all of the strategies and that were found to be important in the recognition process. The accuracy of object identification, feature extraction, and classification approaches, as well as the dataset being used, are all dependent on the techniques used. Their research has the potential to contribute to the advancement of strategies and procedures for detecting violent activity in surveillance videos, which are now under investigation.

In the recent study conducted by Ahmed et al., (2021), it was reiterated that because of its applicability in activities relating to security and law enforcement, the recognition of violence is essential to these endeavors. Existing semi-automated systems have flaws, such as the need for time-consuming manual surveillance, which leads to human error and reduces the effectiveness of the system. Different approaches have been offered, including trajectory-based methods as well as methods that are not centered on objects and deep-learning methods. Previous researches have demonstrated that deep learning algorithms achieve higher

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

accuracy while also having lower error rates than previous techniques. Their overall performance, on the other hand, must be improved.

According to the findings of their study, a deep learning framework for recognizing violent action from video was proposed. When it came to feature selection and classification, the suggested framework utilized the keyframe extraction technique to eliminate duplicate frames and S-CNN and inception v4 CNN for feature selection and classification. Extensive experiments were carried out to verify the validity of the proposed paradigm. The results of their study reveal that key frame extraction can eliminate as much as 25% of duplicate frames in some cases. The classification model achieved an accuracy of roughly 98 percent in the testing process. As a result, sequential CNN and inception v4 are more effective at recognizing violent actions in videos than other methods. In a subsequent investigation, the proposed technique will be applied to the detection of other aberrant activity.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Weapon Detection System

Gun violence has substantial public, physical, psychological, and economic costs. Every year, gun violence kills many. Psychological trauma is common among youngsters exposed to high amounts of violence in their communities or media. Children who are victims, perpetrators, or witnesses of gun violence can suffer short- and long-term psychological repercussions. Studies reveal that handguns are commonly used in crimes like robbery, shoplifting, and rape. These offenses can be reduced through early detection of disruptive conduct and rigorous monitoring of suspicious actions by law enforcement agents (Velaustin, 2006).

Security is always a major worry in any sector, owing to an increase in crime rates in crowded or suspiciously lonely regions, which is especially true in the construction industry. Gun violence is a contemporary human rights issue that affects people all around the world. We are in danger of losing our most fundamental human right, the right to life, as a result of gun-related violence. People all across the world are affected by gun violence on a daily basis, and it is a tragedy that affects their lives. Every day, more than

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

500 individuals are killed as a result of acts of violence committed with weapons. The easy availability of firearms has always been and will continue to be a significant element in the rise in crime and disorder (Akash Kumar et al., 2021).

In the study of Bhatti et al., (2021), it was stated that the ability to maintain control over the law-and- order situation is essential for a country's development. Whether we want to attract investors for investment or generate revenue through the tourism industry, all of these objectives require a tranquil and secure atmosphere to be accomplished. In many places of the world, the crime rate is extremely high as a result of the availability of firearms. It mostly includes countries where it is allowed to own and carry a handgun on one's person. The world has become a global village, and everything we say or write has an impact on the individuals who hear us or read us.

The YOLO V3 object detection model for weapon detection was implemented and trained over the dataset based on the study conducted by Narejo et al., (2021). In their proposal, they suggested a model that delivers visionary sense to a machine or robot in order for it to recognize a dangerous

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

weapon. The model may also inform the human administrator when a gun or a firearm is visible in the vicinity. The experimental results demonstrated that the trained YOLO V3 model outperforms the untrained YOLO V2 model in terms of performance while being computationally less expensive. Improvements in surveillance capabilities, as well as more resources to support monitoring the effectiveness of human operators, are required immediately. Smart surveillance systems would completely replace current infrastructure due to the increasing availability of low-cost storage, video infrastructure, and superior video processing technologies. Smart surveillance systems would completely replace current infrastructure. The increasing availability of low-cost computing, video infrastructure, high-end technology, and improved video processing will eventually lead to the complete replacement of current surveillance systems by digital monitoring systems, which will be implemented as robots.

Bhatti et al., (2021) conducted a study for both monitoring and control objectives; as a result of this work, a revolutionary automatic weapon identification system in

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

real time was presented. This work undoubtedly contributed to the improvement of the security, law and order situation for the betterment and safety of humanity, particularly in the nations that have suffered greatly as a result of these types of violent activities in the past. This may also have a good impact on the economy by attracting investors and tourists, who place a high value on security and safety as their top priorities. In relation to this, the researchers concentrated their efforts on recognizing the weapon in live CCTV streams while simultaneously reducing the number of false negatives and positives. The researchers also used two approaches to achieve high precision and recall: first, they created a new training database for the real-time scenario, then trained and evaluated it on the most recent state-of-the-art deep learning models using two approaches, namely, sliding window classification and region proposal/object detection. In order to achieve high precision and recall, a variety of algorithms were tested. Following a series of trials, it was discovered that object detection algorithms with Region of Interest (ROI) outperform algorithms without ROI in terms of performance. Many models have been tried, but

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

the most effective model was the cutting edge Yolov4, which was trained on the new database and produced the fewest false positive and negative values. As a result, it produced the most successful results of all the models examined.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

CHAPTER 3 RESEARCH DESIGN AND METHODOLOGY

Description of the Proposed System

In this work, the researchers presented a vision- based surveillance system for detecting anomalous human behavior like fall accidents, violence, drowsiness and carrying of weapons for monitoring elderly people. This detection system could be easily integrated into existing surveillance systems and web cameras to provide caregivers with real-time information.

The system made use of live video feedback from the camera installed in a room. Each of the detection systems analyzed the video feedback, the fall detection and violence detection systems. Each went through an LSTM Neural Network to predict the specific anomalous activity, such as the weapon detection system that made use of the YOLOv3 Neural Network, while the drowsiness detection system utilized the Convolutional Neural Network. Once the system detects an anomalous behaviour, this will create an alarm noise to notify the person monitoring the subject.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Assumption and Preconditions

The system outcome was expected to meet the following conditions:

1. The surveillance software would be able to classify movement and detect the specific anomalous behavior; and
2. The emergency notification system would be able to sound an alarm and notify the person assigned to monitor the subject.

Methods

In this section, we introduced the methods of our video-based real-time anomaly detection system followed by details on the Long-short Term Memory (LSTM) algorithm, You Only Look Once Version 3 (YOLOv3) algorithm, Convolutional Neural Network (CNN), and the data sets used for evaluation. The following methods were discussed for the four detection systems, namely fall detection, violence detection, weapon detection, and drowsiness detection systems.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

Long-short Term Memory (LSTM) Algorithm for Fall and Violence Detection

LSTMs according to Hochreiter, S. et al. (2020), are a recurrent neural network architecture. The memory extension that can be viewed as a gated cell is the LSTM's primary characteristic. By gated, it means that the cell, in accordance to the importance it assigns to the information, concludes whether or not to store or delete information. The importance assignment, which is also learned by the algorithm, takes place through weights. In short, it understands which information is important in the course of time.

LSTM architecture comprises of Input layer, hidden layer and output layer as shown in Figure 5. The input and output layers are totally connected to the hidden layers. An LSTM layer consists of blocks, where each block has three gates namely, input, output and forget gates, that are connected to each other and make decisions on each other.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

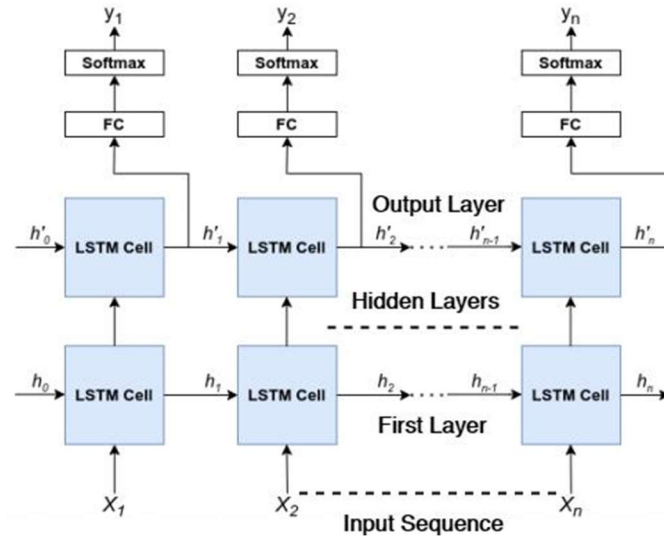


Figure 5. LSTM Neural Network Architecture

The figure above presents the architecture of an LSTM Neural Network in reference to the study of Hochreiter, S. et al. (2020).

Our motivation in using LSTM in this study was dependent on the fact that it allowed us to remember the data sequences. Additionally, we chose LSTM to identify abnormal behavior and precisely predict it from a person's ADL (Activities of Daily Life) in a sequential manner with minimal human intervention.

The LSTM input layer development requires input data to be 3-dimensions as training sample, time step, and features. ReLu activation function is added for this layer. We used the

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

dropout method based on the study of Hinton, G to prevent overfitting issues in LSTM architectures and to enhance the model's performance. The dropout being set to 20% as recommended, was applied between the two hidden layers and between the last hidden layer and the output layer in our proposed model.

The number of outputs that represented the different activities and anomalies was defined in the last layer. The output was viewed as a vector of integers which is transformed into a binary matrix. Prediction of anomaly was formulated as a multi classification problem which needs to produce 7 output values, one for each class, Softmax as activation function and categorical_crossentropy as the loss function.

You Only Look Once, Version 3 (YOLOv3) Algorithm for Weapon Detection

Kadlaskar (2021) discussed the Yolov3 Algorithm, also known as YOLOv3 (You Only Look Once, Version 3), as a real-time object recognition algorithm that recognizes specific items in real-time videos, cameras, and still photos. In order to create better predictions for a single image, region-based convolutional neural networks (R-CNN) require millions of

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

network evaluations, which can be time-consuming and difficult to optimize.

Yolov3 combined the feature extraction and object localization processes into a single, consistent block. YOLO (You Only Look Once) is the name of its single-stage architecture, which results in a very short time to come up with a conclusion. A single example was taken into consideration, and the bounding boxes and class probabilities for these boxes were predicted for an object using only the image data as input. Images were not scanned with a sliding window as is the case with other approaches; instead, the entire image was sent into a convolutional neural network (CNN), which predicted the output in a single pass. Figure 6 shows its architecture:

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

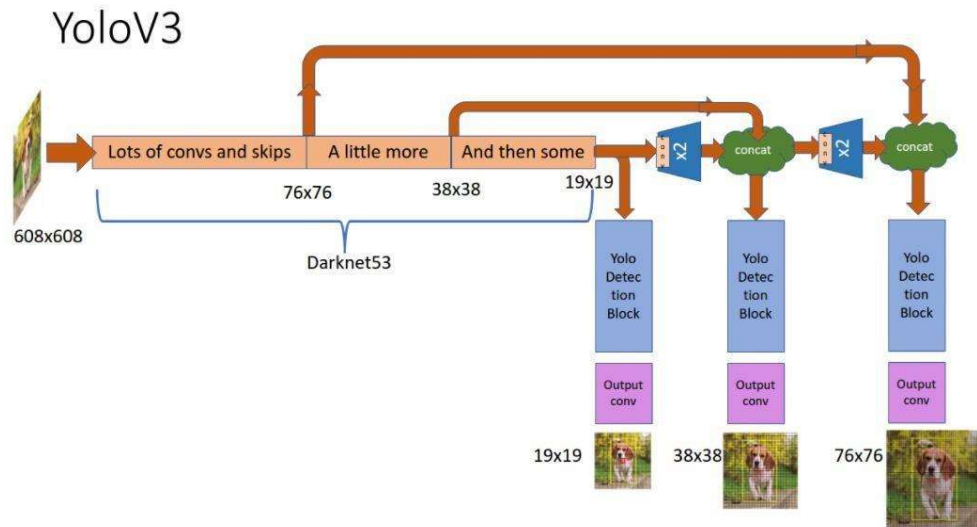


Figure 6. YOLOv3 Network Architecture

The figure above represents the architecture of YOLOv3 Network in reference to the study of Kadlaskar (2021).

Convolutional Neural Network for Drowsiness Detection

In this study, we used a model that was made with Keras using Convolutional Neural Networks (CNN). A CNN is a classification of deep neural networks which is perfectly suitable for image classification purposes. CNN comprises three layers namely input layer, an output layer and a hidden layer. These layers can have multiple layers and on these

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

layers, convolution operation is done using a filter that performs 2D matrix multiplication on the layer and filter.

The layers of the CNN model architecture includes:

- Convolutional layer; 32 nodes, kernel size 3
- Convolutional layer; 32 nodes, kernel size 3
- Convolutional layer; 64 nodes, kernel size 3
- Fully connected layer; 128 nodes

Also, the final layer was a fully connected layer with 2 nodes. We used a Relu activation function on the input and hidden layer while Softmax was used on the output layer. Its architecture as proposed by Singh, K. et al. (2019), is shown in Figure 7.

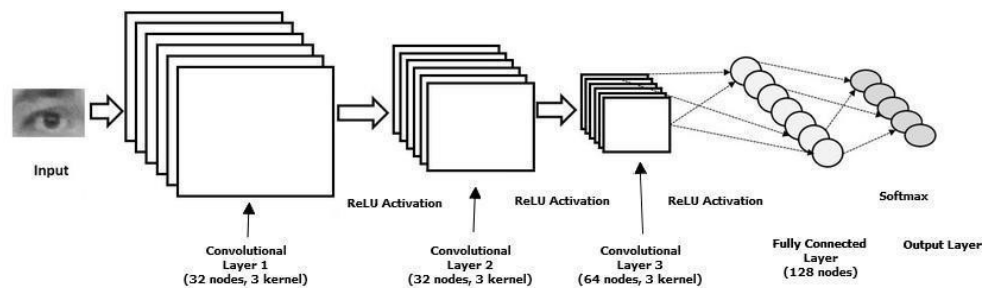


Figure 7. Convolutional Neural Network Model Architecture

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

The figure above describes the architecture of the Convolutional Neural Network Model in reference to the study of Singh, K. et al. (2019).

Optimizing using Adam Optimizer

In a published article of Change, Z. et al. (2019), Adam was an adaptive learning rate method derived from adaptive moment estimation because it uses both first and second moments of the gradient to update the learning rate. Moreover, Adam was a method that calculates individual learning rates for different parameters and incorporates the advantages of AdaGrad and RMSProp methods. Adam updated first the exponential moving averages of the mt gradient and squared the vt gradient. With ReLU activation function, it achieved best results with the values of 80.34% at top-4 and 86.32% at top-5, with learning rate value is 0.1 at 500 epochs during optimization. Compared to Adagrad, Adadelata, SGD, and RMSProp, Adam achieved the highest accuracy based on the results. However, some factors affected the accuracy results such as data used, good preprocessing results, good architecture used, good tuning parameters, etc.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Fall Detection System

The methods we used for our fall detection system were divided into four following parts as depicted in Figure 8:

1. Human Pose Estimation where the input video feedback from the camera was analyzed as a sequence of images and human pose key points were then extracted from the frames using the OpenPifPaf algorithm. OpenPifPaf algorithm as proposed by Kreiss, S. et al. (2019), is a method that uses a Part Intensity Field (PIF) to localize body parts and a Part Association Field (PAF) to associate body parts with each other to form full human poses. For each frame obtained from the camera, a list of key point sets was obtained with each key point set representing the position of joints of a single person with the use of this algorithm.

2. Subject Identification and Tracking was created between the key points derived from the first method. In this stage, a mapping was created between key point sets. The list of key point sets was obtained with each set being mapped to the key point set of the same person in the previous frame.

3. Feature Extraction introduced five features for fall detection. These were based on temporal as well as spatial

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

features and were independent of the pose detection algorithm.

4. We introduced a neural network called Long-short Term Memory (LSTM) and its architecture is depicted in Fig 8. The algorithm followed a step-by-step equation in solving the probability of a person's fall. If the probability of the "falling" activity was highest in the prediction, a fall would be detected for that frame.

Dataset

For our fall detection system we used the UP Fall Detection Dataset that was proposed by Martinez- Villaseñor, L. et al. (2019), which consists of videos from two cameras of 17 subjects performing 11 different activities with three trials for each activity. Table 4 shows the various activities in the data set and for how long the duration of each activity lasts. We grouped the activities 1-5 in one "Fall" class and kept the remaining classes of "No Fall" activities as they are.

Violence Detection System

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

The methods we used for our violence detection system were divided into three parts as depicted in Figure 9.

1. Preprocessing of the two input frames from the live video feedback was acquired using the surveillance camera. The two frames outputs from the bottom layer of the pre-trained model were then concatenated in the last channel and then fed into the additional CNN as depicted in Figure 9. Since the outputs from the bottom layer were regarded as the low-level features, the additional CNN was supposed to learn the local motion features as well as the appearance in variant features by comparing the two frames feature map. The two frames outputs from the top layer of the pre-trained network were also concatenated and fed into the other additional CNN to compare the high-level features of the two frames.

2. The outputs from the two additional CNNs were then concatenated and passed to a fully- connected layer and the LSTM cell to learn the global temporal features.

3. Finally, the outputs of the LSTM cell were classified by a fully-connected layer which contained two neurons that represent the two categories (fight and non-fight), respectively.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Dataset

For our violence detection system, we used the Hockey dataset proposed by Bermejo, E. et. al. (2011), which had 500 fighting clips and 500 non-fighting clips collected from the hockey games. Following the experiment proposed by Ding et. al., the dataset was further split into the following configuration: the 400 clips (including 200 fighting clips and 200 non-fighting clips) for testing, the 500 clips for training and the 100 clips for validation.

Weapon Detection System

The methods we used for our weapon detection system were divided into three following parts as depicted in Figure 10:

1. For object detection, the input video from the surveillance camera was first processed for its frames to be converted. It then preprocesses the video for the detection and localization of objects using the YOLOv3 algorithm which was then subjected to object identification.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

2. For the analysis phase, once the system had detected the frames from the live video from the surveillance camera, the weapon detection system classified whether the object was a weapon or not.

3. Lastly, in the action phase, once the system detected a weapon in the frames, it would sound an alert to notify the person monitoring the subject.

Dataset

YOLO, as defined by Narejo et. al., (2021), is a pre-trained object detector; a pre-trained model is essentially a model that has been trained on another dataset previously. When starting from scratch, training a model can take several weeks or even a month to complete the initial training phase. Before being trained, an object should have already been observed by the model, which means it should know how to classify each thing. It was necessary to train the network on the COCO and Imagenet data sets to generate the weights used in the pre-trained model. Consequently, it can only detect items that belong to the classes represented in the dataset used to train the network; in this system's case, we used a

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

pre-trained model that exclusively detects firearms in frames to achieve this result.

Drowsiness Detection System

The methods we used for our drowsiness detection system were as follows as depicted in Fig 5.4:

1. The researchers collected photographs as input through the use of a camera. We created an infinite loop that captured each frame of the webcam in order to gain access to it.

2. As grayscale images were required as input for the OpenCV object detection algorithm, researchers converted the image to grayscale before attempting to recognize the face in the image. In order to detect the items, the researchers did not require color information. The researchers used a haar cascade classifier to detect faces, and then carried out the detection, as previously stated. This function produced an array of detections with the x and y coordinates, as well as the height, which is the width of the object's boundary box.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

This allowed the researchers to go through the faces one by one and created boundary boxes around each one.

3. The researchers then used the ROI to identify the eyes and provided that information to the classifier. The same process that was used to detect faces was also utilized to detect the presence of eyes. The researchers began by configuring the cascade classifier for eyes. The researchers extracted only the data from the eye's picture from the entire image. This was accomplished by first extracting the bounding box of the eye, after which the researchers extracted the eye image from the frame and displayed it. This information was given into the CNN classifier, which then predicted whether the eyes are open or closed in the image. Additionally, the researchers removed the right eye from its socket.

4. The classifier categorized next whether the eyes are open or closed based on this information. When it came to determining ocular status, the researchers turned to CNN classifiers. It was necessary to do specific operations on the image before it can be fed into the model since the model needs the correct dimensions to begin with in order to function properly. To start, the researchers transformed the

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

color image to grayscale using Adobe Photoshop. Then, because the model was trained on photos with a resolution of 24*24 pixels, resized the image to 24*24 pixels. The data was then adjusted in order to improve convergence. The dimensions were enlarged so that they can be fed into the classifier. The model was loaded to anticipate the appearance of each eye. If the value is equal to 1, it indicates that the eyes are open; if the value is equal to 0, it indicates that the eyes are closed.

5. Finally, scores were generated to determine whether or not a person is sleepy. The score is essentially a numerical value that is used to determine how long a person has closed his eyes for. As a result, while both eyes are closed, researchers continued to increase the score, whereas when both eyes are open, researchers dropped the score. An upper limit was established, for example, if the score rises to larger than 15, it indicates that the person's eyelids have been closed for an extended amount of time. This is the point at which the system sounds the alarm.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Dataset

The researchers developed the data set that was utilized to train this model. The dataset was created using software that captures images of eyes from a webcam and stores them on our local computer's hard drive. We labeled them with the words 'Open' or 'Closed' to distinguish them from one another. A manual cleanup of the data was performed by deleting any undesired photos that were not required for the model's construction. The data set contains approximately 7000 photographs of people's eyes taken under a variety of lighting situations. In this state, the model can be used to determine whether or not a person's eye is open or closed.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Components and Design

Software Architecture

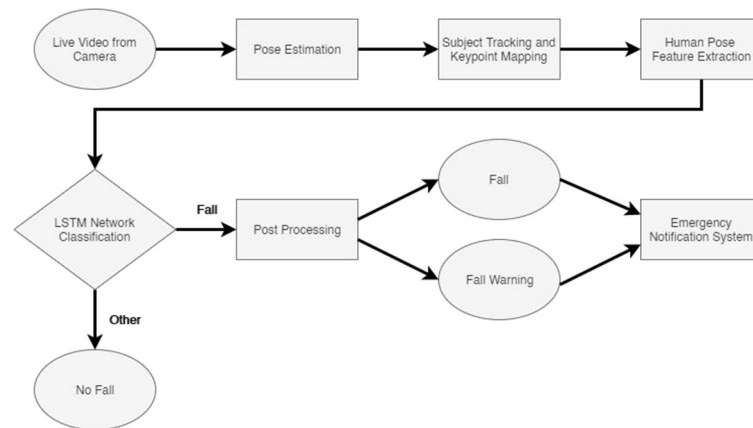


Figure 8. Architectural Design of Fall Detection System

Figure 8 shows the architectural design of our Fall Detection System. As depicted in the figure above, a live video feedback from the cameras was used. The system then extracted poses from the video using OpenPifPaf algorithm which assigns human pose key points on the body of the subject which represents the joints of the body. After the key points are identified, this was used to identify and track the subject. It then extracted features from these designated key points which then fed them to the underlying LSTM network to perform prediction whether a fall was detected or not.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

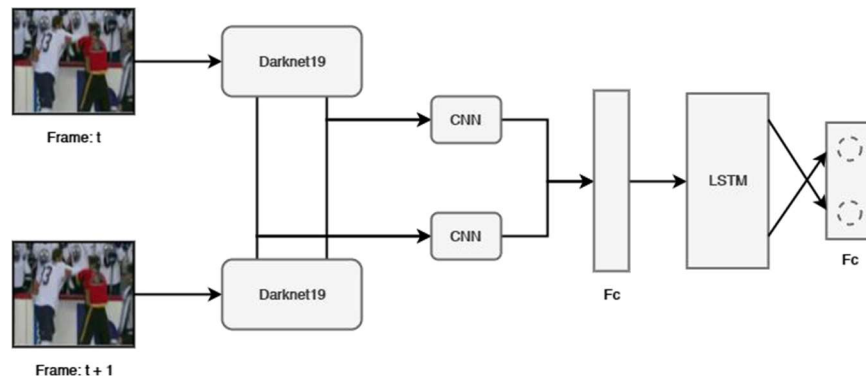


Figure 9. Architectural Design of Violence Detection System

Figure 9 depicts the architectural design of our Violence Detection System. It started with the pre-processing of the input frames from the camera and went through the pre-trained model. The outputs were then concatenated and passed through a fully-connected later LSTM cell for learning global temporal features. Finally, the outputs were classified whether the frame detected violence from the video or not.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

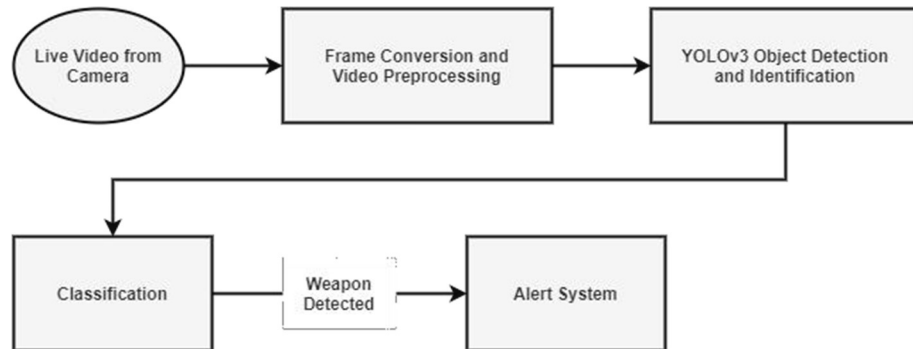


Figure 10. Architectural Design of Weapon Detection System

Figure 10 shows our architectural design for the Weapon Detection System. Initially, for object detection the input video was processed to be converted. It then pre-processed the video for the detection and location of objects using the YOLOv3 algorithm. Then the system analyzed the input once it has detected the frames which was subjected for classification whether the object is a weapon or not. Lastly, a sound alarm will blare once the system detected a suspected weapon

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

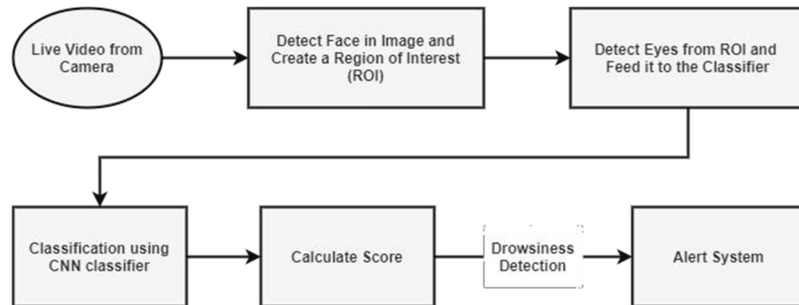


Figure 11. Architectural Design of
 Drowsiness Detection System.

Figure 11 shows the architectural design of our Drowsiness Detection System. This used live video from cameras as input and then processed the frames for the detection of faces, which then created a Region of Interest (ROI) on where the face is located in the frame. Once the system assigns the ROI, it processed the frame again to detect the eyes as output from the ROI. This output was then fed to the CNN for classification by solving the score of how long the eyes of the subject were closed. If the eyes were closed for a long period of time, the system sounds an alarm to alert the person in-charge of monitoring the subject.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Procedural and Object-Oriented Design Procedural Flow

Flow diagram of the proposed Fall Detection System as shown in Figure 12.

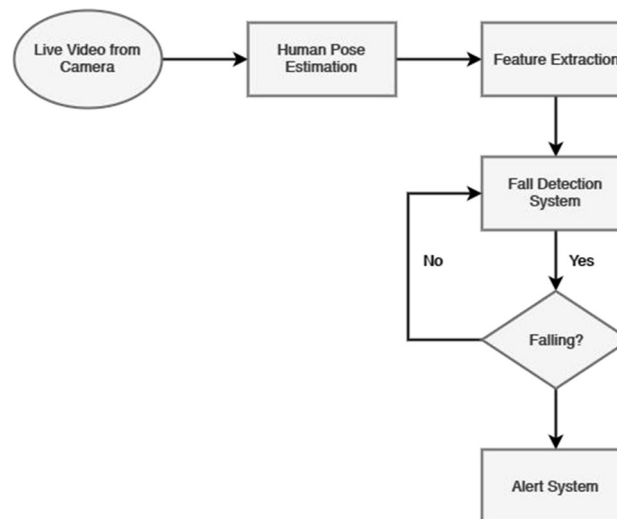


Figure 12. Fall Detection System Flowchart Diagram.

Figure 12 shows the flowchart of the Fall Detection System. The system made use of live video feedback from the camera which then processed for human pose estimation using OpenPifpaf algorithm for the assignment of key points on the human body that represents the joints. It then extracted features from the processed key points, and these features were then fed to the underlying LSTM network for classification of the frames. If the system detects a fall

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

incident it sounds the alert system, and continue to detect an anomaly until investigated.

Flow diagram of the proposed Violence Detection System as shown in Figure 13.

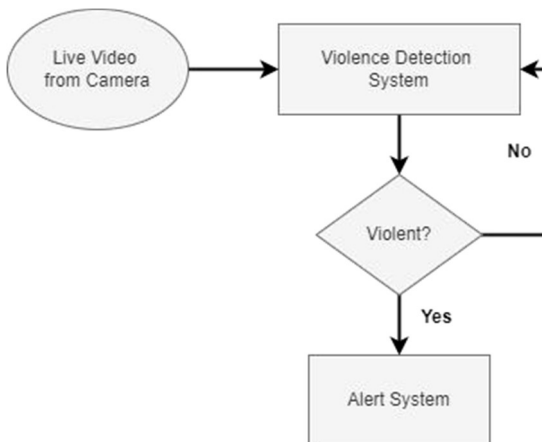


Figure 13. Violence Detection System Flowchart Diagram

Figure 13 shows the flowchart of the Violence Detection System. Using live video from cameras as input and it was then processed by the system for classification. If the system detects violence in the frame it activates the alert system, otherwise it continues to detect possible anomalies.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

Flow diagram of the proposed Weapon Detection System as shown in Figure 14.

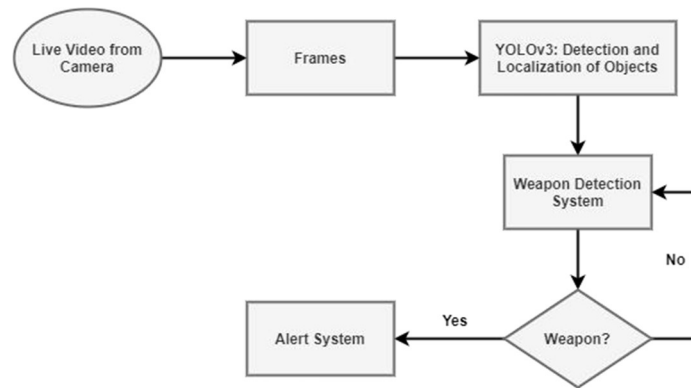


Figure 14. Weapon Detection System Flowchart Diagram.

Figure 14 shows the flowchart of the Weapon Detection System. It made use of live video feedback from the camera. The frames processed and went through the YOLOv3 algorithm for the detection and location of the objects. The output then fed these to the detection system for classification. If the system detects a suspected weapon on the frame, the alert system activates. Otherwise, it continues to detect anomalies.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

Flow diagram of the proposed Drowsiness Detection System as shown in Figure 15.

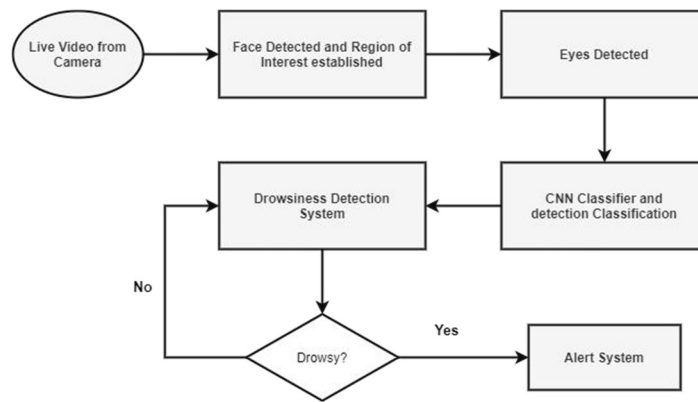


Figure 15. Drowsiness Detection System Flowchart Diagram

Figure 15 shows the flowchart of the Drowsiness Detection System. This made use of the live video from the camera as input and was processed for the detection of faces and assignment of region of interest (ROI). The system then analyzed the frame by detecting the eyes in the established ROI. The output was fed to the CNN for classification. If the system detects a drowsy state from the subject, this activates the alarm system. Otherwise, it continues to detect anomalies.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

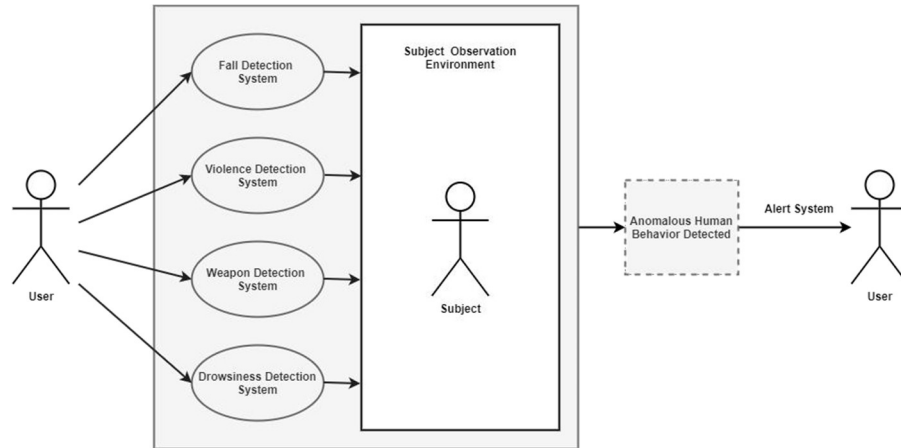


Figure 16. Use Case Diagram of the proposed system

Figure 16 shows the used case diagram of the proposed anomalous human behavior detection system. The user has access to the system to detect anomalous human behaviors, the Fall Detection System for monitoring fall accidents real time, the Violence Detection System for detecting anomalous violent human actions, the Weapon Detection System for detecting dangerous weapons, and the Drowsiness Detection System for monitoring drowsiness.

Each of these systems is connected to the surveillance device that the user used to monitor the subject. Each of these systems has its own alert system that activates whenever it detects an abnormal event within the threshold of the surveillance system. This alert system is set off and notifies

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

the user of the event that has happened to the subject.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

Process Design (DFD)

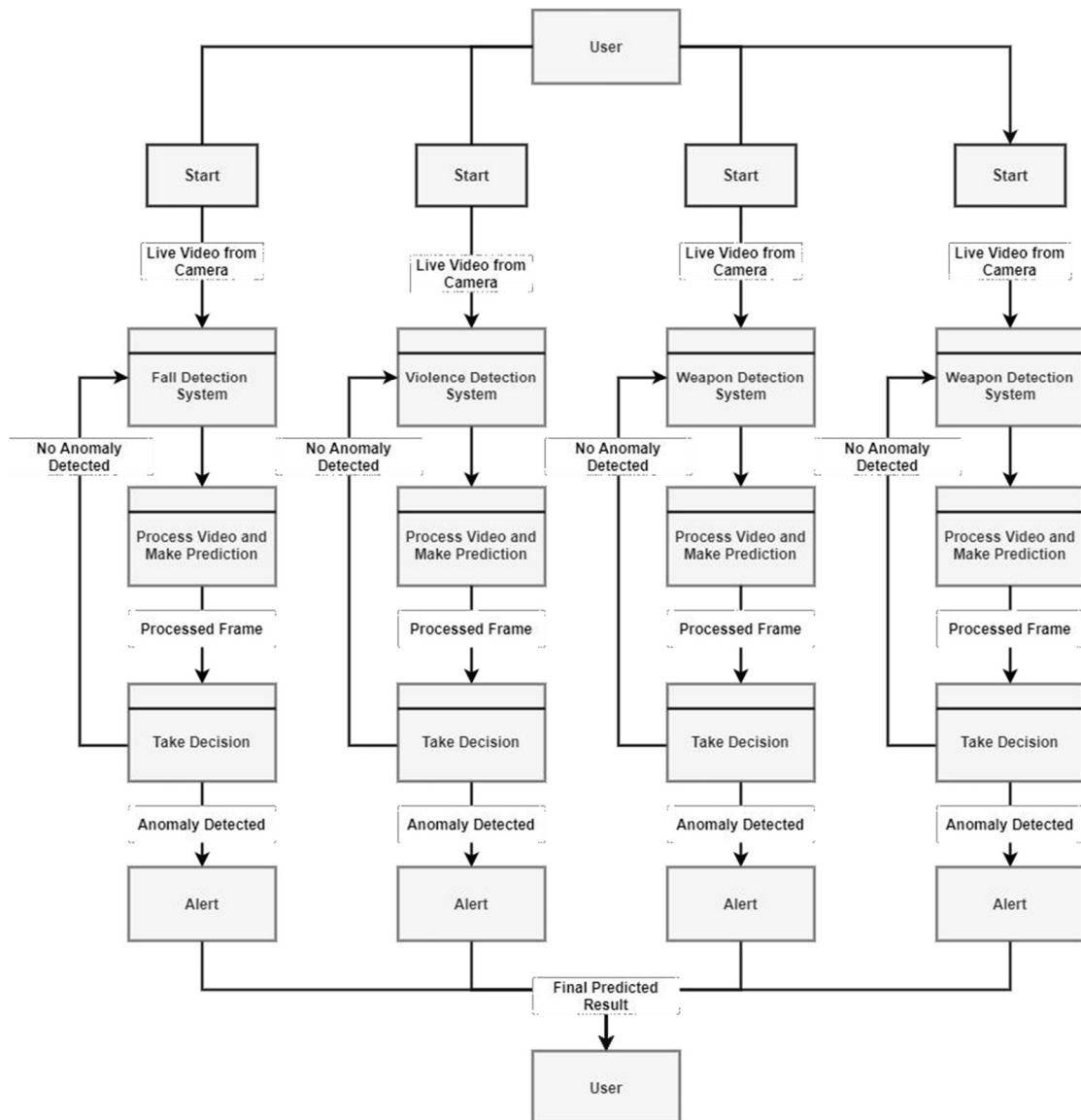


Figure 17. Data Flow Diagram of the proposed anomalous behavior detection system.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Figure 17 shows the data flow diagram of the whole detection system. All systems made use of live video feedback from the camera as input, once the system started each system processed the input videos with their designated models to cater the decision making of each. Once each system has processed the video, it can then make a prediction followed by its decision. If the system detects an anomaly in the frame, then it would activate the alert system, otherwise it would continue to monitor for anomalies. The final result of the prediction would be to notify the user as the alert system sounds an alarm.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

System Development Life Cycle (SLDC)

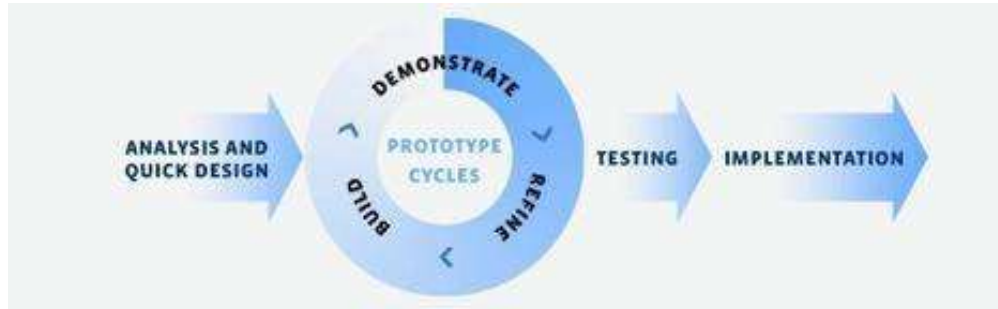


Figure 18. Rapid Application Development Methodology

In this study, the researchers used the Rapid Application Development (RAD) Methodology. This was divided into four phases: (1) Analysis and Quick Design, (2) Prototype Cycles, (3) Testing, and (4) Implementation.

For the first phase, which is Analysis and Quick Design, the researchers looked into the current problems, identified what information and data were needed to meet the requirements and were available for use.

In the Prototype Cycles, wireframes and mock-up designs for the UI/UX of the software were created by the researchers.

For the Testing phase, the researchers created the software, making use of the designs established in the previous cycle and tested the system's functionality and

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

reliability. This was the longest of all the phases to maximize user satisfaction in order to build a working model until the researchers started with an actual prototype.

For the final phase, Implementation was applied through testing and data conversion. Through user testing, the researchers were able to collect insights and look for bugs that can be resolved by the developers of the team.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

CHAPTER 4 RESULTS AND DISCUSSION

System Implementation

The System Implementation described how the system was deployed, installed and transitioned into an operational system. The plan contained an overview of the system, a brief description of the major tasks involved in the implementation, the hours needed to implement and the maintenance operations after the system was fully operationalized.

The following were the four systems presented; Fall Detection, Violence Detection, Weapon Detection, and the Drowsiness Detection. Explanation on how each was implemented as one would be discussed here.

Fall Detection System

The PyTorch library was used to implement the network architecture in this section of the system, which was developed by the researchers. The LSTM architecture depicted in Figure 5 was composed of two LSTM layers, each of which has 48 hidden states of dimension 256 and is made-up of two LSTM layers. In the following step, the output of the LSTM

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

layer was transmitted via a fully-connected layer, which reduced the hidden representation from 256 to 7 nodes in size. This was followed by the softmax layer, which calculated the probability of each activity taking place. The researchers utilized a weighted Cross Entropy loss function, with the terms in the loss function for each label being weighted according to how rare they were. It is supposed that a label must have a relative frequency of f , and the loss that results from using that label has a weight of $1/f$. This ensured that the resulting network was trained in a balanced manner, with equal priority being given to the classes "Fall" and "No Fall," regardless of the amount of videos present in the data set during the training process.

The use of mini-batches of size 64 allows the researchers to accelerate the training process. Every one of the models was trained with the Adam optimizer given by PyTorch, which has a learning rate of 0.0001 for all of them. In order to make the model more regular, an L2 weight decay of 0.01 was applied to all of the weight parameters. The model was further regularized by applying drop-out with a drop-probability of 0.1 on the LSTM layer, with the probability of the drop-out

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

being 0.1. All of the hyper-parameters were discovered by the use of a grid search across the hyper-parameter space.

Accuracy Result on the use of LSTM on Fall Detection System

Methods/ M etrics	LR i-ii	LSTM i- ii	LSTM i- iii	LSTM i- iv	LSTM i-v
Accuracy	92.61	97.85	97.70	98.28	98.22
Precisio n	66.06	83.24	82.90	88.04	89.76
Recall	54.45	80.54	90.01	91.48	95.62
F1 Score	60.65	86.63	85.61	89.61	91.56

Table 3. Results with the Roman Literals Defining the
Features Used for Classification

ID	Description	Dur. (s)			
1	Falling forward using hands	10	6	Walking	60
2	Falling forward using knees	10	7	Standing	60
3	Falling backwards	10	8	Sitting	60
4	Falling sideways	10	9	Picking up an object	10
5	Falling sitting in empty chair	10	10	Jumping	30
			11	Laying	60

Each frame of each video within the UP Fall Detection Dataset was appointed a label from the activities shown in Table 3. thus, the multi-class labels was converted to binary "Fall" vs. "No Fall" labels as shown in Table 4, and the following metrics of evaluation were defined as shown in Table 3:

- Accuracy. Precisely classified frames to all classified frames ration.
- Precision. Accurately classified "Fall" frames to all frames classified as "Fall ratio.
- Recall. Rightly classified "Fall" frames to all frames with ground-truth label of "Fall ratio.
- F1 Score. Precision and recall harmonic mean.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Evaluation

The results using various models on the UP Fall detection dataset with various numbers of features was shown in Table 3. 5-fold cross-validation was performed to obtain results and using all five features with the LSTM network, best results were obtained. Logistic regression classifier was applied as reference and clearly, the LSTM improves all metrics. Moreover, increasing the number of features applied increases the overall performance as can be seen. All five features were used to obtain the best scores of 92.56% F1 Score.

Violence Detection System

In order to create a time series model for activity prediction, we used an LSTM network for the proposed prediction architecture, as shown in Figure 2. We used the LSTM model for violence prediction and then used Adam optimizer during training to optimize the LSTM network.

The proposed architecture in Figure 6.1 for the researcher's Violence Detection System shows the steps that was followed. The acquisition of data was carried out using

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

data that were already available on the internet. The data were emanated from monitoring activities from various sources of violent actions and further labeled as the sequences of activities. These labels were then fed to the embedding layer that was to be processed by LSTM as a sequence modeling module, then the model was optimized with Adam optimization method to help get a good detection.

Accuracy Result on the use of LSTM on Violence Detection System

The classified result per frame in this work was outputted by the proposed model however the preceding research estimates the accuracy at the video-level. The frame-level results were obtained and processed to be able to compare with the past work through the following strategy: If and only if the number of the continuous signals of such a category is bigger than a definite threshold, then the video is classified to a definite category. By scanning the threshold from 0 to the length of the video and perceiving which threshold yields the best accuracy in the validation set, a certain threshold can be acquired. The small one would

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

be chosen If there a are multiple thresholds that are able to yield the same accuracy.

Method	Accuracy
STIP(HOG)+HIK with 1000 vocabulary	84.25%
STIP(HOF)+HIK with 1000 vocabulary	78.00%
STIP(HOG+HOF)+HIK with 1000 vocabulary	78.50%
MOSIFT+HIK with 1000 vocabulary	90.90%
Conv3D	91.00%
Darknet19 + Residual Layers + LSTM	98.50%

Table 5. The comparison between the previous methods and the proposed method

500 fighting clips and 500 non-fighting clips were gathered from the hockey games composed the Hockey data set proposed by Bermejo, E. et. al. The dataset, following the proposed experiment by Ding et. al., was further split into the following configuration: the 400 clips for testing which contained 200 fighting and non- fighting clips, the 500 clips for training and the 100 clips for validation. As one can

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

see, the result given in Table 5 shows that the proposed method in this work outperforms other state-of-the-art methods.

Weapon Detection System

In this system, the researchers made use of live input from the surveillance camera installed in the room where the subject iwa located. This data was then inputted to the system for prediction. The video underwent three phases, the object detection phase, the analysis phase, and the action phase. The input video data got through the object detection phase where it was converted into frames. The frames were then pre-processed for the object detection and identification using the pre-trained model YOLOv3. After the system has detected the frames on the video, the system then classified them whether the object was a weapon or not in the analysis phase. The system then entered the final phase where it takes action to activate the alarm if it detects a weapon on the frames, which then alerts the user of the detected anomaly.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Accuracy Result on the use of YOLOv3 on Weapon Detection System

Example of Image classification includes the class of one object in a picture. The detection kernel's shape is calculated through $1 \times 1 \times (bb \times (4 + 1 + nc))$ wherein bb signifies the number of bounding boxes, "4" is for the 4 bounding box coordinate positions and 1 is object confidence, and nc is the number of classes. The input image down sampling is for three scale predictions and is computed by strides 32, 16, and 8. As presented below in Equation 2, the loss function is composed of three sections, location error (L_{box}), confidence error (L_{cls}), and classification error (L_{obj}).

$$\text{Loss} = L_{\text{box}} + L_{\text{cls}} + L_{\text{obj}}.$$

Equation 6

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

S. no	Models	Accuracy
1	Traditional CNN	95%
2	YOLO V2	96.76%
3	YOLO V3	98.89%

Table 6. Experimental Result for trained
deep learning models

Small object detection is often the problem of YOLO v2 because as the layers down sampled the input, the fine-grained features get lost. In short, to capture low-level features YOLO v2 applies an identity mapping, concatenating feature maps from a previous layer. However, most state-of-the-art algorithms have some of the influential essentials that are enclosed in them which YOLO v2's architecture was lacking. On the contrary, YOLO v3 integrates all of these. We retrained both YOLO V2 and YOLO V3, and as an alternative, comparative analysis of the models also conducted with traditional CNN which was trained from the very scratch with null weights. Table 5 shows the summary of the gathered results.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Drowsiness Detection System

For this system, the researchers utilized OpenCV for gathering the images from a webcam or surveillance camera and fed them into a deep learning model which classified whether the person's eyes were 'Open' or 'Closed'. The system made use of live video feedback from a surveillance camera installed inside the observation environment of the subject.

With the use of OpenCV, the system took images from the live feedback as input. This captured and read each frame and stored those images in a frame variable. In order for the system to detect the faces in the image, the system first needed to convert the image into grayscale using the OpenCV algorithm as its object and detection takes gray images as its input. This removed the color information of the image as it is not needed to detect the objects. In order for the system to detect the faces, it used a haar cascade classifier to detect faces. This allowed the system to perform the detection process and returns an array of detections with x, y coordinates, and height, and width of the boundary box of the object, which in turn allowed the system to iterate over the faces and draw boundary boxes for each faces detected on

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

the frame.

After the system has drawn the bounding box around the face, the next step was to detect the eyes from the region of interest and fed it to the classifier. The same procedure of using a cascade classifier to detect faces were then used to detect the left and right eyes of the subject.

As the system detects the eyes, this would be fed into the convolutional classifier which in turn would predict whether the eyes are open or closed. Whenever the system detects a drowsy state from the subject, it would sound an alarm to alert the user, otherwise it would remain normal.

Accuracy Result on the use of CNN on Drowsiness Detection System

Convolution Neural Networks (CNN) was used to classify the result detecting the fatigue status of the subject. The analysis of experimental results was done based on the accuracy level and error rating in detecting the subject state. The accuracy level of Convolution Neural Network (CNN) was compared with the conventional classifiers of SVM classifier and KNN classifier and its inferred classifiers as

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

an initial analysis. The comparative analysis of CNN with other conventional classifiers regarding the accuracy percentage was illustrated in Table 6. CNN classifies images accurately through the use of multiple convolution operations and it has scalable features for massive datasets.

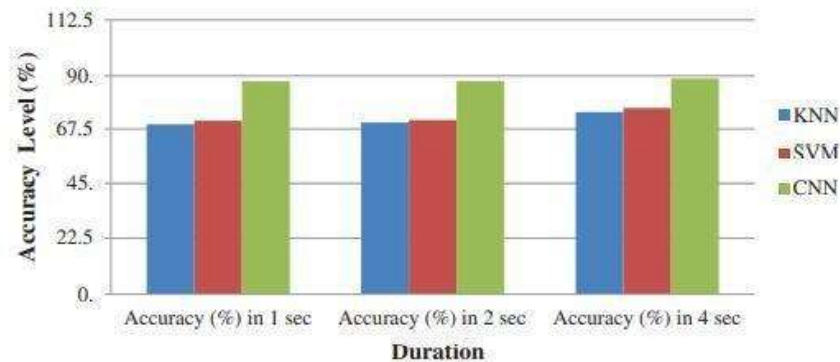


Table 7. Comparison of classifier types
in terms of Accuracy level

It is verified from Table 7 that the multi-layer CNN is more accurate in predicting the state of the driver and successfully classifying the multi-layer state of the driver. By increasing the processing duration, the classifier's level of accuracy improves. The eminent instance in this driver drowsiness system was that in order to distinguish the distraction of the driver, it must have a minimum duration. These models were capable of extracting features that

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

identify the drowsy state of the driver by using two-levels; however, if three or more levels were selected, decrease in accuracy and model overfitting would be the result.

System Inputs and Outputs

In this study, AI/ML for Anomalous Surveillance: A Detection and Alert System, requires data from live feedback from surveillance cameras installed in the observation environment where the elderly subject was located. The data were used to predict the anomalous behavior happening to the subject with the use of the said surveillance cameras. The system made use of the LSTM algorithm for prediction on the fall detection and violence detection, which was further optimized with the Adam optimizer, YOLOv3 algorithm for prediction on the weapon detection, and Convolutional Neural Network for the drowsiness detection.

The following figures show the user interface and visual representations.

Figure 19 shows the GUI when starting the system. We listed four systems, 'Weapon Detection', 'Violence Detection', 'Drowsiness Detection', and 'Falling Detection'.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

There was also a button 'Change System', where its function was to exit the current running system so that the user can switch to a different system. The system was displayed on the white area of the screen.

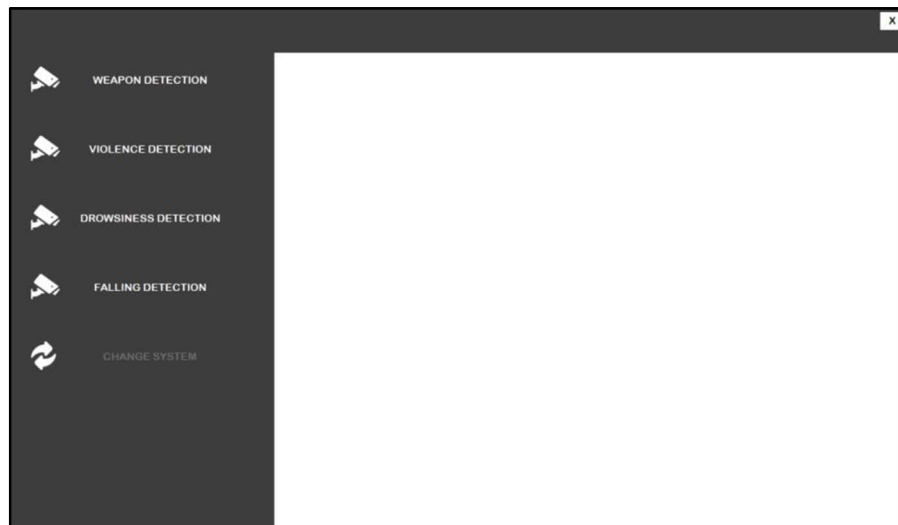


Figure 19. System GUI

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Figure 20 below shows how our Fall Detection System works. We implemented a border on all sides of the detection system for a clear view if the system detects an anomalous activity, which in this case falling. In normal scenarios the border would remain green, which indicates that it is normal, or "No Fall".

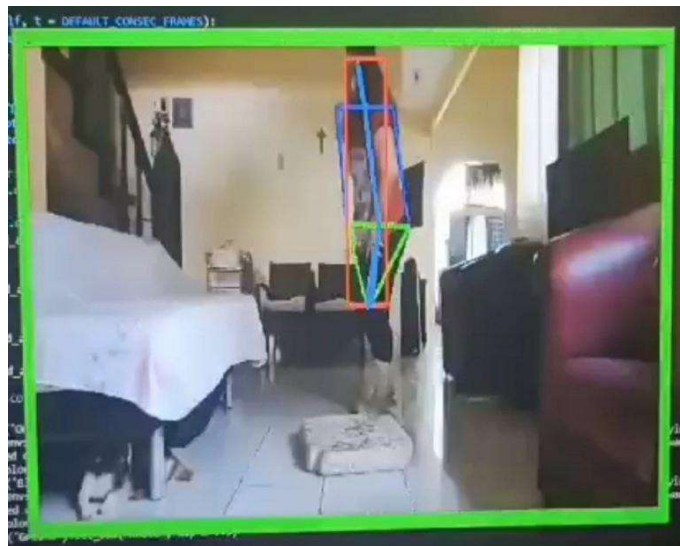


Figure 20. "No Fall" scenario for the Fall Detection System

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Figure 21 shows if our Fall Detection System detects a “Fall” activity on the video. In this scenario, the border would turn red, which would indicate that the system has detected an anomalous behavior, which will then sound an alarm to notify the user.



Figure 21. “Fall” scenario for the Fall Detection System

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Figure 22 shows how our Violence Detection System works. We implemented a border on all sides of the detection system for a clear view if the system detects an anomalous activity, which is violence in this case. In normal scenarios the border would remain green, indicating that it is normal, or "Non-violence".



Figure 22. "Non-violence" scenario for the
Violence Detection System

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Figure 23 shows if our Violence Detection System detects a "Violence" activity on the video. In this scenario, the border will turn red, which would indicate that the system has detected an anomalous behavior, which would then sound an alarm to notify the user.



Figure 23. "Violence" scenario for the
Violence Detection System

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

Figure 24 shows how our Weapon Detection System works. We implemented a border on all sides of the detection system for a clear view if the system detects an anomalous activity, which in this case is the presence of weapons. In normal scenarios the border would remain green, which indicates that it is normal, or "No Weapon".



Figure 24. "No Weapon" scenario for the
Violence Detection System

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Figure 25 shows if our Weapon Detection System detects a "Weapon" on the video. In this scenario, the border would turn red, which indicates that the system has detected an anomaly, causing the alarm to sound and notify the user.



Figure 25. "Weapon" scenario for the
Violence Detection System

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Figure 26 shows how our Drowsiness Detection System works. We implemented a border on all sides of the detection system for a clear view if the system detects an anomalous activity, which in this case the presence of weapons. In normal scenarios the border will remain green, which indicates that it is normal, or “Open” eyes.



Figure 26. “Open” eyes scenario for the
Drowsiness Detection System

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Figure 27 below shows if the Weapon Detection System detects "Close" eyes of the subject in the video. In this scenario, the border would turn red, which indicates that the system has detected an anomaly, which will then sound an alarm to notify the user.



Figure 27. "Close" eyes scenario for the
Drowsiness Detection System

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Evaluation Result

The system was presented to forty (40) surveyees to determine the quality of the proposed system. The 38 said surveyees were caregivers or those who take care of the elderly. As suggested by the panel, the researchers were able to get two IT professionals as surveyees to rate the system in overall set. The researchers constructed a software quality evaluation form which covered seven (7) criteria.

For this purpose, the criteria was designed in reference to ISO 9126 and adjusted its standard.

The criteria for evaluation involved the following: the Functional Suitability which includes the Functional completeness, Functional correctness, and Functional appropriateness; the Performance Efficiency which includes Time behavior, Resource utilization, and Capacity; the Compatibility which includes Co-existence, Interoperability; the Usability which includes Interoperability, Appropriateness recognizability, Learnability, Operability, User error protection, User interface aesthetics, Accessibility.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Additional criterion was also provided whose indicators were as follow: the Reliability which includes Availability, Fault Tolerance, Recoverability, Confidentiality, Integrity, Non-repudiation, Accountability, Authenticity, Modularity, Reusability; the Maintainability includes Analyzability and Modifiability; the Portability includes Adaptability, Installability, and Replaceability.

The surveyee rated each criterion accordingly and compiled the results into area mean. The surveyee also provides their recommendation for the enhancement of the system. The IT professionals that were asked mentioned that the highly recommended solution was to have high performance resources. Similar to the suggestion of one of the panelists, one mentioned using a different optimization tool, however both had the similar suggestion of incorporating the various detections in one video capture, which was why the use of multi-cams to have a surveillance-like output was utilized.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Surveyees' Evaluation Result

CRITERIA	MEAN	DESCRIPTI O N	RANK
Functional Sustainability	4.391666666	Very Good	3
Performance Efficiency	4.166666667	Very Good	6
Compatibility	4.25	Very Good	7
Usability	4.328571429	Very Good	5
Reliability	4.340625	Very Good	4
Maintainability	4.4375	Very Good	1
Portability	4.4	Very Good	2
Overall mean	4.330718537	Very Good	
Median	4.5		
Mode	5		

Table 8. Table of Results from Survey Evaluation

Ranks were used to determine the specific order of the different evaluative criteria-based on the ISO 9126 standards. The researchers were able to gather a total of 40 respondents. The criteria Functional Stability placed third,

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

while Reliability placed fourth. In fifth place, it was the Usability criteria. As for the last two ranks, Performance Efficiency placed 6th and Compatibility placed 7th.

The three highest ranks, which were Maintainability, Portability, and Functional Stability, had a mean of 4.437, 4.4, and 4.392, in order, respectively.

In summary, all of the ranks have a general description of "Very Good." The overall mean garnered was 4.330718537, which is classified as "Very Good." The results also showed a median of 4.5 and mode of 5.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

CHAPTER 5 SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Chapter 5 is composed of four parts: (1) Summary of the Proposed System and Research Design, (2) Summary of the Findings, (3) Conclusions, and (4) Recommendations.

Part One, Summary of the Proposed System and Research Design, presents a brief description about its system design.

Part Two, Summary of the Findings, provides the discussion of comprehensive results of the study.

Part Three, Conclusions, discusses the conclusion of findings of the proposed system.

Part Four, Recommendations, expounds possible improvements could be made by future researchers on the system.

Summary of the Proposed System and Research Design

The researchers proposed a system entitled AI/ML for Anomalous Surveillance: A Detection and Alert System which provides a system that detects anomalous behaviors mainly, fall, violence, detecting weapons, and drowsiness detection

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

for monitoring elderly people. The said system can be easily implemented in existing surveillance systems and web cameras to provide caregivers additional security measures in monitoring the elderlies. The programming language used was Python. The Long-short Term Memory algorithm was used to detect falling in the Fall Detection System and for detecting violence in the Violence Detection System, YOLOv3 algorithm for detecting weapons in the Weapon Detection System, and Convolutional Neural Network for detecting drowsiness in the Drowsiness Detection System.

The system made use of live video feedback from the installed surveillance camera on the observation environment where the elderly are located. This would provide a real time surveillance on the subject which will detect anomalous activities such as falling, violence, detecting dangerous weapons, and monitoring drowsiness. Since the system has a built-in alarm system which would sound if the system detects an anomaly, this may help caregivers to respond on time whenever these accidents happen.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Summary of Findings

This study entitled, AI/ML for Anomalous Surveillance: A Detection and Alert System, aimed to develop a system that can predict and detect anomalous activities mainly, falling, violence, detection of weapons, and monitoring drowsiness on live video feedback from surveillance cameras installed in the observation environment where the elderly is located. The system aimed to predict these anomalies and can sound an alarm to alert caregivers of the anomalous behavior detected.

To check the capability of the system, the researchers selected 38 respondents to test and evaluate the system. After they were presented with the working system, the researchers asked their permission to answer a survey questionnaire via Google Forms. The results of the said survey were automatically accounted for in the researcher's email account. The overall result was recorded and tabulated using Google Sheets. To determine the results of the survey, the researchers computed the Mean, Median, Mode, and Standard Deviation of each criterion given in the questionnaire.

[illegible]

Figure 28 above shows the results of the survey. The survey contains the answers of the 40 respondents which is in accordance with ISO 9126. The mean, median, and mode were calculated.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Conclusion

The researchers were able to achieve the following objectives set for the study after the evaluation and feedback.

1. The researchers were able to design a deep learning model for detecting specific human behavior. The models that were developed were Long-short Term Memory (LSTM) for detecting falling and violence, and Convolutional Neural Network (CNN) for detecting weapon and drowsiness.

2. The researchers were able to develop a desktop software integrated with the proposed models for anomalous human behavior detection. The user can use the software to autonomously detect elderly patients falling, committing violence, weapons, or drowsiness inside homes, hospitals, or retirement homes.

3. The researchers were able to successfully integrate an emergency alert system to the video surveillance desktop

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

software. The built-in alarm system sounded when it detected a certain anomaly.

4. The system was implemented and considered a success in terms of detecting anomalous human behavior such as falling, violence, carrying of weapons, and drowsiness using live video stream from the surveillance camera installed where an elderly subject is located. A user testing and evaluation using ISO-standard Usability Assessment Tool based on system functionality, performance, usability, reliability, and maintainability was performed and has garnered an excellent score.

Recommendations

The following are the recommendation made for future researchers in improving the efficiency and accuracy of the proposed system:

1. Higher computing power with a very good video processor, and multiple high resolution surveillance cameras is extremely recommended so that the system will flawlessly operate.

2. Future researchers should enhance the training of

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

the AI/ML models using a bigger dataset, especially the violence detection algorithm, so that the proposed system will be able to detect anomalous human behaviours more accurately.

3. Future addition of a zooming feature in the Drowsiness Detection System to track the eyes of the subject and detect drowsiness more accurately.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

References

Advani, V. (2021). What is Artificial Intelligence?. How does AI work, Types and Future of it?.

<https://www.mygreatlearning.com/blog/what-is-artificial-intelligence/>

Ahmed, M., Ramzan, M., Khan, H. U., Iqbal, S., Khan, M.A. et al. (2021). Real-Time Violent Action Recognition Using Key Frames Extraction and Deep Learning. CMC-Computers, Materials & Continua, 69(2), 2217-2230.

<https://www.techscience.com/cmc/v69n2/43878>

Akash Kumar. K. S., Akshita. B. P., Arjun. M., & Dr. Geetha N. (2021) Weapon Detection in Surveillance System, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 10. <https://www.ijert.org/weapon-detection-in-surveillance-system>

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

Alsever , S., Karapanagos, G., Diaz , C., Mould, G.,
 Raynes,

P. (2017). AWeD (Automatic Weapons Detection).

<https://digital.wpi.edu/downloads/5712m844v?locale=es>

Azzopardi, G., Ellul, J., & Wang, X. (2020) Elderly Fall
 Detection Systems: A Literature Survey. Front. Robot. AI
 7:71. <https://doi.org/10.3389/frobt.2020.00071>

Baba, M., Gui, V., Cernazanu, C., Pescaru, D. (2019). A
 Sensor Network Approach for Violence Detection in Smart
 Cities Using Deep Learning. [https://www.mdpi.com/1424-
 8220/19/7/1676](https://www.mdpi.com/1424-8220/19/7/1676)

Becker, C., & Chiari, L., Klenk, J., & Palmerini, L.
 (2020). Accelerometer-Based Fall Detection Using Machine
 Learning: Training and Testing on Real-World Falls. Sensors
 (Basel, Switzerland), 20(22), 6479.
<https://www.mdpi.com/1424-8220/20/22/6479>

Bhatti, M., Khan, M., Aslam, M., & Fiaz, M. (2021). Weapon

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Detection in Real-Time CCTV Videos using Deep Learning.

IEEE Access. PP. 1-1. 10.1109/ACCESS.2021.3059170.

Retrieved from

<https://www.researchgate.net/publication/349281117> Weapon D
 etection in Real-Time CCTV Videos using Deep Learning

Brown, S. (2021). Machine learning, explained. Retrieved
 from [https://mitsloan.mit.edu/ideas-made-to-matter/machine-
 learning-explained](https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained)

Burns, E. (2021). Machine Learning. Retrieved from
<https://www.techtarget.com/searchenterpriseai/definition/machine-learning-ML>

Chaudhuri, S., Oudejans, D., Thompson, H. J., & Demiris, G.
 (2015). Real-World Accuracy and Use of a Wearable Fall
 Detection Device by Older Adults. Journal of the American
 Geriatrics Society, 63(11), 2415-2416.
<https://doi.org/10.1111/jgs.13804>

Chelli, A. & Pätzold, M. (2019). A Machine Learning

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

Approach for Fall Detection and Daily Living Activity
 Recognition.

<https://www.researchgate.net/publication/331942521> A Machine Learning Approach for Fall Detection and Daily Living Activity Recognition

Chen, H., Lee, S., Rao, R., Slamani,, M., Varshney, P.,.

(2005). Imaging for concealed weapon detection

- A tutorial overview of development in imaging sensors and processing.

<https://www.researchgate.net/publication/3321594> Imaging for concealed weapon detection -

A tutorial overview of development in imaging sensors and processing

Chen, P.Y., Chiu, H.T., Chiu, H.Y. (2015) Daytime sleepiness is independently associated with falls in older adults with dementia. Geriatr Gerontol Int. 2016 Jul;16(7):850-5. doi: 10.1111/ggi.12567. Epub 2015 Aug 5. PMID:26245804. <https://doi.org/10.1111/ggi.12567>

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Choudhary, P. (2017). Introduction to Anomaly Detection.

<https://blogs.oracle.com/ai-and-datascience/post/introduction-to-anomaly-detection>

City News Service. (2020). Caregiver Accused of Killing Elderly Woman in Accidental Shooting While Intoxicated.

<https://www.nbclosangeles.com/news/local/caregiver-accused-of-killing-elderly-woman-in-accidental-shooting-while-intoxicated/2442176/>

Coconuts Manila. (2020). More elderly Pinoys risk abuse during pandemic, warns Commission on Human Rights.

<https://coconuts.co/manila/news/more-elderly-pinoys-will-be-abused-during-pandemic-warns-commission-on-human-rights/>

Goyala, J. E. P., (2019). Traditional Filipino Family Values that make us Distinct among Others. Retrieved from.

<https://medium.com/@janeellapangoyala/traditional-filipino-family-values-that-make-us-distinct-among-others-7a439fd28a>

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

Hamm, J., Money, A.G., Atwal, A., Paraskevopoulos, I.
 (2016) Fall prevention intervention technologies: A
 conceptual framework and survey of the state of the art. J
 Biomed Inform. <https://doi.org/10.1016/j.jbi.2015.12.013>

Hao, W., Zhang, R., Li, S., Li, J., Fuzhong, Li, Zhao, S.,
 Zhang, W. (2020). Anomaly Event Detection in Security
 Surveillance Using Two-Stream Based Model, Security and
 Communication Networks, vol. 2020, Article ID 8876056, 15
 pages. <https://doi.org/10.1155/2020/8876056>

Harari, Y., Shawen, N., Mummidisetty, C.K. et al. (2021).
 A smartphone-based online system for fall detection with
 alert notifications and contextual information of real-life
 falls. J NeuroEngineering Rehabil 18, 124.
<https://doi.org/10.1186/s12984-021-00918-z>

Hayley, A. C., Williams, L. J., Kennedy, G. A., Holloway,
 K. L., Berk, M., Brennan-Olsen, S. L. & Pasco, J. A.
 (2015). Excessive daytime sleepiness and falls among older
 men and women: cross- sectional examination of a

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

population-based sample. BMC GERIATRICS, 15 (1),

<https://doi.org/10.1186/s12877-015-0068-2>

Hazrati, M., Mashayekh, M., Sharifi, N., Motalebi, S.A.

(2020) Screening for domestic abuse and its relationship with demographic variables among elderly individuals referred to primary health care centers of Shiraz in 2018.

BMC Geriatr. 2020 Aug 17;20(1):291. doi: 10.1186/s12877-020-01667-9. PMID: 32807091; PMCID: PMC7430016.

<https://doi.org/10.1186/s12877-020-01667-9>

Hochreiter, S., Schmidhuber J. (1997) Long Short-Term

Memory. Neural Comput 1997; 9 (8): 1735-1780. doi:

<https://doi.org/10.1162/neco.1997.9.8.1735>

Islam, Z., Rukonuzzaman, M., Ahmed, R., Kabir, H., Farazi,

M. (2021). Efficient Two-Stream Network for Violence

Detection Using Separable Convolutional LSTM.

<https://arxiv.org/pdf/2102.10590.pdf>

Jabbar, R., Shinoy, M., Kharbeche, M., Al-Khalifa, K.,

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Krichen, M., Barkaoui, K. (2020). Driver Drowsiness Detection Model Using Convolutional Neural Networks Techniques for Android Application.
https://www.researchgate.net/publication/339164278_Driver_Drowsiness_Detection_Model_Using_Convolutional_Neural_Networks_Techniques_for_Android_Application

Johnson, J. (2020). Anomaly Detection with Machine Learning: An Introduction. Retrieved from
<https://www.bmc.com/blogs/machine-learning-anomaly-detection>

Kadlaskar, A. (2021). How to create a Threat Detection Model using YOLOv3.
<https://www.analyticsvidhya.com/blog/2021/10/how-to-create-a-threat-detection-model-using-yolov3/>

Marr, B. (2016). What Is The Difference Between Artificial Intelligence And Machine Learning?
<https://www.forbes.com/sites/bernardmarr/2016/12/06/what-is-the-difference-between-artificial-intelligence-and->

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

[machine-learning/?sh=6af8fb776bdf](#)

Marr, B. (2018). The Key Definitions Of Artificial Intelligence (AI) That Explain Its Importance.

<https://www.forbes.com/sites/bernardmarr/2018/02/14/the-key-definitions-of-artificial-intelligence-ai-that-explain-its-importance/?sh=2cba72504f5d>

Narejo, S., Pandey, B., Esenarro vargas, D., Rodriguez, C., Anjum, M.R., (2021). Weapon Detection Using YOLO V3 for Smart Surveillance System. Mathematical Problems in Engineering, vol. 2021, Article ID 9975700, 9 pages, 2021.

<https://doi.org/10.1155/2021/9975700>

Parthasarathy, S. (2019). Key differences between Artificial Intelligence and Machine Learning.

<https://towardsdatascience.com/key-differences-between-artificial-intelligence-and-machine-learning-fe637cd0deca>

PreScouter Editorial Team. (2021). What are the applications of AI in anomaly detection?.

<https://www.prescouter.com/2021/05/how-can-ai-be-used-in->

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

anomaly-detection/

Ramzan, M., Abid, A., Khan, H., Awan, S., Ismail, A., Ilyas, M., & Mahmood, A. (2019). A Review on State-of-the-Art Violence Detection Techniques. IEEE Access. 7. 107560 - 107575.10.1109/ACCESS.2019.2932114.

<http://dx.doi.org/10.1109/ACCESS.2019.2932114>

Ren, Y., & Peng L. (2019). Research of Fall Detection and Fall Prevention Technologies: A Systematic Review, in IEEE Access, vol. 7, pp. 77702- 77722.

<https://ieeexplore.ieee.org/abstract/document/8736227>

Roy, R (2020). AI, ML, and DL: How not to get them mixed!

<https://towardsdatascience.com/understanding-the-difference-between-ai-ml-and-dl-cceb63252a6c>

Shu, F., & Shu, J. (2021). An eight-camera fall detection system using human fall pattern recognition via machine learning by a low-cost android box.

<https://www.nature.com/articles/s41598-021-%2081115-9#Abs1>

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Shiho, K. (2017). Filipino Hospitality and Respect for the Aged. Retrieved from

<https://www.nippon.com/en/features/c02810/>

Singh, A. (2017). Anomaly Detection for Temporal Data using Long Short-Term Memory (LSTM) Silva, A., Bandeira de Mello, R., Schaan, C., Fuchs, F., Redline, S., Fuchs, S.

(2016). Sleep duration and mortality in the elderly: A Systematic review with meta-analysis. BMJ Open. 6. e008119. 10.1136/bmjopen-2015-008119. <https://www.diva-portal.org/smash/get/diva2:1149130/FULLTEXT01.pdf>

Sultani, W. & Chen, C., & Shah, M. (2018). Real-world Anomaly Detection in Surveillance Videos.

<https://www.researchgate.net/publication/322498470> Real-world Anomaly Detection in Surveillance Videos

Sumon, S., Goni, R., Hashem, N., Shahri, T., Rahman, R. (2020). Violence Detection by Pretrained Modules with Different Deep Learning Approaches.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

<https://doi.org/10.1142/S2196888820500013>

Thudumu, S., Branch, P., Jin, J., Singh, J. (2020). A comprehensive survey of anomaly detection techniques for high dimensional big data. J Big Data 7, 42.

<https://doi.org/10.1186/s40537-020-00320-x>

Velastin, S. A., Boghossian, B. A., & Vicencio-Silva, M. A. (2006). A motion-based image processing system for detecting potentially dangerous situations in underground railway stations. Transportation Research Part C: Emerging Technologies, vol. 14, no. 2, pp. 96-113.


<https://doi.org/10.1016/j.trc.2006.05.006>

Zhou, P., Ding, Q., Luo, H., Hou, X. (2018) Violence detection in surveillance video using low- level features.<https://doi.org/10.1371/journal.pone.0203668>

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

Appendices

Appendix A: Letter to the Adviser

Attachment 3			
	INVITATION LETTER FOR ADVISER	Document No.	WVSU-ICT-SOI-03-F03
		Issue No.	1
		Revision No.	0
	WEST VISAYAS STATE UNIVERSITY	Date of Effectivity:	April 27, 2018
		Issued by:	CICT
		Page No.	Page 1 of 1

March 2, 2021

Dr. Evans B. Sansolis
 Instructor I
 College of Information and Communications Technology - West Visayas State University
 La Paz, Iloilo City

Dear Dr. Evans B. Sansolis,

The undersigned are BS in Computer Science Research 1/Thesis 1 students of CICT, this university. Our thesis/capstone project title is *"Video-Based Surveillance for Detecting Anomalous Human Activity using Spatio-Temporal 3D Convolutional Network Autoencoder"*.

Knowing of your expertise in research and on the subject matter, we would like to request you to be our **ADVISER**.

We are positively hoping for your acceptance. Kindly check the corresponding box and affix your signature in the space provided. Thank you very much.

Respectfully yours,

1. Jesmar Troy Adolacion 
2. Nezel Leagogo 
3. Francine Anne Millares 
4. Francis Rey Soldevilla 

PS:

Advisers are tasked to work with the students in providing direction and assistance as needed in their thesis/capstone project. They shall meet with the students weekly or as needed to provide direction, check on progress and assist in resolving problems until such a time that the students passed their defenses and submit their final requirements, as well as, preparing their evaluations and grades.


Action Taken: <input checked="" type="radio"/> I Accept. <input type="radio"/> Sorry. I don't accept.	 EVANS B. SANSOLIS Signature over printed name of the Adviser
---	--

CC:

CICT Dean
 Research Coordinator
 Group
**To be accomplished in 4 copies*

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Appendix B: Letter to the Technical Editor

	ADVISER'S ENDORSEMENT FORM (For Thesis Manuscript)	Document No.	WVSU-ICT-SOI-03-F10
		Issue No.	1
	WEST VISAYAS STATE UNIVERSITY	Revision No.	0
		Date of Effectivity:	April 27, 2018
		Issued by:	CICT
	Page No.	Page 1 of 1	

Respectfully endorsed to the **Technical Editor**, the attached manuscript of the thesis entitled:

AI/ML for Anomalous Surveillance: A Detection and Alert System

Said manuscript has been presented to me for preliminary evaluation and guidance, and after a series of corrections/directions given which was implemented by the proponents whose names are listed hereunder and their thorough research, we have come to its completion.

Now therefore, I hereby **ENDORSE** the said thesis manuscript to the Technical Editor for **TECHNICAL EDITING**.


EVANS B. SANSOLIS Ph.D.
 Adviser's Name & Signature

Date: January 14, 2022

Group Members:

1. Francis Rey M. Soldevilla
2. Francine Anne N. Millares
3. Jesmar Troy A. Adolacion
4. Nezel C. Lesgogo

Note: This form should be accomplished and signed if the corrections and changes made by the adviser have been implemented and a new copy of the document have been printed for checking and submission to the next editor

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

Appendix C: Letter to the English Editor


	TECHNICAL EDITOR'S ENDORSEMENT FORM (For Thesis Manuscript)	Document No.	WVSU-ICT-SOI-03-F11
		Issue No.	1
	WEST VISAYAS STATE UNIVERSITY	Revision No.	0
		Date of Effectivity:	April 27, 2018
		Issued by:	CICT
	Page No.	Page 1 of 1	

Respectfully endorsed to the **English Editor**, the attached manuscript of the thesis entitled:

“AI/ML for Anomalous Surveillance: A Detection and Alert System”

Said manuscript was presented to me and was reviewed and edited in terms of technical specifications, correctness of diagrams and other technical matters. The corrections and suggestions was carried and implemented by the proponents whose names are listed hereunder.

Now therefore, I hereby **ENDORSE** the said thesis manuscript to the English Editor/Grammarian for **English Grammar Editing**.


Dr. Frank U. Eljorje
 Technical Editor's Name & Signature

Date: February 7, 2022


Group Members:

1. Francis Rey M. Soldevilla
2. Francine Anne N. Millares
3. Nezel C. Leagogo
4. Jesmar Troy A. Adolacion

Note: This form should be accomplished and signed if the corrections and changes made by the Technical Editor have been implemented and a new copy of the document have been printed for checking and submission to the next editor.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Appendix D: Letter to the Thesis Format Editor

	ENGLISH EDITOR/GRAMMARIAN'S ENDORSEMENT FORM (For Thesis Manuscript)	Document No.	WVSU-ICT-SOI-03-F12
		Issue No.	1
		Revision No.	0
	WEST VISAYAS STATE UNIVERSITY	Date of Effectivity:	April 27, 2018
		Issued by:	CICT
		Page No.	Page 1 of 1

Respectfully endorsed to the **Thesis Format Editor**, the attached manuscript of the thesis entitled:

AI/ML for Anomalous Surveillance: A Detection and Alert System

Said manuscript was presented to me for English grammar editing, corrections has been made and the proponents whose names are listed hereunder implemented said corrections and changes in the revised manuscript.

Now therefore, I hereby **ENDORSE** the said thesis manuscript for **Thesis Format Editing**.



Prof MARIEVIC M. VIOLETA
 English Editor/Grammarian's Name and Signature

Date: March 15, 2022


Group Members:

1. Francis Rey M. Soldevilla
2. Francine Anne N. Millares
3. Jesmar Troy A. Adolacion
4. Nezel C. Leagogo

Note: This form should be accomplished and signed if the corrections and changes made by the English Editor have been implemented and a new copy of the document have been printed for checking and submission to the next editor.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

*Appendix E: Letter to the Undergraduate Research
Coordinator*

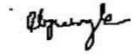
	THESIS FORMAT EDITOR'S ENDORSEMENT FORM (For Thesis Manuscript)	Document No. WVSU-ICT-SOI-03-F13
		Issue No. 1
	WEST VISAYAS STATE UNIVERSITY	Revision No. 0
		Date of Effectivity: April 27, 2018
		Issued by: CICT
	Page No. Page 1 of 1	

Respectfully endorsed to the **Thesis Coordinator**, the attached manuscript of the thesis entitled:

**“ AI/ML FOR ANOMALOUS SURVEILLANCE:
A DETECTION AND ALERT SYSTEM “**

Said manuscript was presented to me and has checked the preliminaries, thesis document convention and end matters, made some corrections which was implemented by the proponents whose names are listed hereunder.

Now therefore, I hereby **ENDORSE** said manuscript to the Thesis Coordinator for appropriate action.



DR. MA. LUCIE P. SABAYLE
 Thesis Format Editor's Name and Signature

Date: May 19, 2022

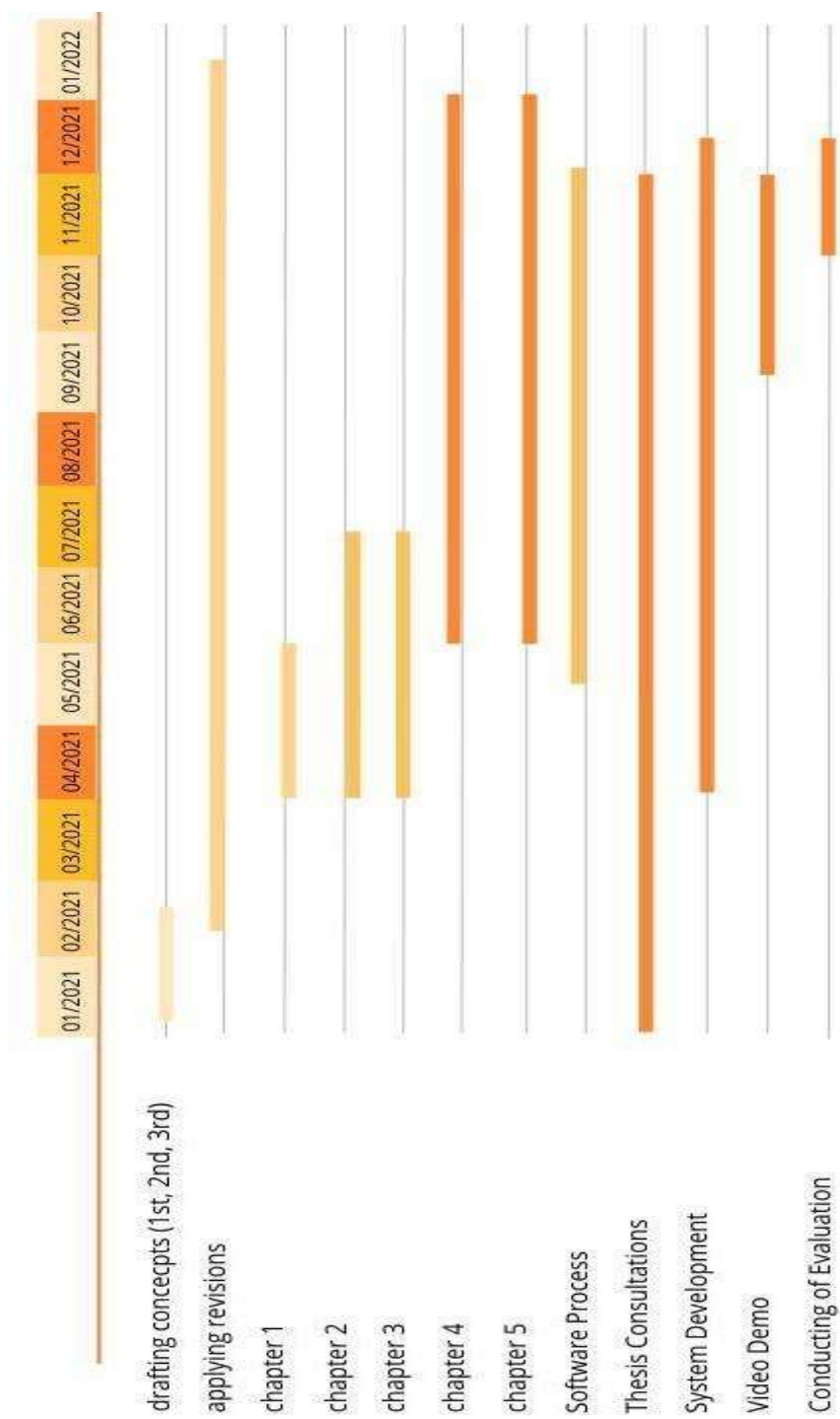
Group Members:

1. Francis Rey Soldevilla
2. Francine Anne Millares
3. Jesmar Troy Adolacion
4. Nezel Leagogo

Note: This form should be accomplished and signed if the corrections and changes made by the Thesis Format Editor have been implemented and the four (4) new copies have been printed ready for bookbinding.

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Appendix F: Gantt Chart



West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

Appendix G: Data Dictionary

Table Name: Fall Detection System

Data Item	Data Type	Description	Example
tagged_df	Boolean		None
cam	Integer	Reads camera index	cv2.VideoCapture(0)
img	Integer	Reads inputted images	1
height	Integer	Sets height of the window	500
width	Integer	Sets width of the window	800
width_height	Integer	Readjusts window based on the size of the inputted video or image	(500 * 800 // 16) * 16
output_video	Boolean	Sets initial	None

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

		value of the output video as False	
frame	Literal	Sets initial value	0
fps	Literal	Sets initial value	0
t0	Float	Reads time	time.time()

Table Name: Violence Detection System

Data Item	Data Type	Description	Example
width	Integer	Sets width of the window	500
height	Integer	Sets height of the window	500
frameRate	Integer	Sets initial value	0

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

capture	Any	Reads camera index	cv2.VideoCapture(0)
flag	Boolean		False
startDetect Time	Float	Reads the start of the detection phase	time.time()
isFighting	Boolean	Returns Boolean if system detects anomaly	True
endDetectTime	Float	Reads the end of the detection phase	time.time()
targetSize	Integer		500
frame	Any	Capture frame- by- frame from camera	capture.read())

Table Name: Weapon Detection System

Data Item	Data	Description	Example
-----------	------	-------------	---------

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

	Type		
mixer	Any	Reads mp3 file	alarm.wav
net	Any	Reads pre-trained models	yolov3_training_2000.weights
classes	String	Reads classes of the weapon detected	
cap	Any	Reads camera index	cv2.VideoCapture(0)
font	Integer	Sets font style	cv2.FONT_HERSHEY_SIMPLEX
starting_time	Float	Reads the start of the detection phase	0
frame_id	Literal	Reads frames from webcam	1

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

center_x	Integer	Detects x coordinate	detection[0] * width
center_y	Integer	Detects y coordinate	detection[1] * height
x	Integer	Rectangle x coordinate	center_x - w / 2

Table Name: Drowsiness Detection System

Data Item	Data Type	Description	Example
sound	Any	Reads mp3 file	alarm.wav
face	String	Reads xml file for the detection of face	haarcascade_f ro ntalface_alt. xm l
leye	String	Reads xml file for the detection of left eye	haarcascade_l ef teye_2splits. xm l

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

reye	String	Reads xml file for the detection of right eye	haarcascade_r ig hteye_2splits .x ml
lbl	String	Labels eyes close or open	Close, Open
model	String	Reads h5 file	Cnncat2.h5
cap	Any	Reads camera index	cv2.VideoCapture ur e(0)
font	Integer	Sets font style	cv2.FONT_HERSHEY HE Y_COMPLEX
count	Literal	Sets initial value for the counter	0
score	Literal	Sets initial value for the scoring	0

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

Appendix H: Sample Program Codes

Sample Code: Fall Detection System

```
import cv2

import logging

import time

import numpy as np

import matplotlib.pyplot as plt

from vis.visual import write_on_image, visualise,
activity_dict,
visualise_tracking

from vis.processor import Processor

from helpers import pop_and_add, last_ip, dist,
move_figure, get_hist

from default_params import *

from vis.inv_pendulum import *

import pandas as pd

from scipy.signal import savgol_filter, lfilter

from model.model import LSTMModel

import torch

import math

import pygame
```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

```

pygame.mixer.init()

pygame.mixer.music.load('./alarm.wav')

def get_source(args):
    tagged_df = None

    cam = cv2.VideoCapture(0, cv2.CAP_DSHOW)

    img = cam.read()[1]

    logging.debug('Image shape:', img.shape)

    return cam, tagged_df

def resize(img, resize, resolution):
    if resize is None:
        height, width = img.shape[:2]
    else:
        width, height = [int(dim) for dim in
                           resize.split('x')]
    width_height = (int(width * resolution // 16) *
16,
                    int(height * resolution // 16) * 16)

```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

```

        return width, height, width_height

def extract_keypoints_parallel(queue, args,
self_counter, other_counter,
consecutive_frames, event):
    try:
        cam, tagged_df = get_source(args)

        ret_val, img = cam.read()
except Exception as e:
    queue.put(None)

    event.set()

    print('Exception occurred:', e)

    print('Most likely that the video/camera doesn\'t
exist')

    return

width, height, width_height = resize(img, args.resize,
args.resolution)

logging.debug(f'Target width and height =
{width_height}')

processor_singleton = Processor(width_height, args)

output_video = None

```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

```

frame = 0

fps = 0

t0 = time.time()

while not event.is_set():

    if args.num_cams == 2 and (self_counter.value >
other_counter.value):

        continue

    ret_val, img = cam.read()

    frame += 1

    self_counter.value += 1

    if tagged_df is None:

        curr_time = time.time()

    else:

        curr_time = tagged_df.iloc[frame-
1]['TimeStamps'][11:]

        curr_time = sum(x * float(t) for x, t in
zip([3600, 60, 1],

curr_time.split(":"))))

    if img is None:

```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

```

print('no more images captured')

print(args.video, curr_time, sep=" ")

if not event.is_set():
    event.set()

break

img = cv2.resize(img, (width, height))

hsv_img = cv2.cvtColor(img, cv2.COLOR_BGR2HSV)

keypoint_sets, bb_list, width_height =
processor_singleton.single_image(img)

assert bb_list is None or (type(bb_list) == list)

if bb_list:
    assert type(bb_list[0]) == tuple
    assert type(bb_list[0][0]) == tuple

if args.coco_points:
    keypoint_sets = [keypoints.tolist() for keypoints
in keypoint_sets]

else:
    anns = [get_kp(keypoints.tolist()) for
keypoints in keypoint_sets]

ubboxes = [(np.asarray([width,
```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

```

height]))*np.asarray(ann[1])).astype('int32')

        for ann in anns]

    lbboxes = [(np.asarray([width,
height]))*np.asarray(ann[2])).astype('int32')

    for ann in anns]

    bbox_list = [(np.asarray([width,

height]))*np.asarray(box)).astype('int32') for box
in bb_list]

    uhlist_list = [get_hist(hsv_img, bbox) for bbox in
ubboxes]

    lhlist_list = [get_hist(img, bbox) for bbox in
lbboxes]

    keypoint_sets = [{"keypoints": keyp[0],
"up_hist":uh,
"lo_hist":lh, "time":curr_time, "box":box}
for keyp, uh, lh, box in zip(anns, uhlist_list,
lhlist_list, bbox_list)]

cv2.polylines(img, ubboxes, True, (255, 0, 0), 2)

```


West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

```

cv2.polylines(img, lbboxes, True, (0, 255, 0), 2)

for box in bbox_list:

    cv2.rectangle(img, tuple(box[0]), tuple(box[1]),
((0, 0, 255)), 2)

    dict_vis = {"img": img, "keypoint_sets":
keypoint_sets, "width":
    width, "height": height, "vis_keypoints":
args.joints,
    "vis_skeleton": args.skeleton, "CocoPointsOn":
args.coco_points,
    "tagged_df": {"text": f"", "color": [0, 0, 0]}}

queue.put(dict_vis)

queue.put(None)

return

```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Sample Code: Violence Detection System

```
import cv2
import os

import sys

Import numpy as np

import time

from src.ViolenceDetector import ViolenceDetector

from settings import DeployLiveSettings as
deploySettings

from settings import DataSettings as dataSettings

from src.data

import ImageUtils as ImageUtils

import pygame #For playing sound

pygame.mixer.init()

pygame.mixer.music.load('./alarm.wav')

class VideoSaver:

    def AppendFrame(self, image_):

        self.outputStream.write(image_)
```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

```
def _init_(self, targetFileName, videoCapture): width
    = int(deploySettings.DISPLAY_IMAGE_SIZE) height =
    int(deploySettings.DISPLAY_IMAGE_SIZE) frameRate =
    int(videoCapture.get(cv2.CAP_PROP_FPS)) codec =
    cv2.VideoWriter_fourcc(*'XVID')

    self.outputStream = cv2.VideoWriter(targetFileName
+ ".avi", codec, frameRate, (width, height))

def DetectViolence(PATH_FILE_NAME_TO_SAVE_RESULT): font
    = cv2.FONT_HERSHEY_SIMPLEX

    violenceDetector = ViolenceDetector() capture =
    cv2.VideoCapture(0, cv2.CAP_DSHOW)

    shouldSaveResult = (PATH_FILE_NAME_TO_SAVE_RESULT !=
    None)

    if shouldSaveResult:

        videoSavor =
        VideoSavor(PATH_FILE_NAME_TO_SAVE_RESULT + "_Result",
        capture)

        listOfForwardTime = []
```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

```
frame_width = capture.get(cv2.CAP_PROP_FRAME_WIDTH)
frame_height = capture.get(cv2.CAP_PROP_FRAME_HEIGHT)
fps = capture.get(cv2.CAP_PROP_FPS)

if capture.isOpened() is False: print("Error opening
    the camera")

flag = False

while capture.isOpened():
    ret, frame = capture.read()
```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

```

    if ret:

        assert not isinstance(frame, type(None)), 'Frame
        not Found'

    if ret is True:

        netInput = ImageUtils.ConvertImageFrom_CV_to
        _NetInput(frame)

        startDetectTime = time.time()

        isFighting = violenceDetector.Detect(netInput)

        endDetectTime = time.time()

        listOfForwardTime.append(endDetectTime -
        startDetectTime)

        targetSize = deploySettings.DISPLAY_IMAGE_SIZE - 2 *
                                deploySettings.BORDER_SIZE

        currentImage = cv2.resize
                                (frame, (targetSize, targetSize))

    if isFighting:

        pygame.mixer.music.play(-1)

        resultImage = cv2.copyMakeBorder(frame,
                                deploySettings.BORDER_SIZE,
                                deploySettings.BORDER_SIZE,

```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

```

                                deploySettings.BORDER_SIZE,
                                deploySettings.BORDER_SIZE,
                                cv2.BORDER_CONSTANT,
value=deploySettings.FIGHT_BORDER_COLOR)

    else:

        pygame.mixer.music.stop()

        resultImage = cv2.copyMakeBorder(frame,

                                deploySettings.BORDER_SIZE,
                                deploySettings.BORDER_SIZE,
                                deploySettings.BORDER_SIZE,
                                deploySettings.BORDER_SIZE,
                                cv2.BORDER_CONSTANT,

value=deploySettings.NO_FIGHT_BORDER_COLOR)

cv2.namedWindow("", flags=cv2.WINDOW_GUI_NORMAL)
cv2.setWindowProperty("", cv2.WND_PROP_FULLSCREEN,
cv2.WINDOW_FULLSCREEN)
cv2.setWindowProperty("", cv2.WND_PROP_TOPMOST, 1)
cv2.resizeWindow("", 800, 600)

```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

```

cv2.moveWindow("", 370, 75)

cv2.imshow("", resultImage)

if shouldSaveResult:

    videoSavor.AppendFrame(resultImage)

    if cv2.waitKey(1) & 0xFF == ord('q'):

        capture.release()

        cv2.destroyAllWindows()

        flag = True

        break

    else:

        isCurrentFrameValid, currentImage = capture.read()

        averagedForwardTime = np.mean(listOfForwardTime)

        if flag:

            break

if __name__ == '__main__':

    print("\n\n\n")

    try:

        PATH_FILE_NAME_TO_SAVE_RESULT = sys.argv[1]+"\"

    except:

```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

```

PATH_FILE_NAME_TO_SAVE_RESULT =
"C:\\Users\\asus\\Desktop\\AnomalyDetectionSystem\\Violence
-Detection\\results\\DeployResults\\"
DetectViolence(PATH_FILE_NAME_TO_SAVE_RESULT)

Sample Code: Weapon Detection System

import cv2

import numpy as np import time

from pygame import mixer

mixer.init() mixer.music.load('./alarm.wav')

net = cv2.dnn.readNet("./yolov3_training_2000.weights",
"./yolov3_testing.cfg")

classes = [""]

layer_names = net.getLayerNames() output_layers =
[layer_names[i[0] - 1] for i in
net.getUnconnectedOutLayers()]

colors = np.random.uniform(0, 255, size=(len(classes),
3))

```


West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

```
cap = cv2.VideoCapture(0, cv2.CAP_DSHOW) font =
cv2.FONT_HERSHEY_SIMPLEX
starting_time = time.time() frame_id = 0

while True:
    _, frame = cap.read() frame_id += 1
    height, width, channels = frame.shape

    blob = cv2.dnn.blobFromImage(frame, 0.00392, (416,
416), (0, 0, 0), True, crop=False)
    net.setInput(blob)
    outs = net.forward(output_layers)

    # Visualising data class_ids = [] confidences = []
    boxes = []

    for out in outs:
        for detection in out: scores = detection[5:]
            class_id = np.argmax(scores) confidence =
            scores[class_id] if confidence > 0.1:
```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

```

        center_x = int(detection[0] * width)
        center_y = int(detection[1] * height) w =
        int(detection[2] * width)
        h = int(detection[3] * height)
        x = int(center_x - w / 2) y = int(center_y -
        h / 2)

        boxes.append([x, y, w, h])
        confidences.append(float(confidence))
        class_ids.append(class_id)

indexes = cv2.dnn.NMSBoxes(boxes, confidences, 0.8,
0.3) if indexes == 0:
    frame = cv2.copyMakeBorder(frame, 10, 10, 10, 10,
cv2.BORDER_CONSTANT, None, value=[0, 0, 255])
    mixer.music.play(-1) else:
        frame = cv2.copyMakeBorder(frame, 10, 10, 10, 10,
cv2.BORDER_CONSTANT, None, value=[0, 255, 0])
        mixer.music.stop()

for i in range(len(boxes)): if i in indexes:
    x, y, w, h = boxes[i]

```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

```

        label = str(classes[class_ids[i]]) confidence =
        confidences[i]

        color = colors[class_ids[i]]

        cv2.rectangle(frame, (x, y), (x + w, y + h),
        color, 2)

    elapsed_time = time.time() - starting_time fps =
    frame_id / elapsed_time

    cv2.namedWindow("Image", flags=cv2.WINDOW_GUI_NORMAL)
    cv2.setWindowProperty("Image",
    cv2.WND_PROP_FULLSCREEN,
    cv2.WINDOW_FULLSCREEN)

    cv2.setWindowProperty("Image", cv2.WND_PROP_TOPMOST,
    1)

    cv2.resizeWindow("Image", 800, 600)

    cv2.moveWindow("Image", 370, 75) cv2.imshow("Image",
    frame)

    key = cv2.waitKey(1)

    if cv2.waitKey(1) & 0xFF == ord('q'): break

cap.release() cv2.destroyAllWindows()

```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

Sample Code: Drowsiness Detection System

```
import cv2 import os

from tensorflow.keras.models import load_model

import numpy as np

from pygame import mixer import time

mixer.init()

sound = mixer.Sound('alarm.wav')

face = cv2.CascadeClassifier('haar cascade
files\haarcascade_frontalface_alt.xml') leye =
cv2.CascadeClassifier('haar cascade
files\haarcascade_lefteye_2splits.xml') reye =
cv2.CascadeClassifier('haar cascade
files\haarcascade_righteye_2splits.xml')

lbl=['Close', 'Open']

model = load_model('model/cnn-cat2.h5') path =
os.getcwd()
```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

```

cap = cv2.VideoCapture(0)

font = cv2.FONT_HERSHEY_COMPLEX_SMALL

count=0

score=0

thicc=2

rpred=[99]

lpred=[99]

while(True):

    ret, frame = cap.read() height,width =
    frame.shape[:2]

    gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY) faces
    =
face.detectMultiScale(gray,minNeighbors=5,scaleFactor=1.
1,minSize=(25,25)) left_eye =
leye.detectMultiScale(gray)

right_eye = reye.detectMultiScale(gray)

for (x,y,w,h) in faces:

    cv2.rectangle(frame, (x,y) , (x+w,y+h) ,
    (100,100,100) , 1 )

```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

```
for (x,y,w,h) in right_eye: r_eye=frame[y:y+h,x:x+w]

count=count+1

r_eye = cv2.cvtColor(r_eye,cv2.COLOR_BGR2GRAY)
```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

```

r_eye = cv2.resize(r_eye, (24,24))

r_eye= r_eye/255

r_eye= r_eye.reshape(24,24,-1) r_eye =
np.expand_dims(r_eye,axis=0) rpred =
model.predict_classes(r_eye) if(rpred[0]==1):
    lbl='Open' if(rpred[0]==0):
        lbl='Closed'
    break

for (x,y,w,h) in left_eye: l_eye=frame[y:y+h,x:x+w]
    count=count+1

l_eye = cv2.cvtColor(l_eye,cv2.COLOR_BGR2GRAY)
l_eye = cv2.resize(l_eye, (24,24))
l_eye= l_eye/255 l_eye=l_eye.reshape(24,24,-1)
l_eye = np.expand_dims(l_eye,axis=0) lpred =
model.predict_classes(l_eye) if(lpred[0]==1):
    lbl='Open' if(lpred[0]==0):
        lbl='Closed'
    break

if(rpred[0]==0 and lpred[0]==0): score=score+1

```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

```

else:

    score=score-1

if (score<0):

    score=0
    if (score>10):
        try:
            sound.play()
        except:
            pass
    if (thicc<16):
        thicc= thicc+2
    else:
        thicc=thicc-2
        if (thicc<2):
            thicc=2

frame = cv2.copyMakeBorder (src=frame,

                             top= 10,

                             bottom= 10,

                             left= 10,

                             right= 10,

                             borderType=cv2.BORDER_CONSTANT,

                             dst=None,

                             value=[0, 0, 255])

```


West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

```

else:

    frame = cv2.copyMakeBorder(src=frame,

                                top= 10,

                                bottom= 10,

                                left= 10,

                                right= 10,

                                borderType=cv2.BORDER_CONSTANT,

                                dst=None,

                                value=[0, 255, 0])

    cv2.namedWindow('frame', flags=cv2.WINDOW_GUI_NORMAL)

    cv2.setWindowProperty('frame',

                           cv2.WND_PROP_FULLSCREEN,

                           cv2.WINDOW_FULLSCREEN)

    cv2.setWindowProperty('frame', cv2.WND_PROP_TOPMOST,

                           1)

    cv2.resizeWindow('frame', 800, 600)

    cv2.moveWindow('frame', 370, 75)

    cv2.imshow('frame', frame)

    if cv2.waitKey(1) & 0xFF == ord('q'): break

cap.release()

cv2.destroyAllWindows()

```

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Paz, Iloilo City, Philippines

Appendix I: Software Quality Survey Form

Name (Optional): _____

Gender: ☐ Male ☐ Female

<u>Scale</u>	<u>Description</u>
5	Excellent
4	Very Good
3	Good
2	Fair
1	Poor

Characteristic s	Sub-characteristics	Rating				
		1	2	3	4	5
Functional Suitability	The system has complete functions					
	The functions are correct					
	The functions are appropriate for the					

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

	system					
Performance Efficiency	The performance time is fast					
	It makes use of available resources					
	The capacity deals with a good amount of information					
Compatibility	It coexists with other systems					
	Interoperability: it provides real time information					

Usability	Interoperability: all information is easy to understand					
	It is easy to recognize its features					
	It is easy to learn its					

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Iga, Iloilo City, Philippines

	functions					
	It is easily be operated					
	User error protection					
	User interface aesthetics					
	It is accessible					
Reliability	It is available and ready to carry its task anywhere					
	It is able to continue its normal operation despite the presence of software faults					
	It recovers after going through any crash					
	It keeps the user's personal data secured					
	It is safe to use					

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Ia Iga, Iloilo City, Philippines

	It is not counterfeit and it is authentic					
	Modularity: the decomposition of a program into smaller programs with standardized interface					
	It is reusable					
Maintainability	It is analyzable					
	It is modifiable					

Portability	It is adaptable and can be used anywhere					
	It is easy to install the system					

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City, Philippines

Appendix J: Disclaimer

Disclaimer

This software project and its corresponding documentation entitled "AI/ML for Anomalous Surveillance: A Detection and Alert System" is submitted to the College of Information and Communications Technology, West Visayas State University, in partial fulfillment of the requirements for the degree, Bachelor of Science in Computer Science. It is the product of our own work, except where indicated text.

We hereby grant the College of Information and Communications Technology permission to freely use, publish in local or international journal/conferences, reproduce, or distribute publicly the paper and electronic copies of this software project and its corresponding documentation in whole or in part, provided that we are acknowledged.

Jesmar Troy A. Adolacion

Nezel C. Leagogo

Francine Anne N. Millares

Francis Rey M. Soldevilla

June 2022