

Anomalous Weapon Detection for Armed Robbery Using Yolo V8 [†]

Adrian Lester E. Reyes ^{*} and Jennifer C. Dela Cruz

School of Electrical, Electronics, and Computer Engineering Mapúa University, Manila 1002, Philippines;
jcdelacruz@mapua.edu.ph

^{*} Correspondence: alereyes@mymail.mapua.edu.ph

[†] Presented at the 2024 IEEE 6th Eurasia Conference on IoT, Communication and Engineering, Yunlin, Taiwan, 15–17 November 2024.

Abstract: Improved surveillance systems provide early warnings and improve public safety. Such systems are desperately needed in light of the rising number of armed robberies in private and public places. A YOLOv8-based system specifically intended for CCTV-based armed robbery detection was developed to meet this demand in this study. The system identified weapons such as handguns, assault weapons, shotguns, and others in real-time, utilizing a custom-trained model. The system demonstrated a strong performance with an overall anomaly detection accuracy of 87.50%. The confidence level was 1.2 m (58.79) and 2 m (59.74) in determining the optimal height and distance considering the positioning of the CCTV camera. The low confidence level was attributed to the mixture of images from a general database from the Internet along with self-captured images that resulted in the overfitting of the datasets. Although improvements are needed to increase the confidence level by using real guns in training the model and reducing false negatives, the potential of YOLOv8 to enhance public safety has been confirmed by providing early warnings of armed robberies.

Keywords: YOLO; robbery; detection; anomalous weapons

1. Introduction

The You Only Look Once (YOLO) v5 algorithm is used to classify objects based on their height and size by generating a dataset for training and testing. The size and height of the detected items are determined depending on the YOLO model's classification. The bounding box of the identified object is used to compute the distance between the object and the camera. The normalized distance is utilized to precisely measure the object's size [1]. By applying YOLO v8's object detection on a closed-circuit television camera for public and private use, we created a model to detect harmful objects, such as guns. We gathered the data, which were processed by a convolutional neural network (CNN) [2].

A concept of gun detection in surveillance videos using deep neural networks was established in [3]. The R-CNN detector was used for testing with 608 images and training with 3000 images. The Mobilenet-SSD, FPN, RefineDet, and M2Det detectors were also used to improve accuracy and precise object localization.

However, limited functionality must be addressed to detect harmful objects such as guns effectively. The YOLOv5's object detection was used to build and construct a leaf blight detection system for image processing, a motion sensor alarm system, and SMS texting [4]. Yolo is also used for object detection, such as for cellphones [5]. To create a model for object detection of guns, the measurement size, including the width and focal length of an object from the footage, was used to prevent a decrease in the accuracy of the



Academic Editors: Teen-Hang Meen,
Chi-Ting Ho and Cheng-Fu Yang

Published: 27 May 2025

Citation: Reyes, A.L.E.; Cruz, J.C.D. Anomalous Weapon Detection for Armed Robbery Using Yolo V8. *Eng. Proc.* **2025**, *92*, 2085. <https://doi.org/10.3390/engproc2025092085>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

information. However, this does not provide additional information in a situation where no one looks at the camera instantly.

2. Literature Review

2.1. Weapons

For detection using the ARMAS Weapon detection dataset and IMFDB Weapon detection system [6], specific data on guns or weapons used in robbery incidents are limited, and the only database available was constructed in the United States. The most common weapon used was the handgun, with 44,086 cases of robberies using handguns in 2022. The second most common weapon was personal weapons. In 11,797 cases, knives or cutting apparatus were used. [7]

2.2. Yolo V8

YOLO is an artificial intelligence model that detects objects in the environment through a camera. The unique feature of YOLO is its ability to detect multiple objects with CNN [8]. The YOLO algorithm utilizes the region of interest (ROI) to detect multiple objects, and CNN of the hidden Markov model is used to detect multiple objects in one image. YOLOV8 is the most advanced version of YOLO, with upgrades in its backbone that improve accuracy and speed compared to the older version [9].

2.3. CNN

CNN is a combination of layers that analyze images, audio, or speech input to predict the output. CNN consists of many layers of neural networks: the convolutional layer, the pooling layer, and the fully connected layer. The convolutional layer extracts a feature from an image, filters it, and processes it for insertion into the pooling layer. The pooling layer filters out the necessary processed data to simplify the result and, lastly, the fully connected layers take in the processed data to output the prediction [10].

3. Methodology

We created a system for an early warning for armed robbery using YOLOV8. To identify anomalous objects, a CCTV camera's optimum height and distance were determined. The data were gathered at actual houses under permission. Voice recognition, people counting, and facial recognition were not included in the system's detection.

The image data were collected by using the CCTV camera. The system consisted of an automatic texting system, an early warning device, and an image processing module. Using Robo flow, the data was annotated and used to train the image processing model. The YOLOv8 image classification model was used to process the data and identify weapons. CCTV GUI, Arduino Nano, and SIM800L V2 were used to manage the system's input and output. The anomaly detection results and a live video feed were displayed on a screen (Figure 1).

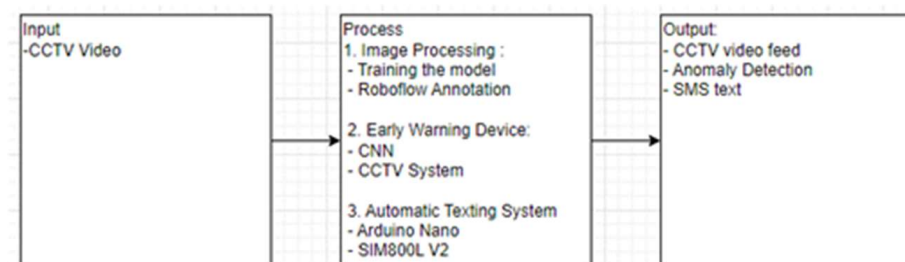


Figure 1. System operation.

3.1. Hardware and Circuit Design

The hardware connection of the system is displayed in Figure 2. The digital video recorder (DVR) system was used with connected CCTV cameras. The laptop or computer was used to receive the streamed and processed videos through the high-definition multimedia interface connection. The automatic texting system and alarm systems were employed to send a text message to warn of unusual objects.

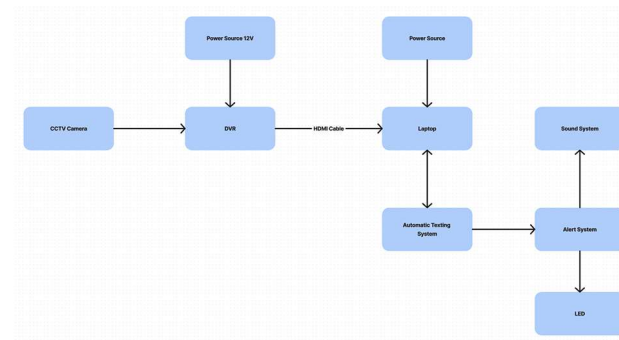


Figure 2. Hardware components.

3.2. System Process Flow

The system started when all of the components were turned on and interconnected. The camera began to take images, and the data were stored in DVR. DVR sent the data from the router to the IoT. The YOLO model processed the image. First, the image was segmented and processed to detect anomaly objects. After detection, the automatic texting system alerted the authorities (Figure 3).

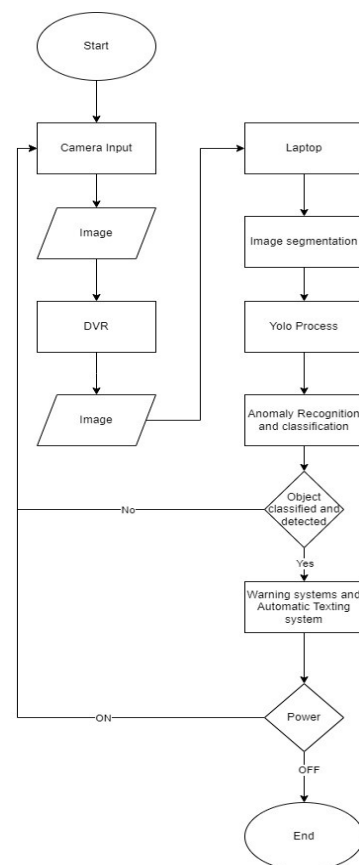


Figure 3. System process.

3.3. YOLO V8

Public databases such as the UCF crime website along with self-made database were used for model testing. The model was trained using Google Collab and images from DVR. 2812 images were used in Robo flow to show a clear view of the objects in each frame or image (Figures 4 and 5).

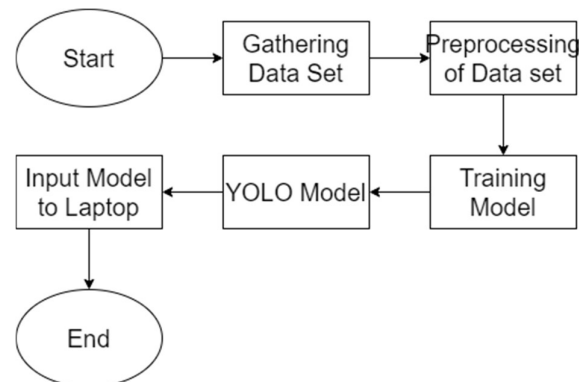


Figure 4. YOLO custom model training.

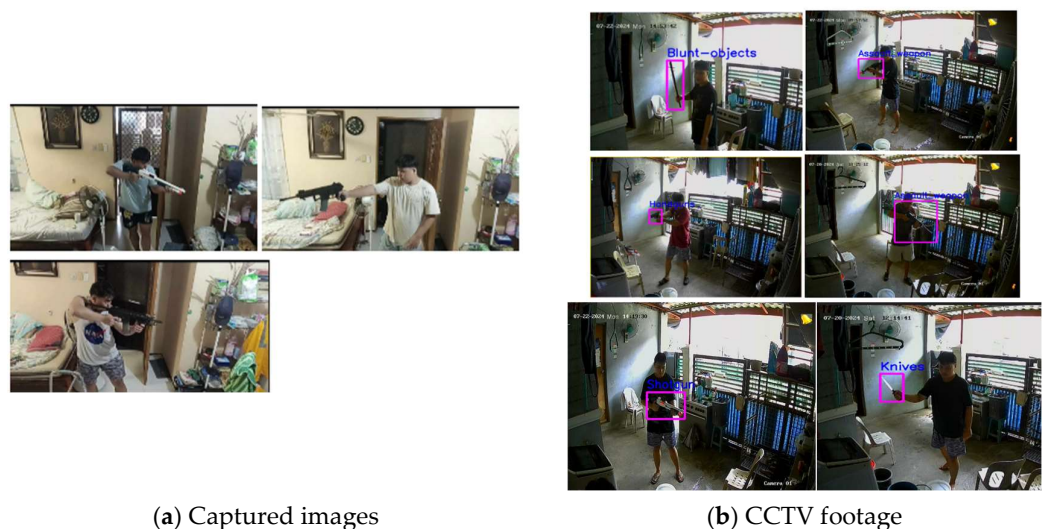


Figure 5. Examples of images and videos from Robo flow for training model.

3.4. Components

In Figure 6, The DVR system was connected to the CCTV cameras and recorded images and videos. The DVR system also streamed the video through the ethernet connected to the router.



Figure 6. DVR in system.

A dome-type CCTV camera recorded video by rotating the optimal viewing angle. It had built-in infrared lights and cameras used at night. The camera has a resolution of 2 MP and 1080p (Figure 7).



Figure 7. CCTV camera.

The Arduino Uno was used to control the automatic texting system. Text messages were sent using an Arduino board with a GSM module. The SIM800L GSM v2 module allowed the Arduino Uno to send and receive short messages and voice chats on the Internet (Figures 8 and 9).

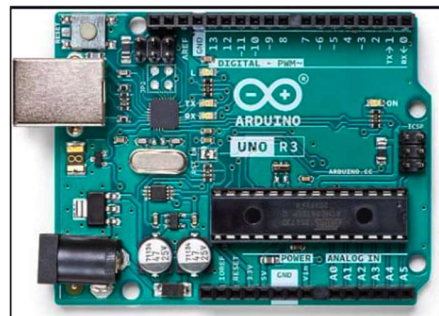


Figure 8. Arduino Uno R3.

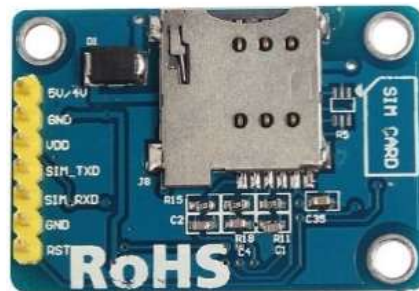


Figure 9. SIM800L GSM v2 module.

3.5. Test

The test was conducted in a moderately sized room that simulated a store (Figure 10). The camera was placed in the corner of the room at the highest position to maximize the viewing angle. Behind the counter where the owner/store clerk was positioned, the actor was assumed to enter the door with various objects to simulate an armed robbery.

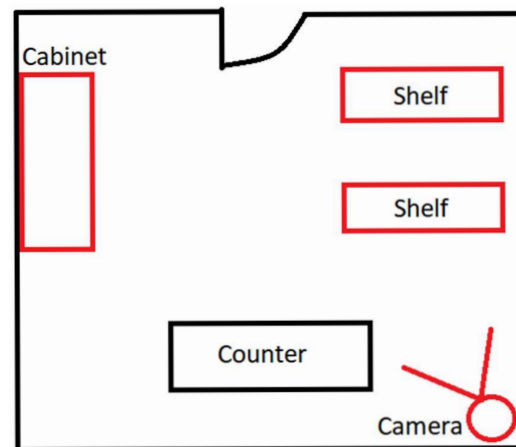


Figure 10. Test room.

3.6. Statistical Analysis

The model's accuracy was calculated at the average confidence level of the model per class. A confusion matrix was used to estimate the accuracy of the anomaly algorithm (1):

$$Accuracy = \frac{\sum(TP \text{ and } TN)}{TP + FP + FN + TN} \quad (1)$$

TP: true positive, correctly identified the anomaly. TN: true negative, correctly identified no anomaly. FP: false positive, incorrectly identified no anomaly. FN: false negative: incorrectly identified an anomaly.

To determine the optimal height and distance, we used different height positionings for the CCTV camera (1, 1.2, 1.4, and 1.8 m). For each height, distances of 1, 1.5, 2, and 2.5 m were used to accurately identify the anomalous weapons.

4. Results and Discussion

Table 1 presents the results of image classification. Eight trials per object were conducted. The confidence level was used as a measure of how confident the model detected an object. Even though the confidence level is low, the model is accurate in determining which anomalous object is present. The low confidence level occurred due to the weapons used when testing, which were not real guns but toy guns. The overfitting of the datasets was mitigated by training the model. The public database was constructed using images from the Internet and self-captured images.

Table 1. Average Image Classification Day time.

Object	Object Detected	Average Confidence Level
Handgun	Handgun	54.52
M16	Assault Weapon	59.01
AK47	Assault Weapon	60.60
SMG	SMG	69.87
Kitchen knives	Knives	53.28
Long stick and Axe	Blunt object	57.36
Shotgun	Shotgun	47.12
Overall Confidence level		57.39

Table 2 presents the performance of the anomaly detection algorithm and its ability to accurately identify anomalies within a controlled environment. The algorithm correctly identified 56 anomalies with no FPs. However, it showed FNs, where anomalies were

not present but an anomaly was detected, which impacted the system's overall accuracy. Despite FNs, the overall accuracy of the system was 87.50%, demonstrating its excellent performance in distinguishing anomalous objects.

Table 2. Anomaly algorithm Confusion Matrix result.

Anomaly Algorithm Confusion Matrix	Sum of Obstacles	Total Trial Observed Object Result Score
Correctly identified the anomaly, True positive.	$\Sigma 8 + 8 + 8 + 8 + 8 + 8 + 8$	56 Object
Correctly identify no anomaly, True Negative.	0	0 Object
Incorrectly identify no anomaly, False Positive.	0	0 Object
Incorrectly identified an anomaly, False Negative.	$\Sigma 8$	8 Object

Determining Optimal Height and Distance Test

Table 3 shows the average confidence level of all objects at given heights and distances. The highest average confidence level was obtained at a height of 1.2 m (58.79) and 1.4 m (57.50) and a distance of 2 m (59.74), and 2.5 m (57.51). Even if the model was accurate in identifying the given weapons, a low confidence level was observed for toy guns. The confidence level was high when using images taken in this study.

Table 3. Overall average of confidence level.

Height	Distance								Average per Height
	1 m	1.5 m	2 m	2.5 m	3 m	3.5 m	4 m	4.5 m	
1 m	50.22	57.125	57.64	57.82	60.38	57.42	55.99	43.87	55.06
1.2 m	55.85	62.19	67.5	65.91	59.51	56.26	57.53	45.59	58.79
1.4 m	54.01	61.65	63.81	60.95	62.76	63.07	44.54	49.19	57.5
1.8 m	48.33	46.14	50.03	45.36	47.2	49.99	49.1	46.04	47.78
Average per distance	52.1	56.78	59.74	57.51	57.46	56.68	51.79	46.17	

5. Conclusions

We developed an early warning system for weapons using a CCTV system and Yolo V8 image processing algorithm. For image classification, the confidence level was 57% in the daytime. The highest average confidence level was obtained at heights of 1.2 m (58.79) and 2 m (59.74). The model showed a low confidence level for images obtained from the Internet and self-capturing, which caused overfitting. The model was accurate in determining weapons. The YOLOv8 model was effective in recognizing various weapons with high accuracy. FNs highlighted the need for refinement to enhance the reliability of the model. Despite these challenges, the system is an effective tool for security measures in public and private settings.

It is recommended to research further in order to optimize the anomaly detection algorithm to reduce FNs, and integrate additional features, such as voice recognition and facial recognition, to create a more comprehensive security system. More datasets are necessary to train the model using real guns to train algorithms and avoid FPs.

Author Contributions: Conceptualization: A.L.E.R. and J.C.D.C.; methodology: A.L.E.R. and J.C.D.C.; software: A.L.E.R.; validation: A.L.E.R. and J.C.D.C.; formal analysis: A.L.E.R. and J.C.D.C.; investigation: A.L.E.R.; resources: A.L.E.R.; data curation: A.L.E.R.; writing—original draft preparation: A.L.E.R.; writing—review and editing: A.L.E.R. and J.C.D.C.; visualization: A.L.E.R. and J.C.D.C.; supervision: J.C.D.C.; project administration: A.L.E.R. and J.C.D.C.; funding acquisition: N/A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ankad, S.S.; Raj, B.S.; Nasreen, A.; Mathur, S.; Kumar, P.R.; Sreelakshmi, K. Object Size Measurement from CCTV footage using deep learning. In Proceedings of the 2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), Bangalore India, 16–18 December 2021; pp. 1–5. [\[CrossRef\]](#)
2. Germanov, A. How to Detect Objects in Images Using the YOLOv8 Neural Network. freeCodeCamp.org, 4 May 2023. Available online: <https://www.freecodecamp.org/news/how-to-detect-objects-in-images-using-yolov8/> (accessed on 13 October 2024).
3. Lim, J.; Al Jobayer, M.I.; Baskaran, V.M.; Lim, J.M.; Wong, K.; See, J. Gun detection in surveillance videos using Deep Neural Networks. In Proceedings of the 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Lanzhou, China, 18–21 November 2019; pp. 1998–2002. [\[CrossRef\]](#)
4. Hanopol, G.L.; Cruz, J.C.D. Design and Development of a Leaf Blight Detection System Using Image Processing with Intrusion Notification via SMS. In Proceedings of the 2023 IEEE 14th Control and System Graduate Research Colloquium (ICSGRC), Shah Alam, Malaysia, 5 August 2023; pp. 175–179. [\[CrossRef\]](#)
5. Babila, I.F.E.; Villasor, S.A.E.; Cruz, J.C.D. Object Detection for Inventory Stock Counting Using YOLOv5. In Proceedings of the 2022 IEEE 18th International Colloquium on Signal Processing & Applications (CSPA), Selangor, Malaysia, 12 May 2022; pp. 304–309. [\[CrossRef\]](#)
6. Afandi, W.E.I.B.W.N.; Isa, N.M. Object Detection: Harmful Weapons detection using YOLOv4. In Proceedings of the 2021 IEEE Symposium on Wireless Technology & Applications (ISWTA), Shah Alam, Malaysia, 17 August 2021; pp. 63–70. [\[CrossRef\]](#)
7. Investigation, F.B.O. Number of Robberies in the U.S. by Weapon 2018 | Statista. *Statista*. 2018. Available online: <https://www.statista.com/statistics/251914/number-of-robberies-in-the-us-by-weapon/> (accessed on 13 October 2024).
8. Llanes, L.A.C.; Ulbis, C.R.H.; Garcia, R.G. Remote controlled unmanned water vehicle with human detection and GPS Using Yolov4 for Flood Search Operations. In Proceedings of the 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 17–18 March 2023; pp. 373–379. [\[CrossRef\]](#)
9. Pullakandam, M.; Loya, K.; Salota, P.; Yanamala, R.M.R.; Javvaji, P.K. Weapon object detection using quantized YOLOv8. In Proceedings of the 2023 5th International Conference on Energy, Power and Environment: Towards Flexible Green Energy Technologies (ICEPE), Shillong, India, 15–17 June 2023; pp. 1–5. [\[CrossRef\]](#)
10. Saha, S. A Comprehensive Guide to Convolutional Neural Networks—The ELI5 Way. *Towards Data Science*, 15 December 2018. Available online: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53> (accessed on 8 November 2024).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.