Available online at :
http://ejournal.amikompurwokerto.ac.id/index.php/telematika/

Telematika

Accreditated SINTA "2" Kemendiktisaintek, No. 10/C/C3/DT.05.00/2025

Violence and Robbery Detection System Using YOLOv5 Algorithm Based on IoT Technology

Hani'atul Khoiriyah[1], Fauzan Abdillah[2], Afris Nurfal Aziz[3], I Gede Wiryawan[4]

[1, 2, 3] Teknik Informatika, Teknologi Informasi,
[4] Teknik Komputer, Teknologi Informasi,
[1, 2, 3, 4] Politeknik Negeri Jember, Jember, Indonesia
E-mail: haniatul122@gmail.com[1], fauzan.abdillah2705@gmail.com[2], afrisaziz@gmail.com[3], wiryawan@polije.ac.id[4]

ARTICLE INFO

History of the article:
Received January 15, 2025
Revised August 15, 2025
Accepted August 26, 2025

Keywords:
Violence
Robbery
Detection
YOLOv5
IoT

Correspondence:
E-mail: haniatul122@gmail.com

ABSTRACT

Violence and robbery are two common forms of crime that often cause material losses, psychological trauma, and insecurity within society. Conventional CCTV systems are limited in preventing such incidents, which highlights the need for more intelligent and responsive security solutions. The primary objective of this research is to design and evaluate SmartGuard, a real-time detection system for violence and robbery based on artificial intelligence (AI) using the YOLOv5 algorithm, integrated with Internet of Things (IoT) technology for remote monitoring. This study employed an experimental design with several stages: dataset preparation, model training, testing, model analysis, and system integration with Raspberry Pi, Firebase, and a mobile application. The dataset consisted of 6,900 labeled images across three classes: violence, robbery, and normal activity. Model evaluation was conducted using a separate test dataset and analyzed with a confusion matrix. The results show that the model achieved an overall accuracy of 70.94%. The system performed relatively well in detecting violence, with a precision of 71.13% and an F1-score of 62.47%. However, recall values for robbery (47.53%) and normal activity (48.99%) were considerably lower, indicating challenges in consistently recognizing these classes. Despite these limitations, SmartGuard allows users to view and receive notifications in emergency situations, enabling them to take quick action and monitor the situation effectively.

INTRODUCTION

Violence and robbery are among the most pressing criminal acts that frequently occur in both urban and rural areas. These crimes not only cause financial losses but also inflict long-term psychological trauma and generate insecurity within society. According to the 2024 Crime Statistics published by Badan Pusat Statistik (BPS), these include violent theft which increased from 4,335 cases in 2022 to 6,573 cases in 2023, domestic violence from 5,526 to 10,783 cases, assault from 34,452 to 51,106 cases, and mob violence from 8,239 to 16,441 cases (Badan Pusat Statistik, 2024). Cases such as the viral CCTV recording of child abuse further highlight the urgency of developing more effective preventive measures.

Although CCTV is widely used as a monitoring tool, it primarily serves as a passive recorder. Experts emphasize that it is inadequate for preventing or responding to spontaneous violent behavior, and issues such as delayed review and potential manipulation weaken its reliability (Houser et al., 2024). A

main issue identified in this study is the lack of a smart real-time system capable of detecting violent acts or crimes as they occur. This study limits its scope to two main categories of crime: violence and robbery, while also considering normal activities as a control class.

To address this limitation, an intelligent solution is needed that can actively detect threats and provide timely alerts. One promising approach is the use of artificial intelligence, particularly the YOLOv5 algorithm, which is recognized for its speed and accuracy in object detection tasks (Boukabous and Azizi, 2023; Bushra et al., 2022). Integrating this model with Internet of Things (IoT) technology allows real-time monitoring and mobile-based notifications, enabling users to take immediate action when suspicious activity is detected (Ganesan et al., 2022). Previous studies have shown the effectiveness of AI-based surveillance for detecting weapons and violent acts, but few have combined it with IoT integration for real-time community-level monitoring (Kumar et al., 2024; Razzaq et al., 2023).

Therefore, this study proposes SmartGuard which is a real-time violence and robbery detection system that combines the YOLOv5 algorithm with IoT and mobile applications. The objectives of this study are to design, implement, and evaluate the system's effectiveness in detecting violent and robbery-related activities. The significance of this work lies in providing a proactive framework that not only demonstrates technical feasibility but also contributes to improving public safety by enabling early detection, real-time alerts, and rapid response.

RESEARCH METHODS

The system development involves stages including dataset collection, dataset training, model testing, model analysis, Firebase integration, Raspberry Pi integration, and mobile integration, with a technical design that incorporates YOLOv5 for real-time object detection, Firebase for data management, and Raspberry Pi for real-time processing. The test data consisted of simulated videos with three scenarios violence, robbery, and normal activity each designed to evaluate the system's ability to detect criminal events in real time. A mobile application developed using Flutter and integrated with Firebase Authentication, Firestore, and Messaging, serves as a user interface for monitoring and analyzing detection data. System performance was assessed using three main parameters: detection success, detection accuracy, and notification display on the mobile application. The observations from these parameters served as the basis for analyzing system performance, which is further elaborated in the results and discussion section. Figure 1 presents the methodological framework underlying this study, which will be explained in detail in the following section.

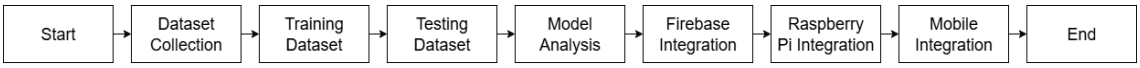

Figure 1. Methodology

1. System Components

The system utilizes a webcam as an alternative to CCTV for the primary data source, integrating it into the detection system for real-time monitoring. A Raspberry Pi 4 is utilized as the processing unit to manage data captured by the camera, leveraging the YOLOv5 algorithm for object detection (Aziz et al., 2024; Harsh, 2025). The Raspberry Pi 4 is chosen for its advantages, including affordability, ease of use, compatibility with various devices, and energy efficiency (Karthikeyan et al., 2023; Palomino et al., 2022). To maintain stable operation, a power adapter provides consistent

electrical power to the Raspberry Pi 4 and other connected devices. Connectivity is established through an Ethernet LAN cable, allowing the system to interface with a local network or the internet for data transfer and remote access. The integration with Firebase enables the detection system to securely store data in a cloud-based database, which can be accessed in real-time from multiple connected devices, ensuring efficient synchronization and user authentication (Ahmed et al., 2020). The development process utilizes computers and software tools such as Roboflow Universe for dataset management, Google Colab for model training, and Visual Studio Code for coding, thereby providing comprehensive support throughout the system's development. Figure 2 depicts the architecture of the proposed system, detailed in the subsequent explanation:

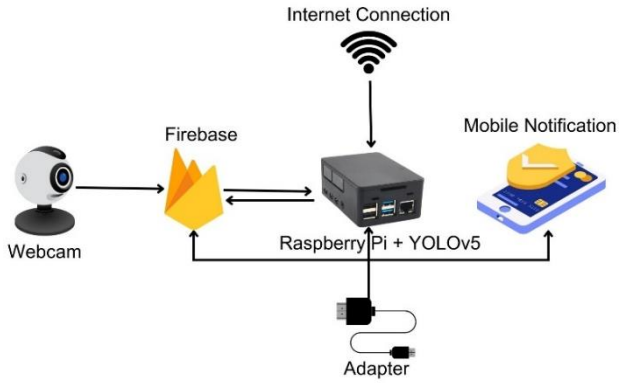

Figure 2. System Architecture

a. Recording and Monitoring

Webcams capture images from the monitored area in real-time. Then the visual data is stored in Firebase and sent to Raspberry Pi 4 for analysis.

b. Activity Detection

Raspberry Pi 4 uses YOLOv5 to analyze visual data. When violent or robbery activity is detected, Raspberry Pi 4 sends event data in the form of images, time, and type of activity to Firebase.

c. Data Management and Storage

Firebase receives the detection results back and stores the event data. Activity logs and details of violent and robbery events are stored and can be accessed through the mobile application.

d. Notification and Response

Firebase sends real-time notifications to registered application users when violent activity or robbery is detected. Users receive notifications in the SmartGuard mobile application and can take quick action.

e. User Access and Event Details

Users can view activity logs and event details that include event images, activity types, detection confidence levels, and time of occurrence.

2. YOLOv5 Algorithm

Artificial Intelligence is used to identify and classify various objects in videos in real-time. YOLOv5 offers an efficient, fast, and accurate real-time object detection solution (Alahdal et al., 2024; Varun and Bhuvanesh, 2023). YOLOv5 was selected because previous studies in the domain of crop monitoring have demonstrated that this model maintains accuracy comparable to YOLOv8 and YOLO-NAS while offering higher detection speed and better computational efficiency than YOLOv6, characteristics that make it relevant for real-time applications (Nnadozie et al., 2024).

YOLO is an approach in image processing that utilizes the entire neural network to analyze the whole image simultaneously. Compared to other detection models, YOLOv5 has several advantages. This method is able to consider the global context of the entire image when making predictions, as it views the image as a whole during testing (He et al., 2025). By processing the entire image in a single pass, YOLO outperforms two-stage detectors like R-CNN in terms of speed, while maintaining competitive accuracy for real-time applications (Reswara et al., 2023). Initially this model was developed within the Darknet framework, but YOLOv5 is the first in the YOLO family to use the PyTorch framework (Kisaezehra et al., 2023). This change makes it lighter and easier to use. However, YOLOv5 does not experience significant architectural changes from YOLOv4 and does not substantially outperform YOLOv4 on standard benchmarks (Murat and Kiran, 2025). YOLOv5 can recognize certain patterns that indicate violence or robbery (Abdillah et al., 2024). Here's how YOLOv5 works:

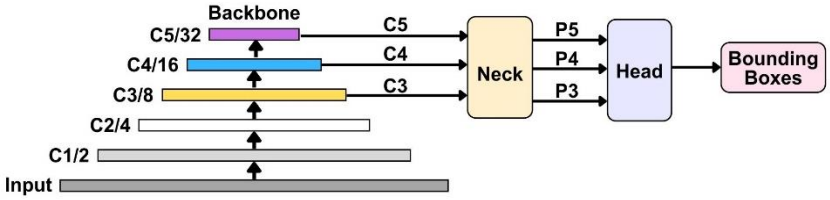

Figure 3. YOLOv5 works

The process as shown in Figure 3 begins by inputting the image into the input layer for further processing. After the image is received, it is forwarded to the backbone, which is responsible for extracting features through several layers: C1/2, C2/4, C3/8, C4/16, and C5/32. Each of these layers applies convolution, reducing the resolution size at each stage while increasing the feature depth, resulting in feature maps that represent object characteristics at various scales. Following feature extraction, the different scale feature maps (C3, C4, C5) are combined in the neck using a Feature Fusion Network, such as FPN or PANet, to produce three main feature maps: P3, P4, and P5, which are utilized for detecting small, medium, and large objects, respectively (Kang et al., 2024; Liu et al., 2022). The processed feature map from the neck is then sent to the prediction head, where calculations for confidence and bounding box regression are performed for each pixel. This includes generating bounding box coordinates, box sizes, and confidence scores regarding the accuracy of the boxes and the object classes. After the prediction is complete, bounding boxes indicating the locations of objects in the image are generated, and Non-Maximum Suppression (NMS) is employed to filter out irrelevant or overlapping boxes, ensuring that only the bounding box with the highest precision is selected (Liu et al., 2022).

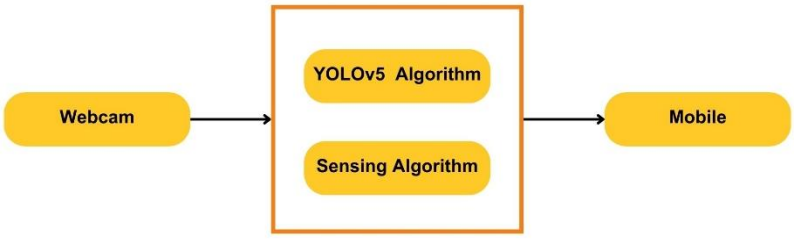

Figure 4. YOLOv5 Algorithm

Figure 4 depicts the YOLOv5 algorithm applied within the system framework (Sung and Park, 2021). In this security system, the model is trained to detect human activity, violent behavior, and robbery. Using deep learning algorithms allows the model to continuously learn and improve its

accuracy from new data collected. Using YOLOv5 algorithm, the system can process data in real time and make decisions quickly and accurately (Khalfaoui et al., 2024).

3. Embedded Systems and Internet of Things (IoT)

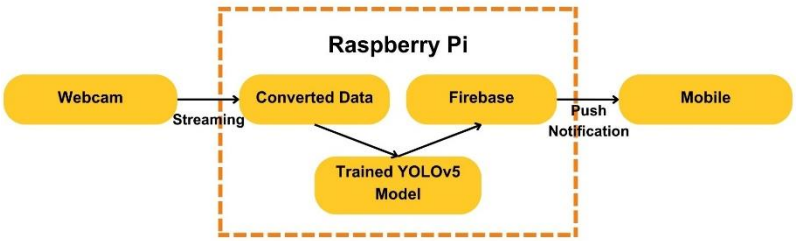

Figure 5. System Integration

Figure 5 illustrates the system integration based on Raspberry Pi 4, which serves as the central data processing unit. This device is equipped with the YOLOv5 model and Firebase, enabling the reception of video data from the camera, image processing for object detection, as well as real-time data storage and synchronization. Firebase, as a cloud-based real-time database with JSON format, allows data to be securely stored and instantly accessed by multiple devices (Prasetyawan et al., 2021). The Raspberry Pi 4 is connected to the camera to capture images as a data source and continuously monitor and detect normal, violent, or robbery activities. The system provides rapid notifications for situations requiring immediate action (Jayasakthi et al., 2023).

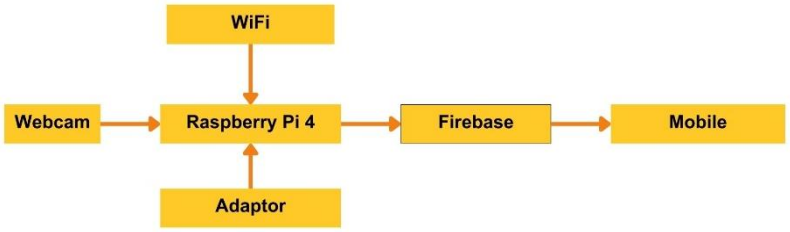

Figure 6. System Configuration

Figure 6 presents the IoT based system configuration, connecting the camera, YOLOv5 model, Firebase integration, Raspberry Pi 4, and a mobile application via Wi-Fi. The system starts by connecting the Raspberry Pi 4 to the network through a Wi-Fi connection. Firebase integration is used to link the data captured by the camera, which is then processed by Raspberry Pi 4 using the YOLOv5 model for violence and robbery detection. This setup enables fast communication between the Raspberry Pi 4 and the central server, supports video data transfer to the detection model, and returns the detection results, subsequently sending real-time notifications to the mobile application, allowing users to monitor the system remotely while connected to the internet (Desnanjaya and Arsana, 2021).

RESULTS AND DISCUSSION

1. Dataset Collection

The dataset used comes from the Roboflow Universe platform, which includes 6,900 images with labels for three classes: "violence" (3,000 images), "robbery" (3,000 images), and "normal" (900 images). The dataset is divided as follows: 5,902 images (86%) for the train set, 853 images (12%) for the valid set, and 145 images (2%) for the test set, this image division can be seen in Figure 7. After that, the dataset is compressed in zip format and uploaded to Google Drive for the model training process.

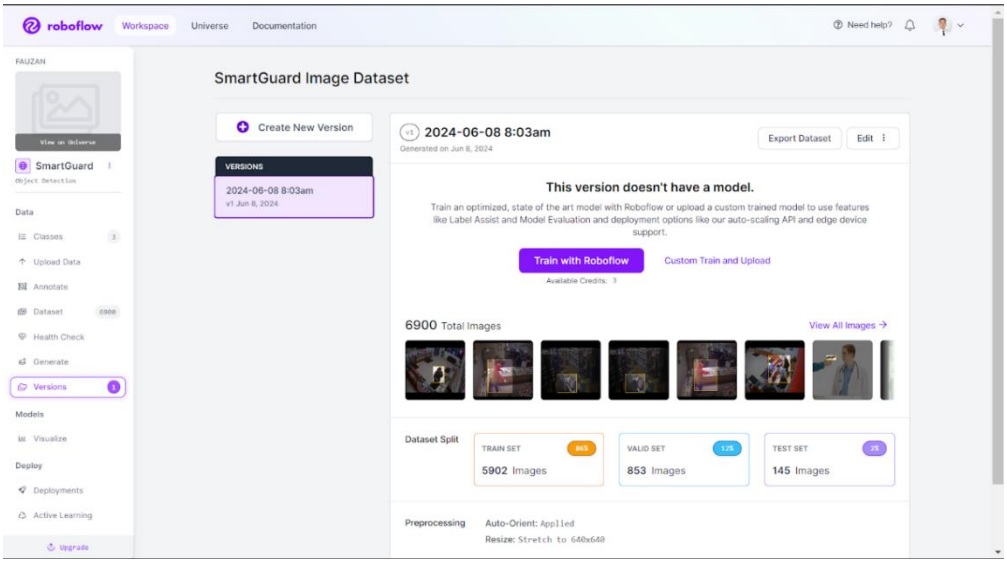

Figure 7. Dataset

Then, another dataset of 1,490 images has been prepared for the needs of model analysis. This division is done by allocating 497 images for the "violence" class, 497 images for the "robbery" class, and 496 images for the "normal" class. The purpose of this division is to ensure that each class has a balanced representation in further analysis, so that the model can be evaluated fairly and comprehensively. Each subset of this dataset is used to test the performance of the model in detecting and classifying objects according to their class, as well as to analyze performance metrics such as accuracy, recall, and precision. With this approach, it is expected that the model can learn and generalize well on new data that has not been seen before.

2. Training Dataset

After the dataset training process is complete, output files are generated to support the analysis and evaluation of the model trained using YOLOv5, as shown in Figure 8. First, the “labels” file provides information about the labels assigned to each image in the dataset, helping in the validation and analysis of the model results. Next, the “labels_correlogram” file presents a visualization of the correlation between labels, helping in understanding the relationship and distribution between classes in the dataset, which is crucial for knowing how the model classifies objects.

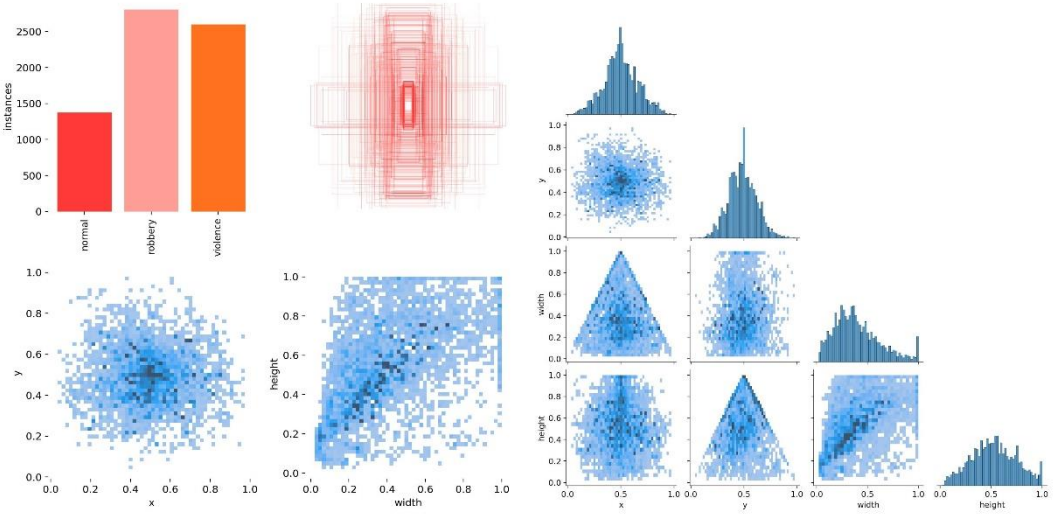

Figure 8. Labels Training and Labels_correlogram Training

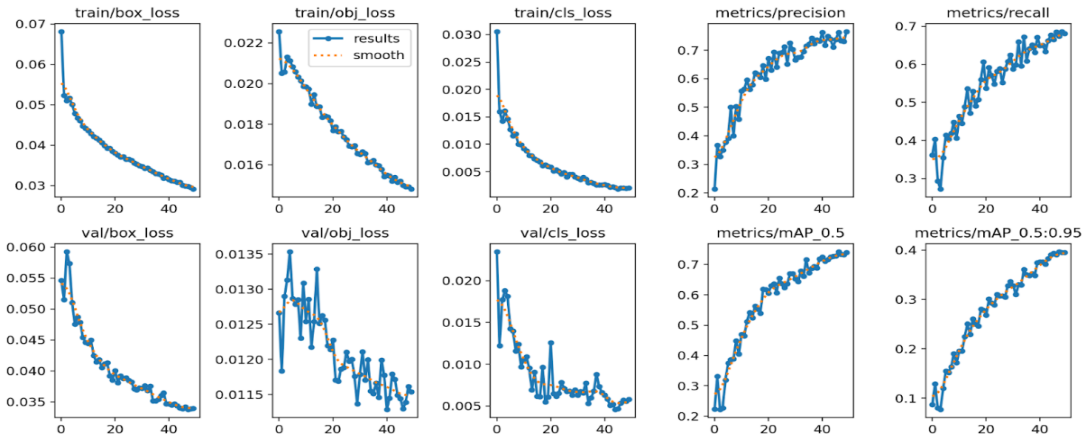

Figure 9. Training Result

After the training process is complete, evaluation is carried out using metrics such as recall, precision, and other indicators as shown in Figure 9. Recall provides an overview of how well the model identifies real criminal objects, while precision measures how accurately the model identifies the objects. By performing fifty rounds of training (epochs), the goal is to achieve the best balance between the model's ability to detect criminal objects and the detection error rate. The results of this training produce an output file with the extension .pt, which is very important for integrating the criminal detection model into devices such as the Raspberry Pi 4. The file contains parameters that have been adjusted during the training process, including weights, which are the core of the detection model.

3. Testing Model

The test results on a dataset consisting of a total of 1,490 images show variations in the model's performance for each class. From 496 images in the "normal" dataset, the model successfully detected 324 images. Of the detected images, 243 were classified as "normal," while 81 images were misclassified, as shown in figure a. In the "violence" dataset, which consisted of 497 images, the model successfully detected 378 images. Of the detected images, 277 were classified as "violence," while 101 images were misclassified, as shown in figure b. As for the "robbery" dataset, also consisting of 497 images, the model successfully detected 355 images. Of the detected images, 236 were classified as "robbery," while 119 images were misclassified as shown in figure 10. Overall, out of the total 1,490 images tested, the model successfully detected 1,057 images, and 433 images were not detected. These results provide an overview of the model's ability to identify and classify images into the correct classes, as well as areas that require further improvement.

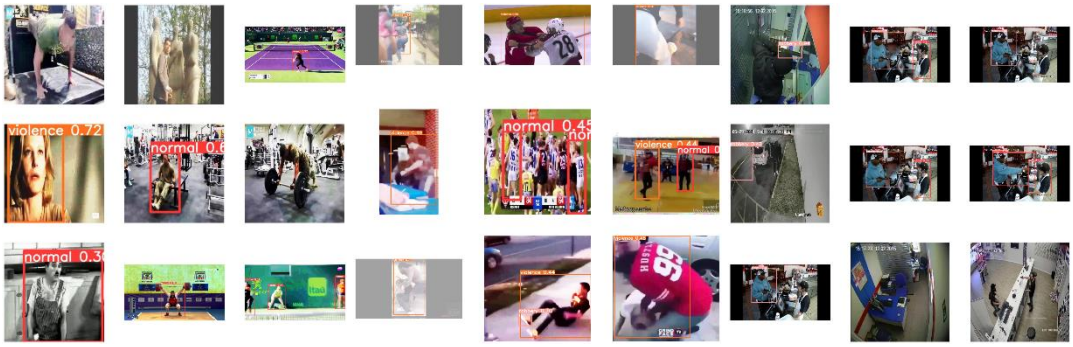

Figure 10. Testing Normal Dataset, Testing Violence Dataset, and Testing Robbery Dataset

4. Model Analysis

Based on the results of the previous tests, the following is the Confusion Matrix in Table 1 which details the results of testing the detection system using a dataset consisting of 993 images.

Table 1. Confusion Matrix of Model Testing Results

		Actual Class		
		Violence	*Robbery*	*Normal*
Predicted Class	*Violence*	**277 (TP)**	9 (FP)	49 (FP)
	Robbery	6 (FP)	**236 (TP)**	32 (FP)
	Normal	95 (FN)	110 (FN)	**243 (TP)**

The table above shows the results of the detection system's performance evaluation in classifying images into three categories: "Violence," "Robbery," and "Normal." In this evaluation, several possible classification outcomes were observed, including True Positive (TP), False Positive (FP), and False Negative (FN). From the evaluation results, it was found that out of the total images that were actually violent, 277 images were correctly predicted as violent by the system (TP), while 95 images that were actually violent were misclassified as normal (FN). Out of the total images that were actually robberies, 236 images were correctly predicted as robberies by the system (TP), while 110 images that were actually robberies were misclassified as normal (FN). On the other hand, of the images that were actually normal, 243 images were correctly predicted as normal by the system (TP), but there were several errors where 49 images that were actually normal were incorrectly predicted as violent (FP), and 32 images that were actually normal were incorrectly predicted as robbery (FP). Other prediction errors included 9 robbery images that were misclassified as violent (FP) and 6 violent images that were misclassified as robbery (FP). This confusion matrix provides a clear picture of the model's ability to classify images into the correct categories, as well as areas that need further improvement, such as reducing false positives and false negatives.

Table 2. Model Performance

Category	*Accuracy*	*Precision*	*Recall*	*F1-Score*
All Class	0.7094	0.6672	0.7094	0.6876
Violence	0.7094	0.7113	0.5571	0.6247
Robbery	0.7094	0.6718	0.4753	0.5556
Normal	0.7094	0.6607	0.4899	0.5614

The confusion matrix was analyzed based on the accuracy, precision, recall, and F1-score values listed in Table 2. From these calculations, the model's performance in distinguishing between "violence," "robbery," and "normal" images can be observed. The model showed an accuracy of 70.94% across all classes, reflecting its ability to classify the data correctly.

The precision for the "violence," "robbery," and "normal" classes is 71.13%, 67.18%, and 66.07%, respectively, indicating the proportion of correctly classified images by the model. The recall for the "violence" class is 55.71%, for the "robbery" class is 47.53%, and for the "normal" class is 48.99%, indicating the proportion of positive data that the model successfully identified. The F1-score for the "violence," "robbery," and "normal" classes is 62.47%, 55.56%, and 56.14%, respectively, representing the harmonic mean of precision and recall for the three classes. Thus, the model shows better results in recognizing "violence" images compared to "robbery" and "normal" images, but there

is still room for improvement in increasing recall and reducing false positives and false negatives for all classes.

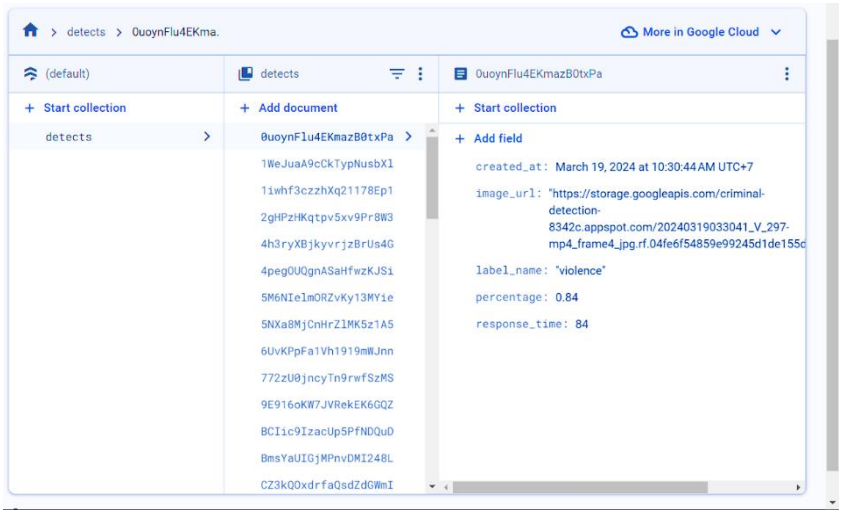

Figure 11. Firebase integration

5. Firebase Integration

As shown in Figure 11, Firebase integration enables the Raspberry Pi 4 to communicate effectively with the detection model, the device can send data from the webcam to the detection model and receive the detection results back. Firebase, as a real-time database hosted in the cloud, stores data in JSON format and synchronizes it instantly with the SmartGuard device. The communication facilitated by Firebase ensures that detection results are sent quickly and on time.

6. Raspberry Pi Integration

The Raspberry Pi 4 serves as the central processing unit, managing data captured by the camera through the implementation of the YOLOv5 algorithm embedded within the Raspberry Pi 4, as illustrated in Figure 12. System connectivity is established via an Ethernet cable, linking the Raspberry Pi 4 to the local network or the internet to ensure smooth data transfer and remote access capability. With the integration of Raspberry Pi 4, the violence and robbery detection system can perform real-time object detection on webcam data and transmit the detection results to the SmartGuard application.

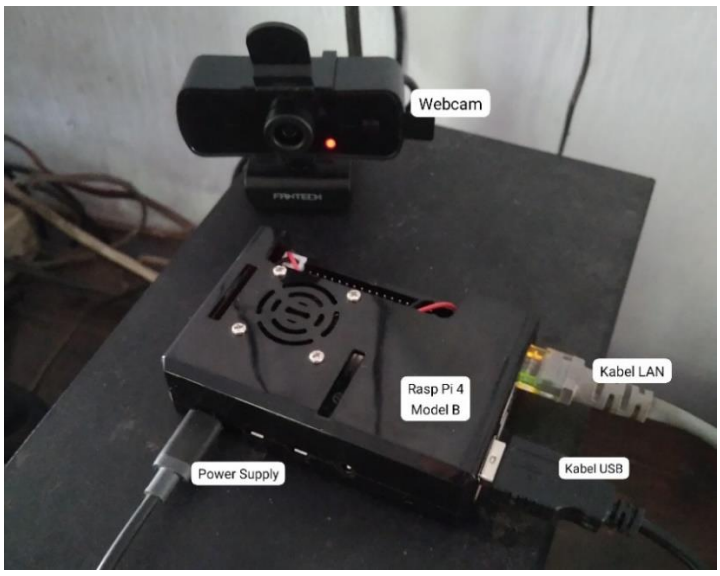

Figure 12. Raspberry Pi integration

7. Mobile App Integration

The SmartGuard mobile application has been developed using the Flutter framework, which successfully integrates with Firebase Authentication, Firestore, and Messaging. The data processed using the YOLOv5 model produces final output displayed on the mobile application. The running system will send incident data to the server via Raspberry Pi 4, allowing users to monitor and respond to acts of violence or robbery in real-time. Mobile integration makes the violence and robbery detection system more accessible and flexible, usable from anywhere and at any time. Using this application, users can view and receive notifications in case of emergencies in real-time. The following section outlines the workflow of the SmartGuard mobile application for the violence and robbery detection system, as depicted in the Figure 13:

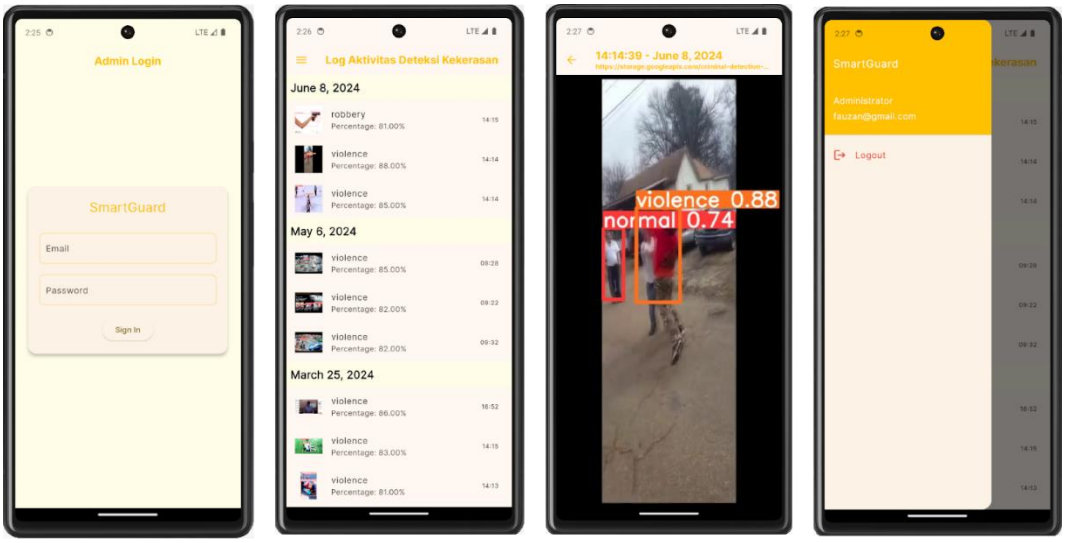

Figure 13. SmartGuard mobile application

a. User Login

Users start by logging into the SmartGuard application using valid credentials, which include an email and password. This process allows users to access various features of the SmartGuard application, such as the activity log and details of each incident.

b. Violence and Robbery Activity Log

After successfully logging into the SmartGuard application, users will be presented with a screen displaying the violence and robbery activity log. This log contains a list of recorded violence incidents, along with important information such as incident images, the type of criminal activity detected, the accuracy level of the detection, and the time of the incident.

c. Incident Details

Users who select an incident from the violence or robbery activity log will be taken to the details feature. In this feature, users can view more detailed information about the incident, such as images captured by the system, the type of criminal activity detected, the accuracy level of the detection, and the specific time of the incident, including the date and hour.

8. System Evaluation

The system evaluation was conducted to measure the performance of SmartGuard when integrated with IoT devices, specifically a Raspberry Pi 4 connected to a webcam for visual data acquisition and Firebase for notification delivery. The experiment was carried out using simulated videos containing

three scenarios: violence, robbery, and normal activity. The observed parameters included detection success, detection accuracy, and notification display on the mobile application.

The experimental results show that the system generates notifications only when a criminal event is detected with a minimum accuracy of 80%. When this criterion is met, the notification successfully appears in the mobile application and is recorded in the activity log. Detection of violent incidents more frequently reached the accuracy threshold compared to robbery and normal activity, which is consistent with previous model evaluation results. These findings confirm that SmartGuard is capable of performing IoT-based real-time detection and delivering automated notifications to assist users in responding to critical situations.

In addition, the integration of Raspberry Pi 4 with the YOLOv5 model has been successfully implemented, enabling real-time object detection. Integration with Firebase has also proven effective for storing processed data and providing real-time access to connected devices. The processed outputs are displayed through the SmartGuard mobile application, demonstrating the effectiveness of the system in delivering timely notifications and supporting rapid responses in emergency conditions.

Despite these promising results, SmartGuard still has several limitations. The main challenge lies in the relatively small and less diverse dataset. Future research should incorporate more varied data, including urban and rural environments, indoor and outdoor settings, daytime and nighttime conditions, as well as different weather scenarios. Moreover, increasing the number of training epochs may improve model accuracy, although it requires higher-specification hardware. Future studies may also consider advanced techniques such as ensemble learning to reduce detection errors without compromising processing speed.

CONCLUSIONS AND RECOMMENDATIONS

This research was conducted to detect criminal acts categorized into two classes, namely violence and robbery. The system successfully demonstrated the effectiveness of the YOLOv5-based violence and robbery detection system with an accuracy rate of 70.94% from the results of the confusion matrix analysis. Although the system showed good performance in detecting the precision of the "violence" class of 71.13%, for the "robbery" class of 67.18%, and "normal" of 66.07% still needs to be improved. Recall for each class is 55.71%, 47.53%, and 48.99%. The F1 score for each class is 62.47%, 55.56%, and 56.14%. Integration with IoT and Firebase technology is carried out for real-time monitoring and providing automatic notifications to the user's smartphone, which has the potential to increase security with a quick response to violent or robbery situations.

It is necessary to add and vary the dataset to improve model performance, especially for robbery and normal activity classes. With a more diverse and representative dataset, the system will be better able to cope with visual variations of real-world events. For further research, it can also focus on reducing false positives and false negatives. Techniques such as fine-tuning the model or using an ensemble learning approach can be implemented to improve detection accuracy without sacrificing speed.

ACKNOWLEDGEMENT

We extend our sincere gratitude to the Directorate of Learning and Student Affairs, Ministry of Education, Culture, Research, and Technology, for funding this project through the Program Kreativitas Mahasiswa - Karsa Cipta 2023 with decree number 0598/D4/KM.01.00/2023.

REFERENCES

Abdillah, F., Khoiriyah, H., Aziz, A.N., Wiryawan, I.G., 2024. Sistem Deteksi Kekerasan Real-Time menggunakan YOLOv5 untuk Keamanan Publik. Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika 198–204. https://doi.org/10.31284/p.snestik.2024.5861

Ahmed, T., Bin Nuruddin, A.T., Latif, A. Bin, Arnob, S.S., Rahman, R., 2020. A Real-Time Controlled Closed Loop IoT Based Home Surveillance System for Android using Firebase. 2020 6th International Conference on Control, Automation and Robotics, ICCAR 2020 601–606. https://doi.org/10.1109/ICCAR49639.2020.9108016

Alahdal, N.M., Abukhodair, F., Meftah, L.H., Cherif, A., 2024. Real-time Object Detection in Autonomous Vehicles with YOLO. Procedia Comput Sci 246, 2792–2801. https://doi.org/10.1016/j.procs.2024.09.392

Aziz, A.N., Khoiriyah, H., Abdillah, F., Wiryawan, I.G., 2024. Prototipe Sederhana Sistem Deteksi Kriminal Berbasis Internet of Things Menggunakan Teknologi YOLOv5. Komputika : Jurnal Sistem Komputer 13, 139–147. https://doi.org/10.34010/komputika.v13i1.12217

Badan Pusat Statistik, 2024. Statistik Kriminal Tahun 2024. Vol. 42. Jakarta: Badan Pusat Statistik. ISSN 2089-5291.

Boukabous, M., Azizi, M., 2023. Image and video-based crime prediction using object detection and deep learning. Bulletin of Electrical Engineering and Informatics 12, 1630–1638. https://doi.org/10.11591/eei.v12i3.5157

Bushra, S.N., Shobana, G., Maheswari, U.K., Subramanian, N., 2022. Smart Video Survillance Based Weapon Identification Using Yolov5, in: Proceedings of the 2022 International Conference on Electronic Systems and Intelligent Computing, ICESIC 2022. Institute of Electrical and Electronics Engineers Inc., pp. 351–357. https://doi.org/10.1109/ICESIC53714.2022.9783499

Desnanjaya, I.G.M.N., Arsana, I.N.A., 2021. Home security monitoring system with IoT-based Raspberry Pi. Indonesian Journal of Electrical Engineering and Computer Science 22, 1295–1302. https://doi.org/10.11591/ijeecs.v22.i3.pp1295-1302

Ganesan, S., Yin Ying, T., Ravi, P., Peng Lean, C., 2022. Designing an Autonomous Triggering Control System via Motion Detection for IoT Based Smart Home Surveillance CCTV Camera. Malaysian Journal of Science and Advanced Technology 2, 80–88. https://doi.org/10.56532/mjsat.v2iS1.120

Harsh, 2025. AI-Powered CCTV Surveillance with Intrusion Detection Using YOLOv5 and Raspberry Pi. International Journal of Computer Techniques 12, 209–212.

He, Z., Wang, K., Fang, T., Su, L., Chen, R., Fei, X., 2025. Comprehensive Performance Evaluation of YOLOv5 on Object Detection of Power Equipment. 2025 37th Chinese Control and Decision Conference (CCDC) 1281–1286. https://doi.org/10.1109/CCDC65474.2025.11090973

Houser, T.E., McMillan, A., Dong, B., 2024. Bridging the gap between criminology and computer vision: A multidisciplinary approach to curb gun violence. Security Journal 37, 1409–1429. https://doi.org/10.1057/s41284-024-00423-7

Jayasakthi, B.G., Varunika, S., Kinzsy Grace, R., Ezhilin Freeda, S., 2023. IoT Based Security Alert System for Children. 7th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2023 - Proceedings 1632–1636. https://doi.org/10.1109/ICECA58529.2023.10395028

Kang, M., Ting, C.M., Ting, F.F., Phan, R.C.W., 2024. ASF-YOLO: A novel YOLO model with attentional scale sequence fusion for cell instance segmentation. Image Vis Comput 147, 1–9. https://doi.org/10.1016/j.imavis.2024.105057

Karthikeyan, S., Aakash Raj, R., Cruz, M.V., Chen, L., Ajay Vishal, J.L., Rohith, V.S., 2023. A Systematic Analysis on Raspberry Pi Prototyping: Uses, Challenges, Benefits, and Drawbacks. IEEE Internet Things J 10, 14397–14417. https://doi.org/10.1109/JIOT.2023.3262942

Khalfaoui, A., Badri, A., Mourabit, I. El, 2024. A lightweight you only look once for real-time dangerous weapons detection. IAES International Journal of Artificial Intelligence 13, 1836–1842. https://doi.org/10.11591/ijai.v13.i2.pp1838-1844

Kisaezehra, Farooq, M.U., Bhutto, M.A., Kazi, A.K., 2023. Real-Time Safety Helmet Detection Using Yolov5 at Construction Sites. Intelligent Automation and Soft Computing 36, 911–927. https://doi.org/10.32604/iasc.2023.031359

Kumar, P., Shih, G.L., Guo, B.L., Nagi, S.K., Manie, Y.C., Yao, C.K., Arockiyadoss, M.A., Peng, P.C., 2024. Enhancing Smart City Safety and Utilizing AI Expert Systems for Violence Detection. Future Internet 16. https://doi.org/10.3390/fi16020050

Liu, H., Sun, F., Gu, J., Deng, L., 2022. SF-YOLOv5: A Lightweight Small Object Detection Algorithm Based on Improved Feature Fusion Mode. Sensors 22, 1–14. https://doi.org/10.3390/s22155817

Murat, A.A., Kiran, M.S., 2025. A comprehensive review on YOLO versions for object detection. Engineering Science and Technology, an International Journal 70. https://doi.org/10.1016/j.jestch.2025.102161

Nnadozie, E.C., Casaseca-de-la-Higuera, P., Iloanusi, O., Ani, O., Alberola-López, C., 2024. Simplifying YOLOv5 for deployment in a real crop monitoring setting. Multimed Tools Appl 83, 50197–50223. https://doi.org/10.1007/s11042-023-17435-x

Palomino, X.P., Paredes, K.R., Tejada, J.E., 2022. Low-Cost Gas Leak Detection and Surveillance System for Single Family Homes Using Wit.ai, Raspberry Pi and Arduino. International Journal of Interactive Mobile Technologies. https://doi.org/10.3991/ijim.v16i09.30177

Prasetyawan, P., Samsugi, S., Prabowo, R., 2021. Internet of Thing Menggunakan Firebase dan Nodemcu untuk Helm Pintar. Jurnal ELTIKOM 5, 32–39. https://doi.org/10.31961/eltikom.v5i1.239

Razzaq, F.A., Chaudary, M.A., Fareed, S., Tariq, W., Waqas, M., Javaid, S., 2023. Enhancing Public Safety: Detection of Weapons and Violence in CCTV Videos with Deep Learning, in: 2023 25th International Multi Topic Conference, INMIC 2023 - Proceedings. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/INMIC60434.2023.10465800

Reswara, E., Suakanto, S., Putra, S.A., 2023. Comparison of Object Detection Algorithm using YOLO vs Faster R-CNN : A Systematic Literature Review. ICBDT '23: Proceedings of the 2023 6th International Conference on Big Data Technologies 419–424. https://doi.org/https://doi.org/10.1145/3627377.3627443

Sung, C.S., Park, J.Y., 2021. Design of an intelligent video surveillance system for crime prevention: applying deep learning technology. Multimed Tools Appl 80, 34297–34309. https://doi.org/10.1007/s11042-021-10809-z

Varun, S., Bhuvanesh, V.M., 2023. Real Time Theft Detection Using YOLOv5 Object Detection Model, in: Proceedings - 2023 3rd International Conference on Innovative Sustainable Computational Technologies, CISCT 2023. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/CISCT57197.2023.10351223