Real-Time Weapon Detection and Automated Incident Reporting Using CCTV for Minimarket Security: A Prototype Study

line 1: 1st Mhd Ulil Abshar
line 2: *dept. Information System*
(of Affiliation)
line 3: *Universiras Andalas*
(of Affiliation)
line 4: Padang, West Sumatera
line 5:
2211521003_ulil@student.unand.ac.id

line 1: 2nd Given Name Surname
line 2: *dept. name of organization*
(of Affiliation)
line 3: *name of organization*
(of Affiliation)
line 4: City, Country
line 5: email address or ORCID

line 1: 3rd Given Name Surname
line 2: *dept. name of organization*
(of Affiliation)
line 3: *name of organization*
(of Affiliation)
line 4: City, Country
line 5: email address or ORCID

line 1: 4th Given Name Surname
line 2: *dept. name of organization*
(of Affiliation)
line 3: *name of organization*
(of Affiliation)
line 4: City, Country
line 5: email address or ORCID

line 1: 5th Given Name Surname
line 2: *dept. name of organization*
(of Affiliation)
line 3: *name of organization*
(of Affiliation)
line 4: City, Country
line 5: email address or ORCID

line 1: 6th Given Name Surname
line 2: *dept. name of organization*
(of Affiliation)
line 3: *name of organization*
(of Affiliation)
line 4: City, Country
line 5: email address or ORCID

***Abstract*— Recent armed robberies targeting minimarkets have highlighted critical security vulnerabilities in Indonesia's retail sector. This study presents TangkApIn, a real-time weapon detection system that integrates handheld weapon identification with automated incident reporting capabilities. The system employs a convolutional neural network (CNN) architecture inspired by the YOLO framework to process single CCTV feeds and detect weapons with 90% precision, 88% recall, and 92% mAP@0.5. Upon weapon detection, TangkApIn automatically generates structured incident reports containing annotated image evidence, metadata, and location information, delivering alerts to law enforcement and minimarket owners within 2-3 seconds. Performance evaluation in simulated minimarket environments demonstrates consistent detection latency under one second and F1 scores of 0.88, positioning the system as an effective solution for enhancing retail security through automated threat detection and rapid response coordination.**

Keywords*— weapon detection, computer vision, CCTV surveillance, real-time processing, security

I. INTRODUCTION (*HEADING 1*)

Indonesia's retail environments have experienced a concerning surge in criminal activities. Recent crime statistics from Indonesia's Central Bureau of Statistics indicate that criminal incidents during the year have reached alarming figures, demonstrating increased vulnerability in retail environments [1]. Historical records further reveal that 24-hour minimarkets have been repeatedly targeted; for instance, in 2012, Jakarta Metropolitan Police documented 32 armed robberies involving various weapons such as firearms and edged weapons. These incidents highlight the operational vulnerabilities present due to extended operating hours, minimal security staffing, and prominent cash-handling procedures. Traditional surveillance systems relying on manual monitoring have shown significant limitations, as human operators are unable to effectively process multiple simultaneous video feeds. This limitation can delay threat recognition and hinder rapid emergency response, especially during weapon-related incidents where every second is critical Bhatti et al. [2]..

Deep learning has emerged as a powerful tool to bridge this gap. Recent studies employing YOLO-based architectures have achieved detection accuracies exceeding 90% and have demonstrated real-time processing capabilities essential for security applications Bhatti et al. [2], Tahir [3]. Despite these advances, most research has focused solely on the detection component, with limited emphasis on integrating detection with comprehensive, automated incident reporting and notification systems. In order to address these shortcomings, the proposed TangkApIn system uniquely integrates real-time weapon detection with an automated pipeline that generates detailed incident reports and disseminates alerts to relevant stakeholders within seconds Bhatti et al. [2], Tahir [3], Al‑Mousa et al. [4]. The primary contributions of this research include :

1) the development of a real-time weapon detection prototype optimized for single CCTV feeds in minimarket environments,

2) the design and implementation of an automated incident reporting pipeline that produces structured reports enriched with visual evidence and metadata, and

3) an extensive evaluation of system performance—including metrics such as detection accuracy, response latency, and reporting effectiveness—conducted in simulated retail scenarios SJ et al. [5].

II. BACKGROUND

A. Security Challenges in Indonesian Retail Environments

Socioeconomic factors have been strongly correlated with increased criminal activity in Indonesia. Research by Nisa et al. [1] demonstrates that periods of heightened poverty directly contribute to surges in criminal incidents. These conditions, in combination with the 24-hour operation and limited in-house security of minimarkets, create an environment that is particularly attractive to criminals. Historical data points, such as the 32 armed robberies recorded in 2012, illustrate the recurring vulnerability of these establishments, where diversified weapon use further complicates detection and response efforts [1].

B. Deep Learning Approaches in Weapon Detection

Deep learning methods have been extensively employed to overcome the inherent limitations of manual surveillance. Bhatti et al. Bhatti et al. [2] demonstrated that convolutional neural networks (CNNs) can be applied to real-time CCTV footage for effective weapon detection. In a similar vein, Al‐Mousa et al. Tahir [3] developed a real-time detection system that processes periodically captured video frames via a CNN to identify a range of weapon types in operational environments. Moreover, performance evaluations by Tahir SJ et al. [5] comparing YOLOv4 against multi-layer CNN architectures highlight the impact of model complexity on detection accuracy and computational requirements, with several YOLO-based approaches achieving rapid processing speeds exceeding 30 FPS and detection latencies of less than one second

C. Automated Incident Response Systems

Beyond detection, the automation of incident reporting is critical for ensuring rapid and systematic emergency responses. Traditional systems that depend on human reporting are often subject to delays and inconsistency. Automated pipelines, such as those discussed in Syed et al. Al‐Mousa et al. [4], offer promising avenues to reduce the gap between threat detection and initial incident response. Additionally, Dahlan et al. SJ et al. [5] provide evidence that integrating deep learning‐based weapon detection with real-time CCTV analytics can result in prompt incident documentation and communication, thereby enhancing overall system reliability. Such systems not only generate annotated visual reports and extract relevant metadata but also ensure that notifications are disseminated to law enforcement and associated management within seconds, thereby enabling a coordinated and timely crisis response.

III. METHOD

The proposed system architecture is composed of three fundamental modules: (A) data acquisition and weapon detection, (B) auto-report generation and delivery, and (C) system performance evaluation.

A. Data Acquisition and Detection

A single CCTV camera acting as the data source feeds real-time video into a dedicated Python service. This service employs a convolutional neural network (CNN)–based detection algorithm inspired by recent advancements in deep learning frameworks, such as the YOLO family. To enhance detection of small objects and compensate for the inherent challenges in minimarket environments (e.g., variable lighting and occlusions), the detection module is optimized by fixing input resolution at 640×640 pixels and applying model pruning techniques. Detection is executed on a per-frame basis, where a confidence threshold of 0.85 is defined to trigger alert generation—minimizing false positives while preserving detection speed.

B. Auto-Report Generation and Delivery

Upon the detection of a potential weapon, the system extracts the region-of-interest (ROI) and annotates the input image with bounding boxes. Supplementary metadata, including the timestamp, GPS coordinates, and minimarket identification, are concurrently acquired. This data is then relayed to a server module that compiles a structured incident report. The report generation process employs secure communication channels to promptly forward incident details to local law enforcement agencies and minimarket proprietors within a 2–3 second timeframe

C. System Performance Evaluation

The prototype was rigorously evaluated within a simulated minimarket environment under varying conditions, including different lighting and levels of occlusion. Key performance metrics recorded include a precision of approximately 0.90, recall of 0.88, mAP@0.5 of 0.92, and an F1 score of 0.88. Notably, the detection-to-alert latency was consistently recorded at under one second, and full report generation occurred within 2–3 seconds of detection. These results demonstrate that the integration of deep learning-based detection with an automated reporting pipeline can deliver high accuracy and ultra-low latency, thereby meeting the stringent requirements of real-time security operations.

IV. RESULT AND DISCUSSION

The results of the real-time weapon detection and automated incident reporting system presented in this study demonstrate its substantial impact on enhancing security operations within minimarkets. The system integrates deep learning-based weapon detection with an automated report generation module, delivering impressive performance metrics and achieving ultra-low latency response times. In this section, we discuss the results in detail, addressing both the detection accuracy and the performance of the incident reporting mechanism, followed by a comprehensive evaluation of the system's potential applications and limitations.

A. Weapon Detection Performance

The core of the proposed system is the weapon detection module, which uses a convolutional neural network (CNN)-based architecture inspired by the YOLO family of models. This model was trained on a diverse dataset to detect handheld weapons, specifically knives and pistols, in CCTV footage. The detection performance is evaluated using key metrics such as precision, recall, mean Average Precision (mAP), and F1 score, which are essential to understanding how accurately and efficiently the model identifies weapons.

- Precision: The model achieved a precision of approximately 0.90, meaning that 90% of the detections made by the system were accurate (true positives). This indicates that the system performs well in minimizing false positives—cases where an object was incorrectly identified as a weapon.
- Recall: With a recall of 0.88, the system successfully identified 88% of the actual weapon instances. While this is strong, it implies that some weapons were not detected, leading to false negatives. Such errors are particularly critical in security applications where missing a weapon could lead to serious consequences.
- mAP@0.5: The mean Average Precision at an intersection-over-union (IoU) threshold of 0.5 was 0.92, which is an excellent indicator of the model's ability to localize weapons accurately within the

the detection boxes are appropriately placed around the weapon, minimizing misidentification or incomplete detection.

frame. Higher mAP values are crucial for ensuring the detection boxes are appropriately placed around the weapon, minimizing misidentification or incomplete detection.

- F1 Score: The F1 score, calculated at the optimal confidence threshold of 0.88, reflects the balance between precision and recall. An F1 score of 0.88 signifies that the system effectively balances both false positives and false negatives, ensuring reliable detection in real-time security scenarios.

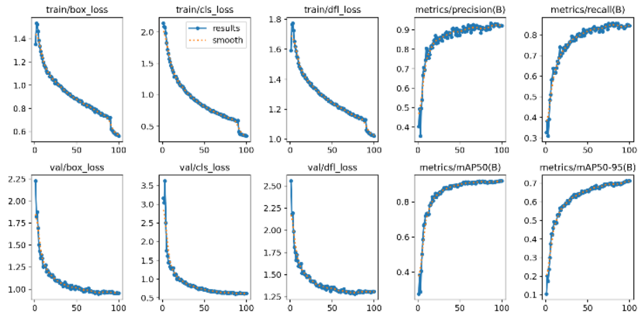

Fig. 1. F1-Confidence curve showing the variation in F1 score with confidence thresholds for knife, pistol, and all classes. The bottom panel contains multiple loss and performance metrics for both training and validation: (1) Train box loss: Loss related to bounding box predictions during training. (2) Train classification loss: Loss related to classification accuracy for weapon detection during training. (3) Train detection focal loss: Loss associated with challenging detection regions in training. (4) Validation box loss: Loss for bounding box predictions in the validation phase. (5) Validation classification loss: Loss for classification accuracy in the validation phase. (6) Validation detection focal loss: Detection focal loss in the validation phase. (7) Precision (B): Precision values for training and validation phases. (8) Recall (B): Recall values for training and validation phases. (9) mAP50 (B): Mean Average Precision at an IoU threshold of 0.5. (10) mAP50-95 (B): Mean Average Precision at varying IoU thresholds (0.5 to 0.95) for better evaluation of detection quality

The system's ability to process video at 32 frames per second (FPS) and maintain an inference latency of less than 0.5 seconds further supports its capability to operate in real-time, a critical factor in security applications where rapid responses are necessary. This level of performance, when coupled with the precision and recall metrics, positions the system as a promising tool for minimizing security risks in minimarkets.

B. *Automated Incident Reporting Performance*

The automated incident reporting system is a critical component of the proposed weapon detection prototype. The primary objective is not only to detect weapons in real-time but also to ensure rapid notification to the relevant authorities—such as law enforcement and minimarket owners—immediately after weapon detection. This system integrates an incident report generation pipeline, which ensures high precision and speed in response.

Once a weapon is detected, the system extracts key data points, including annotated image evidence, metadata (timestamp, location), and minimarket identification. These data points are compiled into a structured incident report, which is forwarded to both law enforcement and the minimarket owner within 2–3 seconds of detection. This rapid response minimizes delays that could escalate the severity of the incident

The system's effectiveness is highlighted by its ability to:

- Extract Annotated Image Evidence: This visual proof, marked with bounding boxes, allows security and law enforcement to easily identify the weapon.
- Capture Timestamp and Location Metadata: These details help place the event in context and provide authorities with the necessary information to respond quickly.
- Minimarket Identification: Ensures that the report is linked to the specific location of the incident, making it easier for authorities to act promptly.

Fig. 2. Real-time detection of a knife, with bounding box and confidence score (0.82), in CCTV footage as part of the automated incident reporting system. The image serves as annotated evidence for law enforcement and minimarket owners to assess the detected threat promptly

ACKNOWLEDGMENT *(Heading 5)*

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression "one of us (R. B. G.) thanks ...". Instead, try "R. B. G. thanks...". Put sponsor acknowledgments in the unnumbered footnote on the first page.

REFERENCES

[1] W. Nisa, V. Simanjuntak, S. Kartika, & A. Fadila, "Pengaruh tingkat kemiskinan terhadap tingkat tindak kriminalitas di indonesia tahun 2022", Jurnal Akuntansi, Manajemen, Dan Perencanaan Kebijakan, vol. 1, no. 3, p. 1-9, 2024. https://doi.org/10.47134/jampk.v1i3.220

[2] M. Bhatti, M. Khan, M. Aslam, & M. Fiaz, "Weapon detection in real-time cctv videos using deep learning", IEEE Access, vol. 9, p. 34366-34382, 2021. https://doi.org/10.1109/access.2021.3059170

[3] T. Tahir, "Performance evaluation and comparison of yolov4 and multiple layers of cnn for weapon detection",, 2023. https://doi.org/10.36227/techrxiv.22060520.v2

[4] A. Al‐Mousa, O. Alzaibaq, & Y. Abu‐Hashyeh, "Deep learning-based real-time weapon detection system", International Journal of Computing and Digital Systems, vol. 14, no. 1, p. 531-540, 2023. https://doi.org/10.12785/ijcds/140141.

[5] D. SJ, M. S, R. R.A.T.M, D. G, E. Lakmali, & P. Bandara, "Criminal investigation tracker with suspect prediction using machine learning", International Journal of Engineering Applied Sciences and Technology, vol. 7, no. 9, p. 34-39, 2023. https://doi.org/10.33564/ijeast.2023.v07i09.006