Available online at : http://jitce.fti.unand.ac.id/
JITCE (Journal of Information Technology and Computer Engineering)
| ISSN (Online) 2599-1663 |

Research Paper

Begal-Detector: A Real-Time Street Crime Detection Framework Combining Human Activity Recognition and Object Detection on Raspberry Pi

*Ismi Batari Agung [1], Ahmad Fadhil Ikram[1], Muhammad Herlan Pratama [1], Nia Mandasari [1], *Pulung Hendro Prastyo [1]*

[1] *Department of Informatics and Computer Engineering, Politeknik Negeri Ujung Pandang, Makassar, Indonesia*

ARTICLE INFORMATION

Received: October 29th , 2025
Revised: November 12th, 2025
Available online: December 3rd, 2025

KEYWORDS

Edge Computing, Object Detection, Human Activity Recognition, YOLOv8, Begal Detection

CORRESPONDENCE

Phone: -
E-mail: pulung.hendro@poliupg.ac.id

A B S T R A C T

Currently, street crime remains a serious challenge in Indonesia, while conventional CCTV systems still function passively as recorders. One of the most concerning types of crime is robbery with violence, commonly known in Indonesia as begal, which remains among the most frequently reported cases. This study proposes the **Begal-Detector**, a YOLOv8-based system that integrates **Human Activity Recognition (HAR)** and **Object Detection** to identify suspicious activities in real time on edge devices. The experiments were conducted on **Raspberry Pi 4, Raspberry Pi 5**, and **Raspberry Pi 5 with Hailo AI Kit**, with variations in distance, camera angle, and lighting conditions. The test dataset consisted of **72 video samples**, including both street crime and non-street crime scenarios, recorded using the **EZVIZ H8C Outdoor CCTV** camera. Experimental results show that the Begal-Detector performs very well, achieving a **100% detection accuracy at a distance of 2 meters**, **94% at 3 meters**, and **94% at a 45° camera angle**. Under low-light conditions supported by infrared light, the system maintained an accuracy of up to **79%**, making it feasible for real-world deployment. In terms of hardware performance, the **Raspberry Pi 5 with Hailo AI Kit** provided the most optimal results, achieving an average of **52.71 FPS** with a stable temperature of **63 °C**, significantly outperforming the **Raspberry Pi 4** and **Raspberry Pi 5** without an accelerator, both of which failed to operate the system in real time. The findings confirm that utilizing **Raspberry Pi 5 with Hailo AI Kit** is an effective solution to ensure that the **Begal-Detector** operates quickly, stably, and reliably.

INTRODUCTION

The crime rate in Indonesia continues to increase at an alarming level, posing a serious threat not only to public safety but also to the comfort of daily life. According to data from the Central Agency of Statistics (BPS Indonesia), the number of criminal cases recorded in 2022 reached 372,965 cases. This figure rose sharply in 2023 to 584,991 cases, representing a 56.85% increase compared to the previous year [1]. However, in 2024, there was a decline in the number of criminal incidents, with the Indonesian National Police (POLRI) reporting 413,037 cases, or a decrease of approximately 29.39% compared to 2023 [2]. Meanwhile, in the first semester of 2025, the Indonesian National Police again recorded an increase, with 287,951 criminal cases, marking a 3.65% rise compared to the same period in the previous year [3].

One of the most concerning types of crime is robbery with violence, commonly known in Indonesia as begal, which remains among the most frequently reported cases. According to a report from National Criminal Information Center (Pusiknas Bareskrim POLRI), from the beginning of 2025 until May 28, 2025, the number of begal victims had reached 2,222 individuals [2]. The nature of begal crimes is particularly alarming, as they often involve acts of violence and the use of weapons. Crime data analysis indicates that 70% of begal incidents occur between 10:00 p.m. and 4:00 a.m., with offenders typically targeting isolated and poorly lit roads. The begal frequently uses sharp weapons such as machetes, knives, and sickles to intimidate their victims, often resulting in serious injuries, and in some cases, loss of life [4].

Law enforcement efforts to address robbery with violence cases have been carried out through continuous patrol intensification and the installation of Closed-Circuit Television (CCTV) systems in areas prone to begal incidents [5]. However, conventional CCTV systems, which have long been relied upon, function merely as passive recording tools without the ability to provide real-time early warnings. By the time the footage is retrieved and analyzed, the perpetrators have often already fled, leaving victims physically and emotionally traumatized. This reactive approach

https://doi.org/10.25077/9.2.40-50.2025

underscores the urgent need for a proactive surveillance system capable of detecting and responding to begal activities as they occur in real time.

Advancements in Artificial Intelligence (AI) and edge computing offer unprecedented opportunities to transform traditional surveillance systems. Modern AI-powered computer vision can analyze video streams in real time, detecting suspicious activities and hazardous objects with high accuracy. This technology has been widely applied in various domains, such as vehicle license plate detection [6] and helmet usage monitoring [7]. Furthermore, through Human Activity Recognition (HAR), systems can identify human movement patterns such as fall detection in the elderly [8], enabling early detection of potentially dangerous situations. However, implementing such sophisticated algorithms on resource-constrained edge devices continues to present significant computational challenges.

The trade-off between accuracy and efficiency remains a significant challenge when deploying AI systems on edge devices. High-performance devices offer fast and accurate detection but are limited by cost and power consumption. Conversely, low-cost devices are more energy-efficient but may compromise detection reliability. This challenge becomes even more complex when Human Activity Recognition (HAR) and Object Detection are combined in real time, as it requires a delicate balance between accuracy, processing speed, and power efficiency to ensure that CCTV-based surveillance systems operate effectively in real-world conditions.

Based on the aforementioned problems, this study proposes a Real-Time Street Crime Detection Framework Combining Human Activity Recognition and Object Detection on Raspberry Pi. The proposed system integrates AI models to simultaneously detect human activities and objects using YOLOv8-pose and YOLOv8 models.

The main contributions of this research are as follows:

1. Development of a multi-model YOLOv8 framework capable of performing human activity recognition and object detection simultaneously to identify begal (robbery with violence) incidents.
2. Performance evaluation of the system on Raspberry Pi 4, Raspberry Pi 5, and Raspberry Pi 5 with Hailo AI Kit, comparing processing speed and power efficiency under various lighting (day/night), angle, and distance conditions.
3. Recommendation of the optimal hardware configuration to support a real-time AI-based begal detection system on edge devices.
4. Provision of a benchmark dataset that can serve as an important reference for future research in edge computing and intelligent surveillance systems.

Related Works

One of the approaches increasingly applied in AI-based surveillance is the use of object detection models to recognize potential threats visually. Pradana et al. explored the use of Faster R-CNN to detect handheld weapons, which are key elements in begal (robbery with violence) incidents. The results demonstrated high accuracy in detecting sharp weapons. However, the implementation was conducted on the Google Colab cloud platform, which requires internet connectivity and high computational resources. This study highlights the importance of weapon detection as an early indicator of potential threats; however, its reliance on cloud computing makes it unsuitable for real-world field deployment [9].

To strengthen the evaluation aspect of object detection on edge devices, Zagitov et al. compared three lightweight object detection models, such as YOLOv8, EfficientDet Lite, and SSD MobileNet, on Raspberry Pi and Jetson platforms. Their findings revealed that although SSD MobileNet demonstrated better energy efficiency, YOLOv8 achieved the highest accuracy [10].

On the other hand, Nurfail et al. developed a criminal activity detection system using the YOLO model, implemented on a Raspberry Pi 4 as a representative edge device. The study demonstrated that the system was capable of operating and detecting activities directly from live camera input, highlighting the potential of using edge devices, such as the Raspberry Pi, for real-time activity detection. However, performance aspects such as detection speed, power consumption, and system responsiveness under varying lighting conditions were not comprehensively evaluated [11].

Furthermore, the study by Feng et al. expanded the understanding of enhancing edge device performance through the use of external accelerators. They evaluated YOLOv3 and YOLOv4 on a Raspberry Pi 4B paired with an Intel Neural Compute Stick 2 (NCS2). They concluded that combining edge devices with accelerators can compensate for computational limitations, particularly when input size and model complexity are properly optimized [12].

The most recent study by Humes et al. emphasized the importance of model optimization through compression techniques to improve power efficiency and throughput on edge devices. Although their approach was not directly applied to criminal detection, the findings serve as an important reference for designing efficient AI-based detection systems. This study supports the adoption of model conversion strategies in the Begal-Detector system, ensuring that inference processes remain optimal even when executed on resource-constrained devices [13].

Although previous studies have made significant contributions, most remain limited to the use of a single detection model. Few studies have integrated sharp weapon detection and human activity recognition within a single system, even though both serve as key indicators in begal incidents. In addition, most research has focused on evaluating a single version of an edge device, without comparing the performance of AI models across different generations of Raspberry Pi. Environmental variations such as lighting conditions (day and night) and object distance have also been rarely examined. Therefore, this study addresses these gaps by developing the Begal-Detector system, a lightweight multi-model framework that integrates YOLOv8 for detecting dangerous objects and YOLOv8-pose for analyzing threatening human movements. Furthermore, this study comprehensively evaluates system performance across multiple generations of Raspberry Pi devices and under various environmental conditions, demonstrating the system's feasibility and efficiency in real-world scenarios.

METHOD

The research methodology was systematically designed, following the workflow illustrated in Figure 1, which begins with data collection, data preprocessing, model training, evaluation, inference, and real-time testing. This process is divided into two main components: The human activity recognition model (YOLOv8-Pose), which recognizes human gestures, and the object detection model (YOLOv8), which identifies sharp weapons.

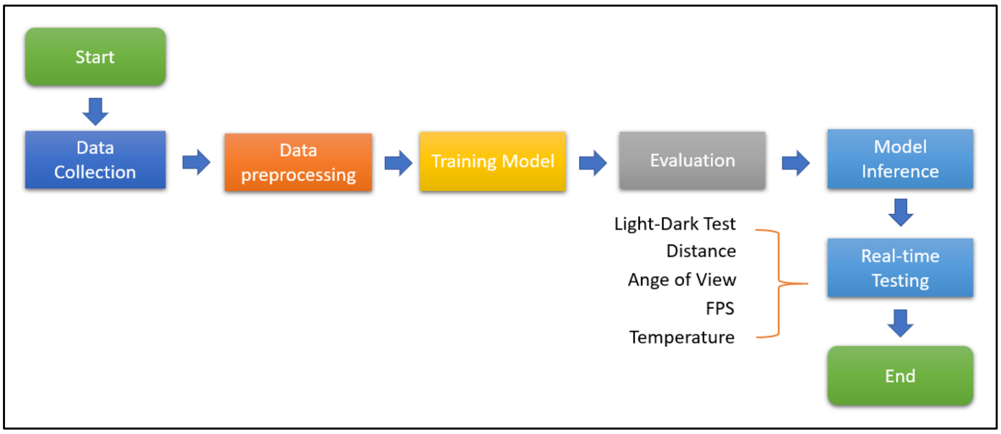


Figure 1. Research Methodology

Data Collection

The initial stage focuses on collecting datasets required for both AI models. First, the human activity recognition model was not trained from scratch; instead, a pre-trained YOLOv8-Pose model was utilized, which is already capable of detecting 17 human skeletal key points. Therefore, no specific data collection process was conducted for this model. Second, for the object detection model, which targets the classes knife, machete, and sickle, a total of 789 images were collected from three primary sources to ensure data diversity. Specifically, 395 images were obtained from a public dataset on Roboflow, 242 images were retrieved from Google Search, and 152 images were captured directly, covering various distances and angle scenarios. All collected images were then annotated using the Roboflow platform, as illustrated in Figure 2, where each sharp weapon object was assigned a bounding box and a class label according to its category.

Figure 2. Example of weapons for the dataset

Data Preprocessing

The human activity recognition model did not require any preprocessing stage, as it utilized a pre-trained model. Meanwhile, the annotated object detection dataset underwent a series of preprocessing steps in Roboflow, which included the following:

1. Resizing: All images were resized to 640 × 640 pixels using the fit (white edges) method to preserve the original aspect ratio.
2. Data Augmentation: Each image was augmented to produce ten new variations using techniques such as horizontal flipping, cropping, rotation, shearing, exposure adjustment, and noise addition. This process expanded the dataset to a total of 6,153 images.
3. Data Splitting: The final dataset was divided into training data (5,960 images), validation data (120 images), and testing data (73 images) for model training and evaluation. The data can be seen in Figure 3.

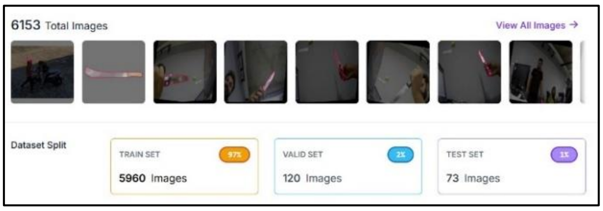


Figure 3. Number of datasets

Training Model

The training process was applied exclusively to the object detection model. The training was conducted on Google Colab using the Ultralytics YOLOv8 library. The model was trained with an image size of 640×640 pixels, a batch size of 16, and 150 epochs. During training, metrics such as box loss, class loss, and DFL loss were monitored to ensure effective learning and to prevent overfitting. The outcome of this process was a best-performing weight file (best.pt), which demonstrated the optimal performance during the validation phase.

Evaluation

After the training process was completed, the object detection model was evaluated to assess its performance quantitatively prior to implementation. This evaluation stage focused on standard metrics commonly used in object detection tasks, namely precision, recall, and mean Average Precision (mAP). Precision was employed to measure the model's accuracy in correctly identifying objects that were truly weapons. At the same time, recall assessed the model's ability to detect all sharp weapon objects present in an image. As an overall performance indicator, mAP provided a comprehensive measure of model accuracy by calculating the average precision across various recall levels. The primary objective of this evaluation was to ensure that the resulting model demonstrated reliable and accurate detection

 https://doi.org/10.25077/9.2.40-50.2025

capabilities before proceeding to the conversion and deployment stages on edge devices.

Inference

After the YOLOv8 object detection model was successfully trained and produced the best weight file in .pt format, the next step was to convert the model for efficient execution on edge devices. This process was carried out in two main stages. First, the model was converted from its original PyTorch format (.pt) to the ONNX (Open Neural Network Exchange) format. ONNX was chosen because it is an industry-standard format that supports interoperability, enabling the model to run on various Raspberry Pi hardware configurations. The second stage involved converting the ONNX file into HEF (Hailo Executable Format), a crucial step specifically for devices equipped with AI accelerators such as the Hailo AI Kit. This conversion was performed using the official Hailo toolkit and aimed to optimize the model for high-performance inference on accelerator hardware. The outcome of this process ensured that the model could perform real-time inference on edge devices with low power consumption while maintaining accuracy, eliminating the need for cloud-based computation.

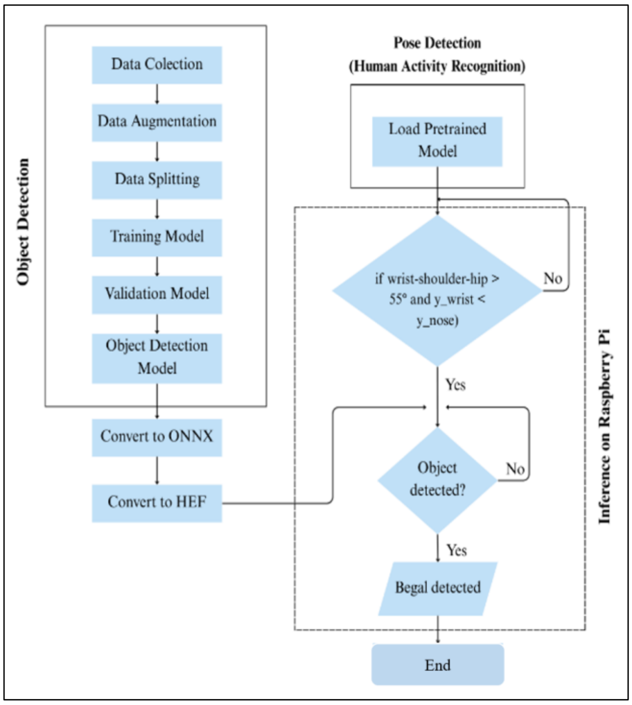


Figure 4. Our proposed method

The next stage involved integrating the two main models, the human activity recognition model (threatening gestures) in HEF format and the sharp weapon detection model in HEF format, into a single intelligent pipeline for real-time begal detection, as shown in Figure 4. The video stream from the CCTV was first processed through a GStreamer pipeline running the human activity recognition model. This model extracted 17 human skeletal keypoints from each frame, including the shoulders, elbows, wrists, hips, and knees.

The coordinates of these keypoints were then utilized in a rule-based human activity recognition logic. At this stage, the system did not employ additional machine learning components; instead, it relied on mathematical computations derived from the detected keypoints to infer human movements and potentially threatening actions.

The system evaluates threatening gestures based on two logical conditions:

1. The arm angle (wrist–shoulder–hip) exceeds 55°, indicating that the hand is raised in a suspicious position.
2. The vertical position of the wrist is higher than the nose (y_wrist < y_nose), indicating that the arm is lifted upward, a gesture commonly associated with intimidation.

If both conditions are satisfied, the system considers the presence of a potential threat. At this point, the sharp weapon detection model is activated. The model is called dynamically, meaning it does not run continuously but only analyzes the area around the hand when a threatening gesture is detected, thereby conserving computational resources.

The outputs from both models are then combined through a decision-making mechanism. An incident is classified as a criminal act if three conditions are simultaneously and consistently met over 20 consecutive frames:

1. A threatening gesture is detected based on the keypoint angle calculations.
2. The object detection model successfully identifies a sharp weapon.
3. At least two people are present in the frame, validating a street crime scenario.

When all these conditions are met, the system classifies the situation as a begal incident.

Real-Time Testing

The system was tested under various scenarios involving distance variations (2 m, 3 m, 4 m, and 5 m), camera angles (45° and 90°), and lighting conditions (bright, dark with infrared light, and dark without infrared light). For each device, a total of 72 test data were used, consisting of 48 begal incident videos and 24 non-begal videos. These datasets were utilized to evaluate the system's ability to accurately detect actual criminal activities while ensuring that non-begal scenes were not falsely identified as threats. The CCTV used for data collection and testing was the EZVIZ H8C Outdoor camera. The example of begal and non-begal data can be seen in Figure 5.

Figure 5. Examples of begal and non-begal datasets

RESULT AND DISCUSSION

The results of the object detection model evaluation and the overall performance analysis of the begal-detector system under various real-time testing scenarios are presented in this section. The discussion focuses on two main aspects: the reliability of the trained AI models and the system's performance across different hardware platforms under varying environmental conditions.

https://doi.org/10.25077/9.2.40-50.2025

Evaluation of Object Detection Model Performance

Prior to deployment on edge devices, the YOLOv8-based sharp weapon detection model was evaluated to assess its performance. The evaluation employed standard object detection metrics, including precision, recall, and mean Average Precision (mAP). Figure 6 illustrates the model's performance trends throughout the training process.

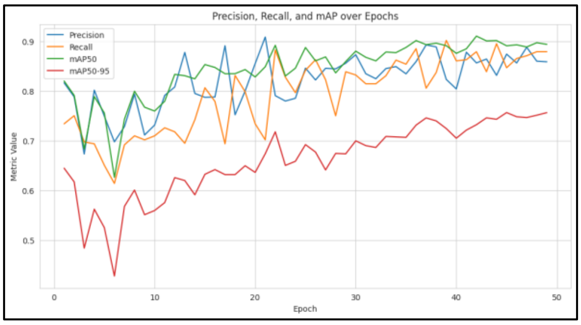


Figure 6. Precision, recall, and mAP

Figure 6 presents the final evaluation results, demonstrating that the trained model achieved excellent performance. The model attained a precision of approximately 0.88 and a recall of 0.87. The high precision value indicates that the model is highly accurate in identifying actual weapon objects, thereby minimizing false detections. Meanwhile, the high recall value reflects the model's ability to detect the majority of sharp weapon objects present in the images.

The overall performance of the model is further confirmed by the mean Average Precision (mAP) scores, with mAP@0.5 reaching 0.90 and mAP@0.5:0.95 reaching 0.76. These results demonstrate that the object detection model possesses high accuracy and reliability, making it suitable for implementation in the Begal-Detector system.

Evaluation of System Performance in Real-time Testing

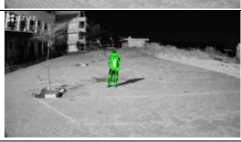
Real-time testing was conducted to comprehensively evaluate the overall system performance, including hardware efficiency and detection accuracy under various conditions. The experiments were performed using three hardware configurations: Raspberry Pi 4, Raspberry Pi 5, and Raspberry Pi 5 equipped with the Hailo AI Kit. A summary of the test results is presented in Table 1.

Table 1. Summary of begal detection scenarios and results

Label	Test Conditions			Video Frame	Raspberry Pi 4		Raspberry Pi 5		Raspberry Pi 5 + Hailo AI Kit		Result
	Lighting	Distance	Angle		Avg FPS	Temp (ºC)	Avg FPS	Temp (ºC)	Avg FPS	Temp (ºC)	
Begal	Bright	2m	45º		0.34	86.9	1.52	83.6	54.95	64.5	Yes
			90º		0.47	85.3	1.34	83.7	53.78	63.9	Yes
		3m	45º		0.22	86.9	1.62	82.0	54.32	65.0	Yes
			90º		0.40	86.88	1.30	83.7	55.96	64.5	Yes
		4m	45º		0.25	85.3	1.52	83.7	48.16	73.2	Yes
			90º		0.41	85.5	1.37	82.8	53.99	62.8	No
		5m	45º		0.30	85.7	1.54	84.0	54.88	56.8	No
			90º		0.46	86.2	1.52	83.7	56.64	57.9	Yes

 https://doi.org/10.25077/9.2.40-50.2025

	2m	45º		0.23	86.4	1.32	82.7	42.55	63.5	Yes
		90º		0.47	86.0	1.24	85.7	51.47	65.6	Yes
	3m	45º		0.30	85.1	1.63	82.5	57.83	63	Yes
Dark with Infrared		90º		0.42	85.6	1.35	83.8	52.76	63.5	Yes
	4m	45º		0.29	85.7	1.42	83.2	52.76	63.5	Yes
		90º		0.36	86.9	1.57	84.8	53.53	61.8	No
	5m	45º		0.49	85.6	1.54	84.0	54.76	58.9	No
		90º		0.25	85.2	1.42	84.7	50.81	61.2	No
	2m	45º		0.21	86.5	1.62	82.6	52.78	66.8	Yes
		90º		0.34	85.6	1.34	83.9	51.06	56.8	Yes
	3m	45º		1.62	85.2	1.62	82.0	50.68	69.0	Yes
Dark without Infrared		90º		1.35	86.0	1.35	83.8	55.53	62.9	No
	4m	45º		1.42	85.1	1.42	83.2	51.59	56.2	Yes
		90º		1.37	85.6	1.37	82.8	52.16	64.7	No
	5m	45º		1.54	87.0	1.54	84.0	50.84	68.4	Yes

			90º		1.52	85.6	1.52	83.7	51.31	59.4	No
Non-Begal	Bright	2m	45º		0.40	86.8	1.30	83.7	55.96	64.5	Yes
			90º		0.25	85.3	1.52	83.7	48.16	73.2	Yes
		3m	45º		0.41	85.5	1.37	82.8	53.99	62.8	Yes
			90º		0.30	85.7	1.54	84.0	54.88	56.8	Yes
		4m	45º		0.46	86.2	1.52	83.7	56.64	57.9	Yes
			90º		0.23	86.4	1.32	82.7	42.55	63.5	Yes
		5m	45º		1.42	85.1	1.42	83.2	51.59	56.2	Yes
			90º		1.37	85.6	1.37	82.8	52.16	64.7	Yes
	Dark with Infrared	2m	45º		1.54	87.0	1.54	84.0	50.84	68.4	Yes
			90º		1.52	85.6	1.52	83.7	51.31	59.4	Yes
		3m	45º		0.34	86.9	1.52	83.6	54.95	64.5	Yes
			90º		0.47	85.3	1.34	83.7	53.78	63.9	Yes
		4m	45º		0.22	86.9	1.62	82.0	54.32	65.0	Yes
			90º		0.40	86.8	1.30	83.7	55.96	64.5	Yes

 https://doi.org/10.25077/9.2.40-50.2025

	5m	45º		0.25	85.3	1.52	83.7	48.16	73.2	Yes
		90º		0.41	85.5	1.37	82.8	53.99	62.8	Yes
Dark without Infrared	2m	45º		1.54	87.0	1.54	84.0	50.84	68.4	Yes
		90º		1.52	85.6	1.52	83.7	51.31	59.4	Yes
	3m	45º		0.34	86.9	1.52	83.6	54.95	64.5	Yes
		90º		0.47	85.3	1.34	83.7	53.78	63.9	Yes
	4m	45º		0.22	86.9	1.62	82.0	54.32	65.0	Yes
		90º		0.40	86.8	1.30	83.7	55.96	64.5	Yes
	5m	45º		0.25	85.3	1.52	83.7	48.16	73.2	Yes
		90º		0.41	85.5	1.37	82.8	53.99	62.8	Yes

Hardware Performance (FPS and Temperature)

The comparison of hardware performance serves as a key benchmark for determining the feasibility of real-time implementation. Figure 7 summarizes the average Frame Per Second (FPS) and operational temperature across the three tested devices.

The experimental results revealed a substantial difference in performance across devices. The Raspberry Pi 5 equipped with the Hailo AI Kit demonstrated superior capability, processing video streams at an average speed of 52.71 FPS while maintaining a stable operating temperature of 63.08°C. In contrast, the Raspberry Pi 5 without the accelerator achieved only 1.46 FPS and operated at a significantly higher temperature of 83.53°C. The Raspberry Pi 4 performed even worse, reaching just 0.63 FPS with an operating temperature of 85.91°C.

These findings suggest that the AI accelerator is not only crucial for achieving real-time processing speed but also vital for maintaining thermal stability and preventing overheating. Without the accelerator, neither the Raspberry Pi 4 nor the Raspberry Pi 5 can effectively execute the system of begal-detector.

Detection Performance Analysis

The system's accuracy in detecting begal incidents was evaluated based on three main variables: distance, viewing angle, and lighting conditions.

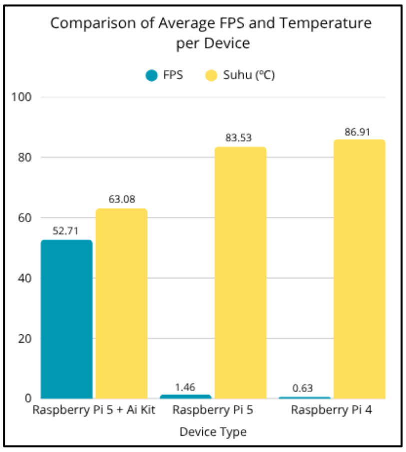


Figure 7. Comparison of average FPS and temperature across device types

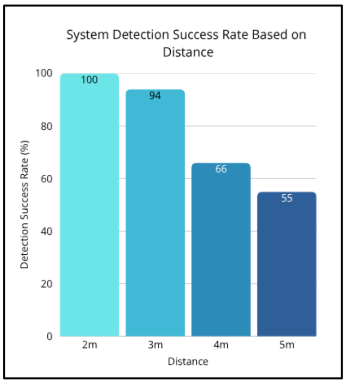


Figure 8. System detection success rate based on distance

The distance between the camera and the object proved to be the most critical factor influencing detection success. As shown in Figure 8, the system achieved perfect performance at a distance of 2 meters, with a 100 percent success rate. This performance remained high at 3 meters, reaching a 94 percent success rate. However, the system's effectiveness began to decline as the distance increased, dropping to 66% at 4 meters and 55% at 5 meters. This decrease is reasonable since objects, particularly weapons, appear smaller and lose visual detail at greater distances, making them more difficult for the model to detect accurately.

Next, the camera placement angle also had a significant impact on system performance. As shown in Figure 9, the system performed considerably better at a 45° angle, achieving a 94 percent success rate. In contrast, at a 90-degree angle, performance dropped to 66%. The 45° angle provided a clearer view of body posture and hand-raising gestures for the human activity recognition model, allowing it to detect threatening actions more accurately.

Finally, the system was tested under three lighting conditions to simulate both daytime and nighttime environments. As shown in Figure 10, the system performed best under bright lighting conditions, achieving an accuracy of 83%. In dark conditions with infrared light, performance decreased slightly. However, it remained high at 79% accuracy, demonstrating that the infrared feature of the CCTV is highly effective in maintaining system reliability during nighttime operations. Meanwhile, under complete darkness without infrared light, the system achieved an accuracy of 75%.

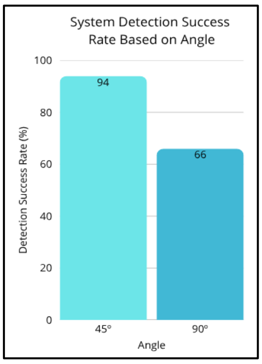


Figure 9. System detection success rate based on angle

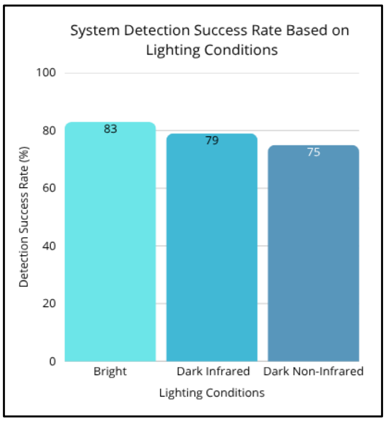


Figure 10. System detection success rate based on lighting conditions

Confusion Metric Analysis

Based on the model performance comparison chart presented in Figure 11, the analysis across the three lighting conditions (bright, dark, and dark with infrared) yields critical findings regarding the strengths and weaknesses of this begal detection model. Overall, the Precision metric exhibits perfect consistency, achieving 100% across all operational environments. This ideal precision value constitutes a major strength of the model, indicating that every instance classified as a begal is correctly identified, thereby effectively minimizing False Positives. However, this perfect consistency often suggests a trade-off with the recall (sensitivity) metric, which represents the model's primary weakness and exhibits the greatest variability. Under the ideal bright condition, the model achieved its optimal performance with an accuracy of 83.34% and a recall of 75%, resulting in an f1-score of 85.71%. The model's vulnerability became distinctly evident in the dark condition, where recall sharply declined to 62.75%, consequently reducing accuracy to 75% and the f1-score to 76.92%. This decline highlights the model's significant susceptibility to visual degradation caused by low light, which ultimately leads to an increase in False Negatives (missed begal incidents). The proposed solution utilizing infrared light in the dark condition proved effective, successfully recovering recall to 68.75% and accuracy to 79.17%. In summary, this indicates that while the model is highly reliable (100% precision) in confirming detected street crime events, it lacks the necessary sensitivity (low recall) to comprehensively identify all actual street crime occurrences.

 https://doi.org/10.25077/9.2.40-50.2025

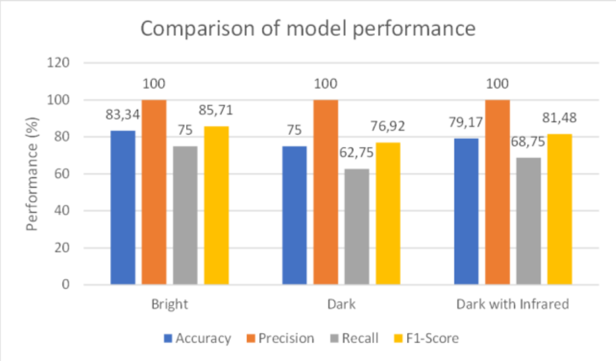


Figure 11. Comparison of model performance

.

False Positive Analysis

In addition to detecting threats, the reliability of a surveillance system is also measured by its ability to distinguish between normal activities and threats accurately. In this experiment, the system demonstrated perfect performance. Across all non-robbery scenarios tested, the system accurately recognized normal activities with a 100% accuracy rate, indicating no false-positive detections. This result is significant as it confirms that the designed human activity recognition logic is sufficiently specific to detect threatening gestures without being triggered by normal movements, ensuring dependable operation and minimizing false alarms. In contrast, if the system relied solely on a single model, either for human activity recognition or object detection, it would likely produce a high false positive rate, as it could be misled by benign activities such as people cutting grass, chopping vegetables, or children playing.

Overall, this experiment successfully established the optimal configuration and operating conditions for real-time robbery detection. In practical field implementation, the Raspberry Pi 5 equipped with the Hailo AI Kit is recommended as the primary platform, as it is the only device capable of achieving sufficient processing speed for real-time video analysis. The highest detection effectiveness was achieved when the system monitored targets at an ideal distance of 2 to 3 meters with the camera positioned at a 45° angle. Furthermore, for nighttime surveillance, the use of infrared light proved to deliver significantly higher accuracy and reliability compared to operation without infrared support. Therefore, these configurations are key to maximizing the accuracy and dependability of the system in real-world scenarios

CONCLUSIONS

This study successfully developed and evaluated the Begal-Detector system, a multi-model YOLOv8 framework that integrates human activity recognition and object detection on an edge platform. The evaluation utilized primary video data, consisting of 72 recordings that represented both begal and non-begal activities. Based on the experimental results, the proposed method performed best at a distance of 2 to 3 meters, achieving success rates of 100% and 94%, respectively. Optimal performance was also observed at a 45-degree camera angle (94%) and under bright lighting conditions (83%). For nighttime scenarios, the use of infrared light proved significantly more effective (79%) compared to operating without infrared light (75%).

From the hardware perspective, the Raspberry Pi 5 equipped with the Hailo AI Kit proved to be the most optimal configuration. This device achieved an average of 52.71 FPS while maintaining a stable temperature of 63°C, enabling reliable real-time system operation. In contrast, both the Raspberry Pi 4 and the Raspberry Pi 5 without an accelerator were inadequate due to limitations in processing speed and thermal stability.

Overall, this study confirms that AI-based begal detection on edge devices is highly feasible, supported by AI accelerators (**Hailo AI Kit**) and proper camera configuration. For future research, improving detection accuracy at longer distances and under low-light conditions should be a key priority, either through model optimization or the integration of additional supporting technologies.

REFERENCES

[1] bps, “Statistik Kriminal 2024,” https://www.bps.go.id/. Accessed: Aug. 17, 2025. [Online]. Available: https://www.bps.go.id/id/publication/2024/12/12/13317138a55b2f7096589536/statistik-kriminal-2024.html

[2] pusiknas, “Tiap Bulan, Lebih 400 Kasus Curas Ditangani Polri,” https://pusiknas.polri.go.id/. Accessed: Aug. 15, 2025. [Online]. Available: https://pusiknas.polri.go.id/detail_artikel/tiap_bulan,_lebih_400_kasus_curas_ditangani_polri

[3] Human Kemenko Polkam RI, “Pemerintah Perkuat Sinkronisasi Nasional,” https://polkam.go.id/. Accessed: Aug. 15, 2025. [Online]. Available: https://polkam.go.id/kejahatan-meningkat-penyelesaian-menurun-pemerintah-perkuat-sinkronisasi-nasional/

[4] Z. Hasan, D. A. AS, A. Febriyanti, and S. Mariska, “Kriminalitas Pencurian Sepedah Motor Di Desa Gandri Kecamatan Penengahan Kabupaten Lampung Selatan,” *Jurnal Rectum*, vol. 5, no. 3, pp. 245–252, Sep. 2023, doi: 10.46930/jurnalrectum.v5i3.3501.

[5] M. R. N. Tjaya, M. R. Permana, and M. J. Prasetiyo, “Penegakan Hukum Terhadap Tindak Pidana Pencurian Kendaraan Bermotor Dengan Kekerasan,” *Jurnal Kewarganegaraan*, vol. 8, no. 1, pp. 324–330, 2024, [Online]. Available: http://repository.uir.ac.id/id/eprint/10028

[6] A. Meirza and N. R. Puteri, “Implementasi Metode YOLOV5 dan Tesseract OCR untuk Deteksi Plat Nomor Kendaraan,” *Jurnal Ilmu Komputer dan Desain Komunikasi Visual*, vol. 9, no. 1, 2024.

[7] M. A. Meidyan and W. Yustanti, “Implementasi Metode You Only Look Once (Yolov5) dalam Deteksi Pelanggaran Helm,” *Journal of Emerging Information Systems and Business Intelligence*, vol. 5, pp. 214–222, 2024.

[8] A. Muhammad, A. Alfajra, D. Syauqy, and G. E. Setyawan, “Perancangan Wearable Vest Pendeteksi Jatuh Lansia Berbasis Esp-32 dan Sensor Mpu9250 Menggunakan Algoritma Support Vector Machine,” vol. 1, no. 1, pp. 1–9, 2024.

[9] I. C. Pradana, E. Mulyanto, and R. F. Rachmadi, “Deteksi Senjata Genggam Menggunakan Faster R-CNN Inception V2,” *Jurnal Teknik ITS*, vol. 11, no. 2, 2022, doi: 10.12962/j23373539.v11i2.86587.

[10] A. N. Aziz, atul Khoiriyah, F. Abdillah, and I. Gede Wiryawan, “Prototipe Sederhana Sistem Deteksi Kriminal Berbasis Internet Of Things Menggunakan Teknologi YOLOv5,” *Jurnal Sistem Komputer*, vol. 13, no. 1, pp. 140–147, May 2024, doi: 10.34010/komputika.v13i1.12217.
[11] H. Feng, G. Mu, S. Zhong, P. Zhang, and T. Yuan, “Benchmark Analysis of YOLO Performance on Edge Intelligence Devices,” *Cryptography*, vol. 6, no. 2, Jun. 2022, doi: 10.3390/cryptography6020016.
[12] E. Humes, M. Navardi, and T. Mohsenin, “Squeezed Edge YOLO: Onboard Object Detection on Edge Devices,” Dec. 2023. Available: http://arxiv.org/abs/2312.11716