Robust Firearm Detection in Low-Light Surveillance Conditions Using YOLOv11 with Image Enhancement

Ram Pravesh*[ID], Bikash Chandra Sahana[ID]

Electronics & Communication Engineering Department, National Institute of Technology, Patna 800001, India

Corresponding Author Email: ramp.phd19.ec@nitp.ac.in

ABSTRACT

In modern surveillance systems, real-time detection of security threats such as firearms in low-light environments remains a significant challenge. This study presents a robust firearm detection framework based on the YOLOv11 object detection model, enhanced with a three-stage image pre-processing pipeline tailored for dark conditions. The proposed system integrates adaptive gamma correction, Gaussian noise reduction, and min-max normalization to improve visual clarity before detection. Images from publicly available datasets were synthetically darkened to simulate real-world low-light scenarios. A custom dataset with 3,107 images was used to train and evaluate the model. The enhanced YOLOv11 model achieved a detection accuracy of 97.28%, with a mean F1-score of 95.78%, significantly outperforming the baseline YOLOv11 under dark conditions. This study demonstrates that strategic image enhancement improves detection robustness and reduces false positives and false negatives in low-light surveillance applications.

1. INTRODUCTION

Firearms, though essential for law enforcement and defense, pose serious threats when accessed by unauthorized individuals such as anti-social elements or mentally unstable persons. Several mass shooting incidents have occurred globally in recent years, particularly in public spaces such as schools, hospitals, and market areas [1, 2]. These events have intensified the need for advanced, real-time surveillance systems capable of identifying security threats, especially the presence of firearms.

Closed-Circuit Television (CCTV) systems have emerged as crucial components in public safety infrastructure, assisting in early threat detection and forensic investigation [3]. However, the efficiency of these systems remains limited due to the challenges posed by manual monitoring, especially under low-light or nighttime conditions. Surveillance effectiveness is significantly reduced in such environments, increasing the likelihood of missed detections [4].

Traditional firearm detection models, including those based on Faster R-CNN, SSD, and earlier YOLO variants, have demonstrated success in controlled lighting conditions [5-7]. However, their performance deteriorates under low illumination due to insufficient feature visibility, resulting in a high rate of false positives and missed detections [8-10]. Some recent studies have applied pre-processing techniques like brightness tuning and contrast normalization to mitigate these issues, but many of them compromise real-time performance or fail to generalize across varying lighting levels [11]. This research addresses a critical gap in the literature by focusing on firearm detection in low-light surveillance environments. It proposes an integrated solution combining a three-stage image enhancement pipeline adaptive gamma correction, Gaussian noise reduction, and min-max normalization with the state-of-the-art YOLOv11 detection model [12]. YOLOv11, presented at the YOLO Vision 2024 Conference (YV24), represents the latest evolution in the YOLO series and incorporates architectural innovations that significantly improve accuracy, detection speed, and computational efficiency [13, 14].

A recent advancement, such as the YOLOv7-DarkVision model [15], has demonstrated the merit of targeted pre-processing for object detection in dark conditions. Building upon this idea, the proposed system aims to improve real-time firearm detection accuracy by enhancing visual clarity before detection, rather than merely post-processing detection outputs.

The substantial advantages of using the proposed firearm detection system are:

- **Enhanced Safety with Robust Surveillance:** The system effectively detects firearms even in poorly lit environments, enabling timely alerts for law enforcement.
- **Reduced Dependence on Manual Monitoring:** The automated detection pipeline minimizes human fatigue and error.
- **Improved Situational Awareness:** Especially valuable in fenced zones or areas with limited lighting, it enhances visibility where artificial lighting is impractical.

In summary, this work contributes a computationally efficient, high-accuracy framework tailored for real-time firearm detection in dark environments. It fills a critical void

in existing research and opens new avenues for the deployment of AI-enhanced surveillance in sensitive civilian and military applications.

Section II presents a comprehensive review of existing firearm detection methods and highlights gaps in performance under low-light conditions. Section III describes the dataset preparation process, including image sourcing, annotation, and the simulation of low-light environments through controlled darkening transformations. Section IV details the proposed methodology, including the selection and tuning of YOLOv11, the three-stage image enhancement pipeline, system architecture, and the end-to-end detection workflow. Section V provides a structured evaluation of the model, including quantitative results, enhancement effectiveness, visual performance, and comparative analysis with and without enhancement. Finally, Section VI concludes the study and outlines future research directions.

2. RELATED WORK

While the motivation for intelligent firearm detection in low-light surveillance environments is well established, numerous approaches have been proposed over the years ranging from traditional machine learning techniques to recent deep learning-based object detectors. However, many of these methods either rely on fully visible weapons or struggle with performance degradation under poor lighting conditions. The following section reviews these existing approaches and highlights the limitations that this study aims to address through an enhancement-integrated YOLOv11 framework.

Numerous research efforts have explored the use of computer vision and deep learning for firearm recognition in surveillance imagery [16-18]. These approaches range from traditional X-ray image analysis to real-time object detection using deep neural networks.

Early research in security-focused weapon detection predominantly relied on X-ray scanning techniques for concealed object detection in controlled environments such as airports and transport hubs [19-21]. While effective in luggage scanning, these methods are not applicable to open, real-world settings such as public areas or street surveillance, especially when detecting firearms carried openly or concealed on individuals at a distance.

Recent advances have shifted toward using visible-spectrum CCTV imagery combined with machine learning and deep learning models. For example, Tiwari and Verma [18] applied handcrafted feature extractors like HIPD and FREAK combined with K-means clustering, but required full firearm

visibility. Olmos et al. [9] used Faster R-CNN for handgun detection but faced performance limitations due to high computational cost and difficulty detecting small objects. Bhatt and Ganatra [12] tested multiple models including YOLOv4, SSD, and Inception ResNetV2, yet still reported high false negative rates.

In a related study, Wang et al. [7] introduced brightness-guided preprocessing (DaCoLT) to improve detection under poor lighting, but their model achieved a low frame rate (1.3 FPS), limiting real-time applicability. Similarly, Wang et al. [20] proposed improvements to YOLOv4 for detecting small weapons, but their model's performance on synthetic and darkened images remained suboptimal.

Amado-Garfias et al. [21] fused YOLOv4 outputs with Random Forest classifiers to infer if a person is armed, yet achieved an accuracy of only 85.44%, highlighting the limitations of hybrid models in complex scenarios. Basit et al. [8] explored human-object interaction features, but the detection accuracy depended heavily on posture and body position, limiting robustness in dynamic environments.

More recently, studies have explored the integration of human pose estimation to improve handgun detection, as shown in the works of Ruiz-Santaquiteria et al. [4] and Velasco-Mata et al. [10]. Further, Lim et al. [22] proposes a strengthened deep multi-level feature pyramid network that takes into account the challenge of determining small firearms from a non-canonical standpoint. V.P. Manikandan et. al. proposes a CNN Attuned Object Detection Scheme (AODS) for harmful object detection from closed-circuit television surveillance images. Though this approach helps reduce false positives, it adds complexity and computational overhead.

A promising direction was proposed by Yadav et al. [23] with their YOLOv7-DarkVision framework, where firearm detection was improved using targeted preprocessing in dark environments. However, YOLOv7, while powerful, lacks the latest architectural optimizations introduced in YOLOv11.

A consolidated summary of major studies, including strengths and limitations, is shown in Table 1. From this comparative analysis, it is evident that no existing model directly addresses robust firearm detection in low-light CCTV conditions with a real-time deployment focus.

The proposed research fills this gap by integrating adaptive image enhancement techniques with YOLOv11, the latest state-of-the-art model in the YOLO family, to achieve accurate and efficient firearm detection under varying lighting conditions. This hybrid strategy enables robust feature extraction from darkened images without compromising inference speed making it highly suitable for practical surveillance applications.

Table 1. Current methods for firearms recognition based on the technique used, strength and limitations

Ref.	Published Year	Techniques/Strengths	Limitations	Dark Environment Suitability
[15]	2015	Combines HIPD, FREAK, and K-means to enhance accuracy in color-based segmentation.	Gun should be fully visible, high miss detection with images having limited visibility.	No
[16]	2018	Faster R-CNN Two-stage deep learning method for object detection tasks, with advantage of precisely detection of small firearms.	Time-consuming, difficult, and computationally expensive.	No
[18]	2019	Improves detection using brightness-directed reprocessing using Inception Res NetV2 and DaCoLT.	Lower FPS (1.3) not fit for low-latency performance.	Limited to special types of Firearms
[4]	2020	Faster-RCNN with ResNet-101 used	Only detects weapons if	No

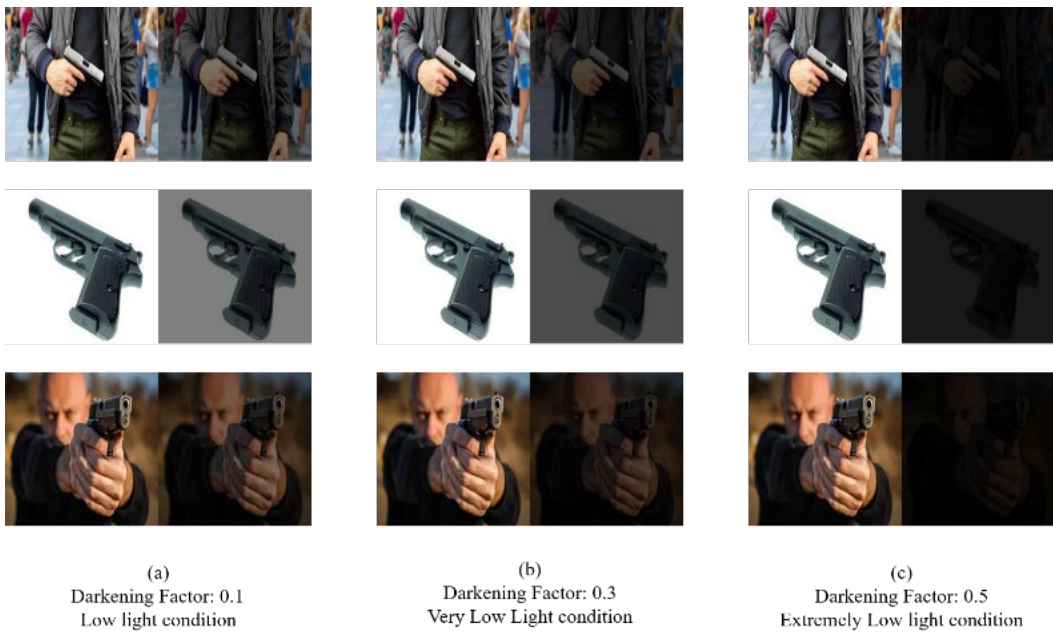|		paired-bounding-boxes to identify human carrying the firearm.	carried by a person.	
[17]	2021	Faster RCNN-Inception ResNetV2, SSD, YOLO4 is suggest based on the detection results.	High false negatives despite 88% recall rate with YOLOv4.	No
[6]	2021	Handgun detection using human pose and appearance.	Computationally expensive; sensitive to pose estimation quality.	No
[7]	2021	Incorporates human pose info to reduce false positives.	Requires precise skeletal tracking; not real-time ready.	No
[20]	2023	YOLOv4 Optimization for enhanced detection against complex backgrounds using spatial attention and multi-scale dilation.	Less effective on synthetic data and images with complex backgrounds.	No
[19]	2024	The ML models (YOLOv4 with Random Forest Classifier) can work together with YOLO to identify automatically armed people.	Performance accuracy (85.44%) is low, Higher version of YOLO model is need to improve the same.	No
[24]	2024	Fuzzy Discernible Feature Selection, an automatic lightweight artificial intelligence solution	Only detects weapons if carried by a person.	No
[25]	2024	YOLOv7-DarkVision with custom enhancement for dark environments.	Good improvement in dark settings, but lacks YOLOv11 advancements.	Yes
Proposed Method	--	YOLOv11 + adaptive gamma correction, Gaussian filtering, and normalization.	Balanced for accuracy and speed; tailored for real-time low-light firearm detection.	Yes

3. DATABASE

The dataset used in research work are legally and ethically collected from database platforms like Kaggle, Linksprite, VBS3, IMFDB, and UGR, and through independent internet searches. The details of image collection are provided in Table 2. A darkening transformation is used to create a low-light scenario on the collected images. To simulate a real-world scenario, various darkening factors have been used. The concept of darkening factors refers to the controlled reduction of pixel intensity values in an image to simulate darker environments or low lighting conditions. This is achieved by multiplying the pixel values of an image by a scalar factor, which reduces the brightness uniformly across all pixels. The darkening transformation is applied to the images after annotation and splitting. The Dataset is divided into training (70%), validation (15%), and testing (15%) sets. Figure 1 shows the sample images pairs from dataset with various darkening factor. In each image pair, the right side shows the darken image of the original image (left side). There are three darkening factor is considered for experimentation, to cover the complete dark environment scenario.

Table 2. Detail of collection of image dataset

Images details	Images	Classes	Sources
Pistols	1048		VBS3 [25], IMFDB [26], UGR [27], Linksprite [28], Kaggle [29], Internet
Shotguns	1023	Firearms	
Handguns	1036		
Total	3107		

(a) Darkening Factor: 0.1 Low light condition

(b) Darkening Factor: 0.3 Very Low Light condition

(c) Darkening Factor: 0.5 Extremely Low light condition

Figure 1. Image samples pairs from dataset with various darkening factor

4. METHODOLOGY

This section presents the design and implementation of the proposed firearm detection system that operates effectively under low-light surveillance conditions. The system integrates an image enhancement pipeline with a state-of-the-art YOLOv11-based object detector to improve visibility, reduce false detections, and ensure high-speed performance. The proposed methodology comprises of:

- Selection and tuning of the YOLOv11 model,
- Design of the three-stage image enhancement pipeline,
- Architecture of the proposed system,
- End-to-end firearm detection procedure.

The combination aims to reduce detection failures caused by poor lighting, enhancing system reliability for real-world applications.

4.1 Selection and tuning of the YOLOv11 model

YOLOv11 is the latest version of the "You Only Look Once" series, introduced at the YOLO Vision 2024 Conference (YV24) [30, 31]. It provides substantial advancements in object detection accuracy, speed, and scalability. "You Only Look Once" model foundation was established by Joseph Redmon et al. [32]. YOLOv11 inherits the efficient codebase of YOLOv8 [33], while integrating architectural improvements from YOLOv9 [34] and YOLOv10 [35]. Five model variants of YOLOv11 are available (nano to x-large), suited for different deployment scenarios [36]. Table 3 presents the five deployment variants of the YOLOv11 model, categorized based on their architectural depth, width, number of parameters, and computational cost [37]. Among these, the YOLOv11n (nano) model is selected for the proposed system due to its lightweight design (2.62 million parameters and 6.6 GFLOPs), making it highly suitable for real-time firearm detection in resource-constrained surveillance environments. Despite being the smallest variant, YOLOv11n retains essential architectural features such as multi-scale detection, spatial pyramid pooling, and attention mechanisms enabling it to perform effectively on low-resolution and low-light imagery. This balance between speed and accuracy makes YOLOv11n ideal for real-world deployment in smart CCTV systems.

The architecture of YOLOv11n in shown in Figure 2, which serves as the core detection engine in the proposed system. The YOLOv11 pipeline consists of:

- **Backbone:** Multi-scale feature extraction through convolutional and C3k2 modules.
- **Neck:** Feature fusion using SPPF and C2PSA blocks.
- **Head:** Object detection using bounding boxes, confidence scores, and class probabilities at three scales (P3, P4, P5).

Table 3. YOLO11 as per deployment requirements [37]

Variant	Purpose	Depth	Width	Max Channels	Layers	Parameters	Gradients	GFLOPs
YOLO11n Nano version	For small and lightweight tasks.	0.50	0.25	1024	319	2,624,080	2,624,064	6.6
YOLO11s Small version	An upgrade of Nano, offering slightly improved accuracy.	0.50	0.50	1024	319	9,458,752	9,458,736	21.7
YOLO11m Medium version	General-purpose use.	0.50	1.00	512	409	20,114,688	20,114,672	68.5
YOLO11l Large version	Higher accuracy, with increased computational cost.	1.00	1.00	512	631	25,372,160	25,372,144	87.6
YOLO11x Extra-large version	For maximum accuracy and performance.	1.00	1.50	512	631	56,966,176	56,966,160	196.0
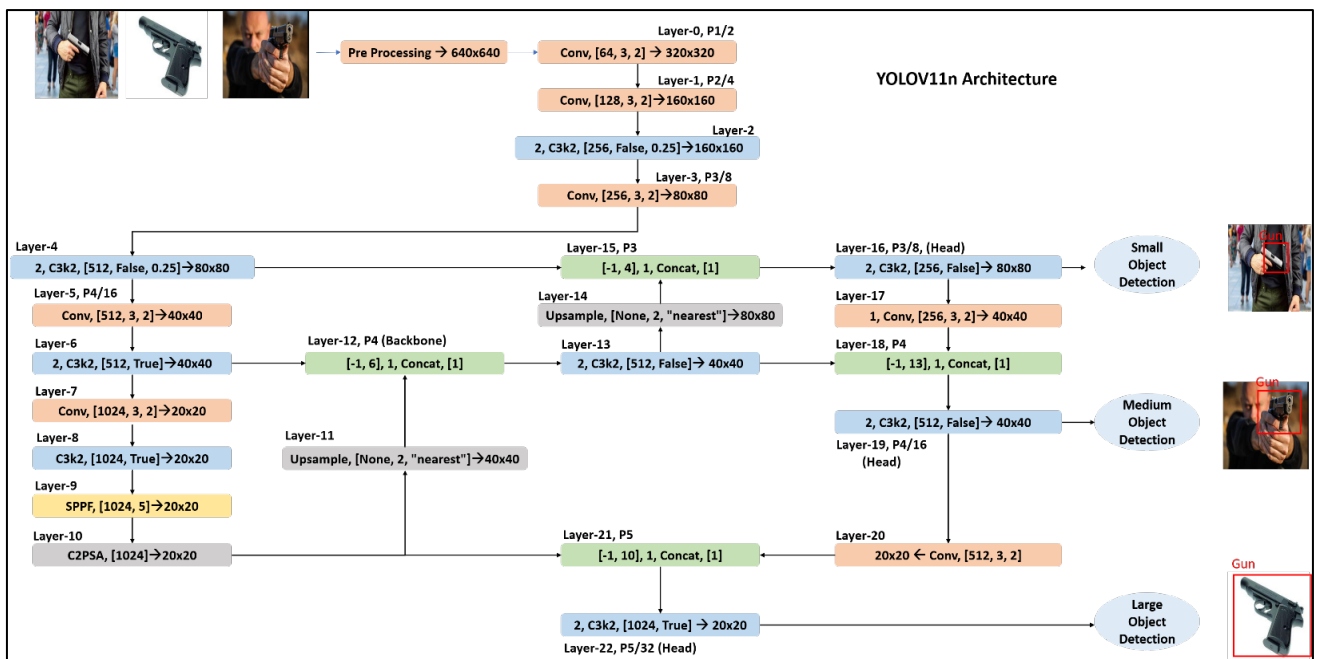
Figure 2. Architecture of YOLO11n [37]

The backbone (Layer-1 to Layer-10) is for pull out multi-scale features from the input image through various convolutional layers, C3k2 modules, and specialized operations like SPPF and C2PSA. These layers perform down sample the input image, generating hierarchical feature maps (P1 to P5) of various resolutions, with P3, P4, and P5. These are the key outputs for small, medium, and large objects. The neck refines these features using up sampling, concatenation, and additional C3k2 modules to fuse information across scales, ensuring effective feature representation for objects of varying sizes. Finally, the head processes the refined features at P3, P4, and P5 to produce bounding box predictions, confidence scores, and class probabilities for object detection. This multi-scale approach allows YOLO11 to detect objects of different sizes efficiently while maintaining a balance between accuracy and computational cost [38].

4.2 Design of the three-stage image enhancement pipeline

The core objective of the image enhancement pipeline is to boost the visibility and contrast of surveillance images captured under low-light conditions. This preprocessing enables YOLOv11n to extract discriminative features more accurately, necessary for firearm detection [39, 40]. The proposed enhancement pipeline includes three sequential stages:

- **Adaptive Gamma Correction** for brightness enhancement,
- **Gaussian Noise Reduction** for edge-preserving smoothing,
- **Min-Max Normalization** for contrast enhancement.

Each step is tailored to be lightweight and computationally efficient to maintain real-time detection feasibility.

4.2.1 Adaptive gamma correction (brightness enhancement)

Purpose: Gamma correction does a non-linear transformation to fine-tune image brightness. This technique used to adjust the brightness of an image by changing the pixel intensity values according to a power-law function. Due to the use of non-linearity gamma correction enhances darker regions without excessively brightening already illuminated areas. The transformation is given by:

Process Overview: Gamma correction adjusts pixel intensity using a power- law transformation:

$$I_{Gamma}(x,y) = 255 \times \left(\frac{I(x,y)}{255}\right)^{\gamma} \tag{1}$$

where,

- γ is the gamma correction factor.
- $I(x,y)$ represents the pixel intensity of the (x,y).
- $I_{Gamma}(x,y)$ is the output pixel intensity after gamma correction.

The choice of γ plays an important role in the enhancement process; γ values less than 1 brighten the image, while greater than 1 darken it. Instead of using a fixed value, an adaptive gamma correction strategy was employed to cater the varying image darkness. In this case γ is computed dynamically for each image based on the Mean Pixel Intensity (MPI):

$$MPI = \frac{1}{N \times M} \sum_{i=1}^{N} \sum_{j=1}^{M} I(i,j) \tag{2}$$

where,

- $N \times M$ is the size of image,

- $I(i,j)$ is the intensity at each pixel.

The gamma correction factor is further calculated by using Eq. (3) given below.

$$\gamma = 2.5 - \frac{MPI}{255} \tag{3}$$

- For dark images (low MPI), γ approaches 2.5 $\rightarrow$ stronger brightening.
- For brighter images, γ decreases $\rightarrow$ minimal enhancement.

Implementation Notes:

- Gamma is applied to each channel independently in RGB images.
- γ is constrained to [1.0, 2.5] to avoid overexposure.

4.2.2 Gaussian filtering (noise suppression)

Purpose: Gamma correction improves the brightness of the image, but removal of noise artifact requires further processing of the image using Gaussian filtering. It is required to remove noise while preserving critical features such as object edges in the image. The Gaussian filter is a low-pass filter. It smooths the image by convolving with a Gaussian kernel as per Eq. (4).

Filtering Operation:

$$G(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{4}$$

where,

- σ represents the standard deviation.
- (x,y) represents the pixel coordinates of the image.

Parameter Configuration:

- Kernel size: 5×5.
- Standard Deviation (σ): 1.0.

The above parameter configuration is chosen experimentally to ensure the suppression of high frequency noise (especially from dark regions) and preservation of object edges for reliable bounding box prediction. It is observed that selection of standard deviation and kernel size is optimum to maintain the essential firearm-related details. This step ensures that YOLOv11n focuses on actual object features rather than pixel variations introduced during gamma correction.

Implementation Note: Filtering is applied post gamma-correction only.

4.2.3 Min-max normalization (contrast enhancement)

Purpose: To map pixel values within the range of 0 to 255, Min-Max Normalization is applied. This has been done to standardize image intensity distribution and enhance contrast further. Normalization ensures that images maintain a consistent dynamic range without overexposure or underexposure.

Normalization Operation: The normalization function is defined as Eq. (5).

$$I_{Norm}(x,y) = \frac{I_{gauss}(x,y) - I_{min}}{I_{max} - I_{min}} \times 255 \tag{5}$$

where,

- I_{max} and I_{min} are the minimum and maximum intensity values of the Gaussian-filtered image $I_{gauss}(x,y)$.
- $I_{Norm}(x,y)$ is final enhanced output image.

The transformed image $I_{Norm}(x,y)$ is scaled to occupy the full intensity range while preserving structural information. By performing normalization, the image's contrast is optimized, improving the distinction between firearms and the background. This ensures that YOLOv11 receives a well-balanced input for more accurate object detection.

4.3 Architecture of proposed model

The proposed firearm detection system integrates a preprocessing enhancement pipeline with the YOLOv11n object detector to ensure accurate and efficient firearm detection under varying illumination conditions. Figure 3 illustrates the end-to-end architecture and the adaptive processing logic used for image handling. This block diagram illustrates the adaptive bifurcation of the input based on MPI. It shows the three-stage enhancement pipeline feeding into YOLOv11n and emphasizes the distinction between processing paths for dark and well-lit images

4.3.1 Overview of the processing flow

The proposed model processes each input image through a series of stages depending on the mean brightness level, determined using the Mean Pixel Intensity (MPI). Based on MPI, the image follows one of two paths:

(a) Input and Classification (Brightness-based Routing)
The system accepts RGB CCTV image frames. Each image is converted to grayscale (if needed) and its MPI is computed. A threshold of MPI = 55 is used:
- If MPI ≤ 55: Image is considered dark.
- If MPI > 55: Image is considered well-lit.

This classification allows the model to skip unnecessary enhancement for already bright images which saves computation time and preserves original features.

(b) Image Enhancement Path (For Dark Images)
If the input image is dark, the following three stages are applied sequentially:
- Adaptive Gamma Correction: Dynamically brightens the image using MPI-based gamma tuning.
- Gaussian Filtering: Removes noise and smooths

transitions without blurring firearm edges.
- Min-Max Normalization: Ensures consistent contrast and dynamic range.

This produces a visually enhanced image suitable for robust detection.

(c) Shortcut Path (For Well-Lit Images): If the image is bright enough, only Gaussian filtering is applied. Gamma correction and normalization are skipped to preserve natural lighting and minimize unnecessary processing.

(d) YOLOv11n Detection Layer: The filtered image is passed into the YOLOv11n model. Detection heads process multi-scale features (P3–P5) and return:
- Bounding boxes,
- Objectness scores,
- Class confidence (firearm detection only, as it is a single-class task).

(e) Output Layer: The model overlays bounding boxes on the input image to mark detected firearms. Each detection is displayed with a confidence score. The results can be saved locally or forwarded to a live monitoring system or alert mechanism.

4.4 End-to-end firearm detection procedure

The proposed firearm detection system is designed to operate adaptively based on lighting conditions to ensure high detection accuracy and real-time performance. The overall workflow begins with the acquisition of RGB surveillance images from CCTV feeds. Each input image is first converted to grayscale, and its Mean Pixel Intensity (MPI) is calculated to determine whether the image is captured in a dark or well-lit environment. If the image is classified as dark (MPI ≤ 55), it is passed through a three-stage image enhancement pipeline comprising:
- Adaptive Gamma Correction to improve brightness,
- Gaussian Filtering to suppress noise while preserving firearm contours,
- Min-Max Normalization to enhance contrast and prepare the image for detection.
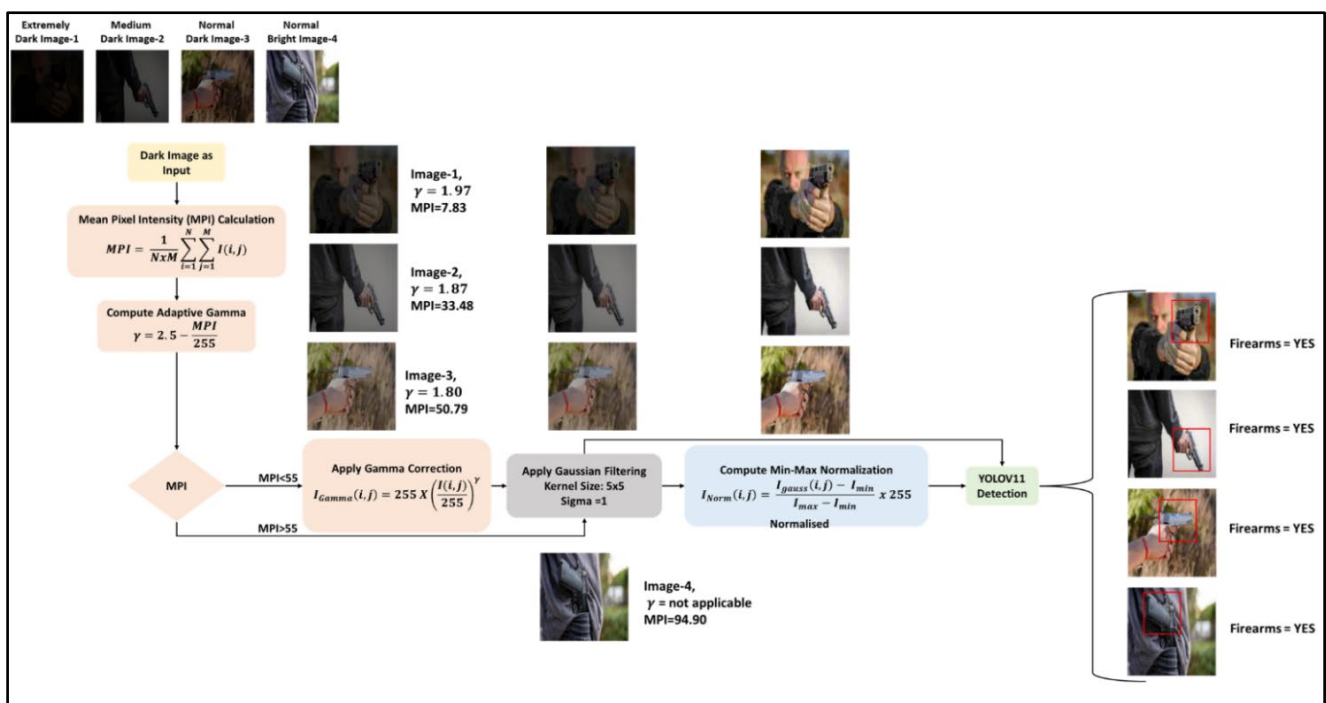
Figure 3. Architecture of the proposed firearm detection system
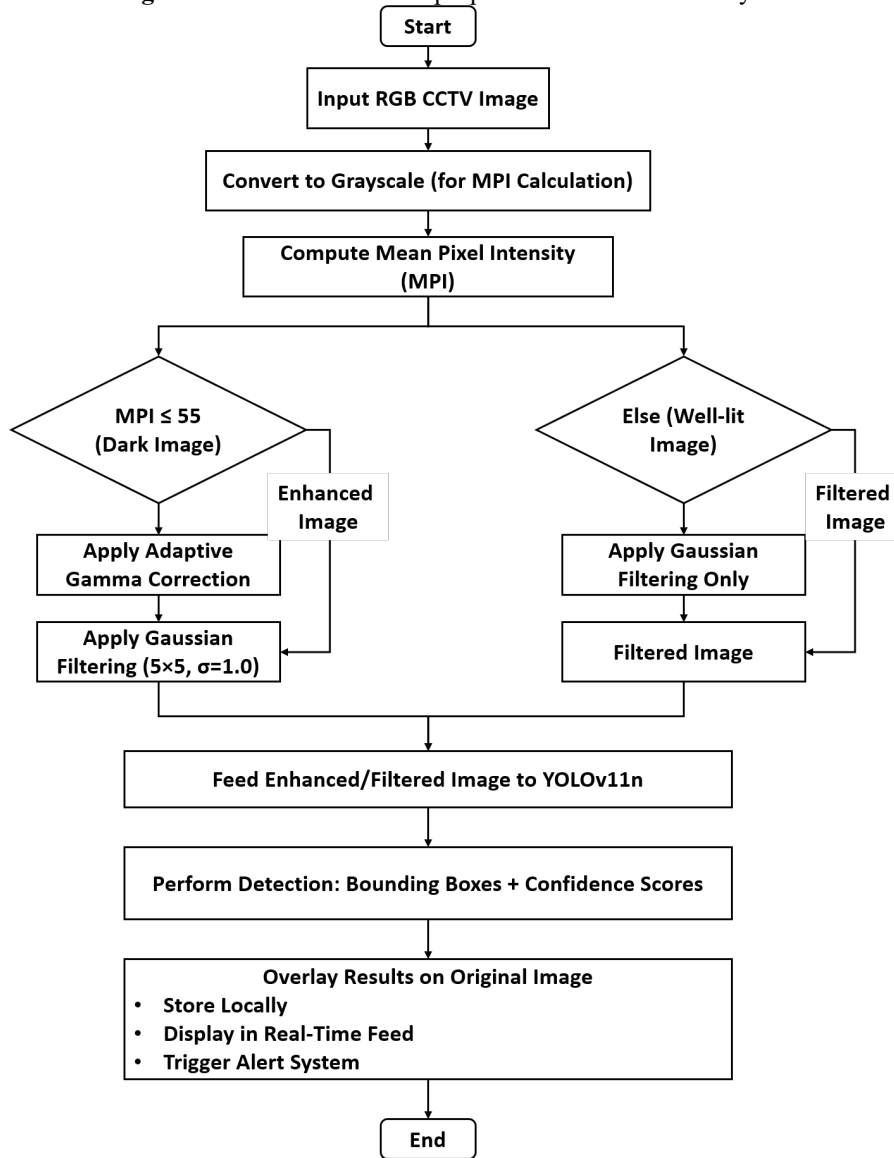
Figure 4. Complete workflow of the proposed model

In contrast, well-lit images (MPI > 55) bypass the gamma correction and normalization stages to retain natural brightness and undergo only Gaussian Filtering for basic smoothing. The enhanced or filtered image is then fed into the YOLOv11n object detection model. YOLOv11n performs multi-scale inference and predicts bounding boxes along with confidence scores for firearm presence. The output image is annotated with detection results and may be either:

- Stored locally for further analysis,
- Displayed on a monitoring interface for live tracking,
- Or used to trigger automated alerts for immediate response by security personnel.

The complete workflow of the proposed model is visually summarized in the flowchart as shown in Figure 4. Flowchart illustrating the end-to-end firearm detection procedure of the proposed system. The workflow begins with input image acquisition and grayscale conversion, followed by the computation of Mean Pixel Intensity (MPI). Based on the MPI threshold, the system adaptively routes each image through either a three-stage enhancement pipeline (for dark images) or a simplified filtering path (for well-lit images). The processed image is then fed into the YOLOv11n model for firearm

detection. The final outputs bounding boxes with confidence scores are visualized on the original image and can be stored, displayed in real-time, or used to trigger alerts in surveillance systems.

5. RESULTS AND DISCUSSIONS

Having described the design and implementation of the proposed firearm detection system, the next section presents a comprehensive evaluation of its performance. The results are structured to demonstrate how the three-stage enhancement pipeline improves detection accuracy across varied lighting conditions. Both quantitative metrics and qualitative visualizations are used to validate the system's effectiveness in achieving robust, real-time firearm detection in low-light environments.

5.1 Evaluation metrics

To assess the performance of the proposed firearm detection model under varying lighting conditions, a standard set of

classification and object detection metrics is used:

- True Positive (TP): The number of images correctly identified as containing firearms.
- False Positive (FP): The number of images incorrectly classified as containing firearms.
- True Negative (TN): The number of images correctly identified as not containing firearms.
- False Negative (FN): The number of images containing firearms but classified as not having them.

From these metrics proposed model Accuracy, Precision, Recall, F1-score, and mean Average Precision (mAP) are derived which are computed using Eqs. (6) to (9).

Precision (P): Indicates how many of the predicted firearm detections are actually correct.

$$Precision\ (PR) = \frac{TP}{TP+FP} \tag{6}$$

Recall (R): Measures how many actual firearm instances were correctly detected.

$$Recall\ (RC) = \frac{TP}{TP+FN} \tag{7}$$

F1-Score: Harmonic mean of precision and recall, providing a balance between them.

$$F1\ Score = 2 \times \frac{PR\ x\ RC}{PR+RC} \tag{8}$$

Mean Average Precision (mAP@0.5): Represents the average precision computed at an IoU (Intersection over Union) threshold of 0.5, which evaluates the quality of bounding box predictions. Since this is a single-class (firearm) task, the mAP is directly computed over the firearm class only.

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i$$
$$Since\ N = 1\ (for\ Single\ class);$$

Hence,

$$mAP = \sum_{i=1}^{N} AP_i \tag{9}$$

Intersection over Union (IoU): Measures the overlap between the predicted bounding box and the ground truth box:

$$IoU = \frac{Area\ (Intersection)}{Aera\ (Union)} \tag{10}$$

These metrics collectively provide a quantitative evaluation of how accurately and robustly the model detects firearms, especially in dark or low-visibility surveillance scenarios.

5.2 Experimental setup and dataset summary

To evaluate the proposed low-light firearm detection framework, a custom dataset consisting of 3,107 images was curated from publicly available sources including Kaggle, VBS3, IMFDB, UGR, and Linksprite. The dataset comprises images containing pistols, shotguns, and handguns, representing real-world surveillance scenarios. To simulate varying low-light environments, a darkening transformation was applied to original images by scaling down their pixel intensity values using three different darkening factors. This approach effectively mimics diverse night-time or poorly lit conditions typically encountered in outdoor and perimeter surveillance. The dataset was partitioned as follows:

- 70% for training,
- 15% for validation,
- 15% for testing.

Images were further categorized into three lighting condition groups based on Mean Pixel Intensity (MPI):

- S1 – Dark Conditions: MPI ≤ 20
- S2 – Moderate Darkness: 20 < MPI ≤ 55
- S3 – Normal Lighting: MPI > 55

For model performance comparison:

- A baseline YOLOv11n model (without enhancement) was trained and tested on all image categories.
- The proposed model integrated a three-stage enhancement pipeline (adaptive gamma correction, Gaussian filtering, and min-max normalization) prior to YOLOv11n detection.

Each experiment (S1, S2, S3) involved evaluating model accuracy, precision, recall, F1-score, and mAP@0.5 under these three lighting conditions. This stratified evaluation provides clear insight into how the proposed enhancements impact detection robustness in low-light surveillance.

Table 4. Parameters observed before and after enhancement for selected images

Exp	Image	MPI	Gamma	Gaussian Kernel	Sigma	Mean Before	Mean After	Std Dev Before	Std Dev After	I_min	I_max
S1	image_33.jpg	7.38	1.97	(5, 5)	1	36.56	36.56	18.8	18.39	0	82
S1	image_32.jpg	8.41	1.97	(5, 5)	1	40.08	40.09	21.22	20.97	0	83
S1	image_8.jpg	12.71	1.95	(5, 5)	1	50.59	50.59	17.44	16.74	0	79
S2	image_55.jpg	32.18	1.87	(5, 5)	1	77.5	77.5	33.77	32	0	136
S2	image_82.jpg	29.26	1.89	(5, 5)	1	73.88	73.88	33.76	32.43	0	137
S2	image_88.jpg	33.31	1.87	(5, 5)	1	81.23	81.23	24.7	24.33	0	134
S3	image_55.jpg	107	Skipped	(5, 5)	1	106.64	106.64	64.89	61.37	Skipped	Skipped
S3	image_93.jpg	38.44	1.85	(5, 5)	1	83.21	83.21	40.36	38.96	4	255
S3	image_75.jpg	145.53	Skipped	(5, 5)	1	145.63	145.63	82.89	82.06	Skipped	Skipped

Table 5. Comparative detection results for YOLOv11n with and without enhancement

Lighting Condition	Model Variant	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	mAP@0.5
S1 (Dark)	YOLOv11n (Original)	65.38	73.42	24.39	36.50	0.46
	YOLOv11n + Enhancement	85.38	84.21	52.44	65.68	0.81
S2 (Moderate-Dark)	YOLOv11n	85.71	85.71	69.23	76.53	0.84

	Model					
	(Original)					
	YOLOv11n + Enhancement	92.85	95.00	79.48	86.48	0.91
	YOLOv11n (Original)	97.14	95.23	90.00	92.53	0.92
S3 (Normal)	YOLOv11n + Enhancement	97.85	96.87	92.30	94.53	0.93

5.3 Enhancement effectiveness analysis

The effectiveness of the proposed image enhancement pipeline was evaluated by comparing key image statistics before and after preprocessing across various lighting conditions. Table 4 summarizes representative results from the three experimental categories (S1: dark, S2: moderate-dark, S3: well-lit). For dark images (S1), gamma correction significantly improved brightness while preserving structural details.

For dark images (S1), gamma correction significantly improved brightness while preserving structural details. For example, images with a Mean Pixel Intensity (MPI) as low as 7.38 were brightened using a computed gamma value of 1.97, making firearm contours more visible to the detector. In these cases, standard deviation (representing contrast spread) was preserved, indicating that enhancement did not introduce over-smoothing.

Moderately dark images (S2) showed improved mean intensities with stable contrast metrics post-enhancement. In contrast, well-lit images (S3) bypassed gamma correction and normalization, undergoing only Gaussian filtering. This ensured natural lighting was preserved while still removing high-frequency noise.

Key Observations from Table 4:
- **Gamma values adaptively decreased** as MPI increased, ensuring controlled brightness enhancement.
- **Mean pixel values remained stable** after enhancement, validating the non-destructive nature of the pipeline.
- **Standard deviation differences were minimal**, confirming preservation of contrast details.
- **I_min and I_max expansion** in darker samples reflects improved dynamic range for detection.

These enhancements result in more consistent input quality, which directly supports improved model performance, particularly in low-light settings where traditional detection approaches tend to fail.

5.4 Detection performance comparison

To assess the effectiveness of the proposed image enhancement pipeline, performance metrics were compared for the baseline YOLOv11n (without enhancement) and the enhanced YOLOv11n model across three lighting conditions S1 (dark), S2 (moderate-dark), and S3 (normal light). The key metrics considered are Accuracy, Precision, Recall, F1-Score, and mAP@0.5, as summarized in Table 5.

Key Observations from Table 5:
- In S1 (dark images), the proposed enhancement pipeline improved F1-score by 29.18%, recall by 28.05%, and mAP@0.5 by 35 points, demonstrating substantial gains in firearm detection accuracy under extremely low visibility.
- In S2 (moderate-dark conditions), improvements were also significant: a ~10% gain in F1-score and notable increases in recall and mAP.

- For S3 (well-lit conditions), only marginal improvement was observed. This validates the pipeline's adaptive nature, where enhancement is selectively applied only to necessary cases, thus preserving original image quality.

5.5 Visual inspection of results

In addition to quantitative improvements, qualitative evaluation confirms the effectiveness of the proposed enhancement pipeline in enabling accurate firearm detection under challenging lighting conditions. Figures 5, 6, and 7 present representative samples from the three experimental categories:

Figure 5 showcases dark images (S1) undergoing full enhancement—gamma correction, Gaussian filtering, and normalization—followed by successful YOLOv11n detection. Step-by-step visualization of the proposed enhancement pipeline applied to dark surveillance images (S1). The sequence includes: original image, gamma-corrected image, Gaussian-filtered image, final normalized image, and YOLOv11n-based firearm detection. The enhancement stages significantly improve visibility and enable accurate localization of firearms in low-light conditions.

Figure 6 includes moderately dark images (S2) where enhancement significantly improves visibility and localization, especially in cluttered or partially occluded scenes. Demonstration of enhancement and detection results for moderately dark images (S2). The proposed preprocessing pipeline improves feature visibility, enabling YOLOv11n to accurately detect firearms that are difficult to localize in the original frames. Red bounding boxes highlight correct detections with improved edge clarity.

Figure 7 demonstrates well-lit images (S3) where gamma correction and normalization are skipped, and detection is performed directly after minimal filtering. Detection outputs for images captured under normal lighting conditions (S3). As per the adaptive pipeline, gamma correction and normalization are skipped, and only Gaussian filtering is applied. YOLOv11n performs accurately with minimal enhancement, confirming the efficiency of the adaptive strategy.

Key Observations:
- In S1, firearm regions that were not visually distinguishable in the original images were successfully detected post-enhancement. This aligns with the improved F1-scores reported in Table 5 (from 36.5% to 65.68%).
- In S2, YOLOv11n consistently detected handguns across varying poses and backgrounds after enhancement, reflecting the precision and recall improvements.
- In S3, the model performed well with and without enhancement, validating the decision to bypass unnecessary processing in naturally lit scenes.

Figure 8 further supports the model's robustness through a set of performance curves. The Precision-Confidence Curve shows a steady increase, reaching peak precision at a

confidence threshold of 0.906. The Recall-Confidence Curve indicates high recall across most thresholds, with a gradual decline near 1.0. The Precision-Recall Curve confirms consistent performance with an mAP@0.5 of 0.895. The F1-Confidence Curve reveals the model's optimal detection threshold at 0.482, where the F1-score reaches 0.84. Together, these curves demonstrate reliable model behavior and help inform threshold selection for real-time deployment.

These visual inspections confirm that the proposed enhancement strategy directly contributes to addressing the primary research challenge: robust firearm detection under low-light surveillance conditions, where traditional detectors often fail.

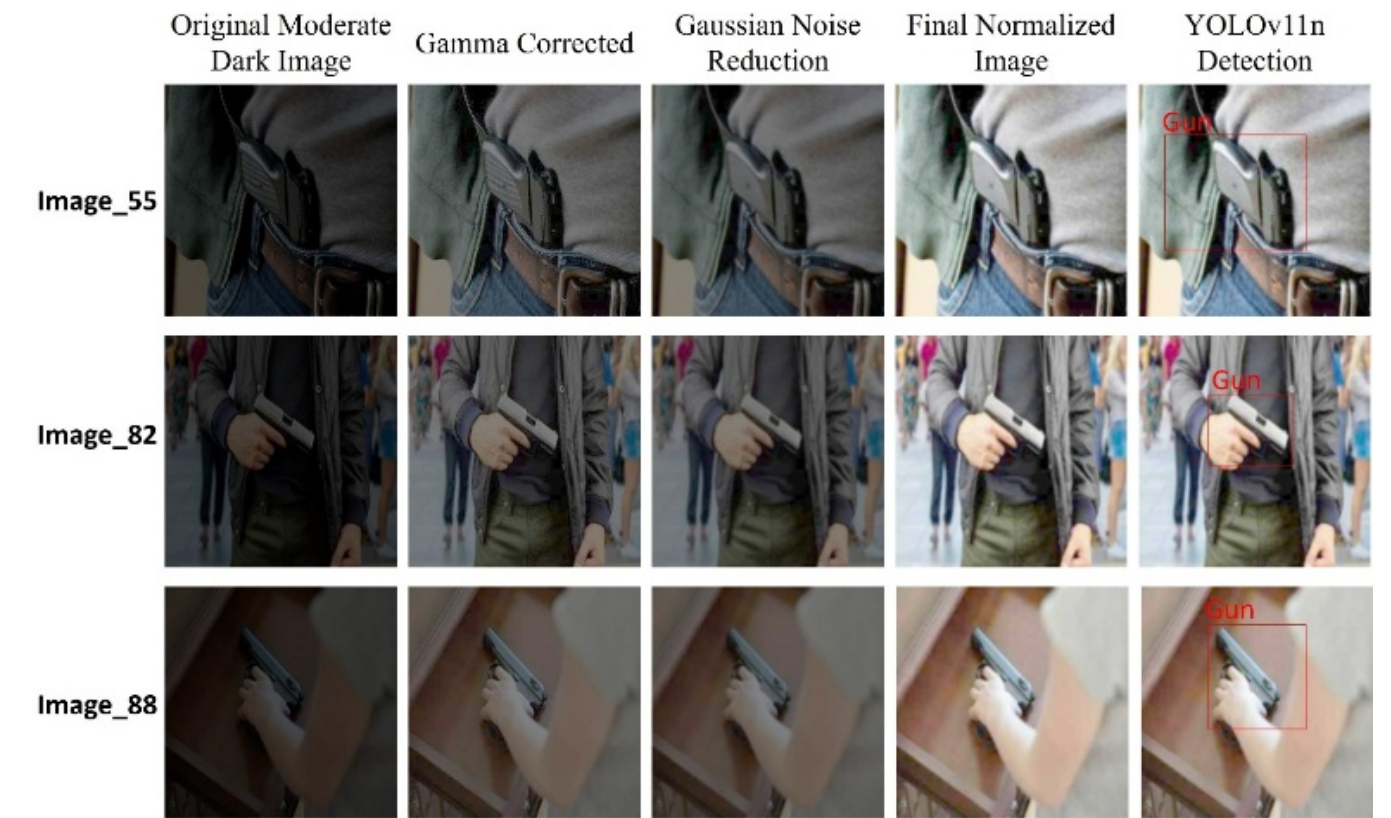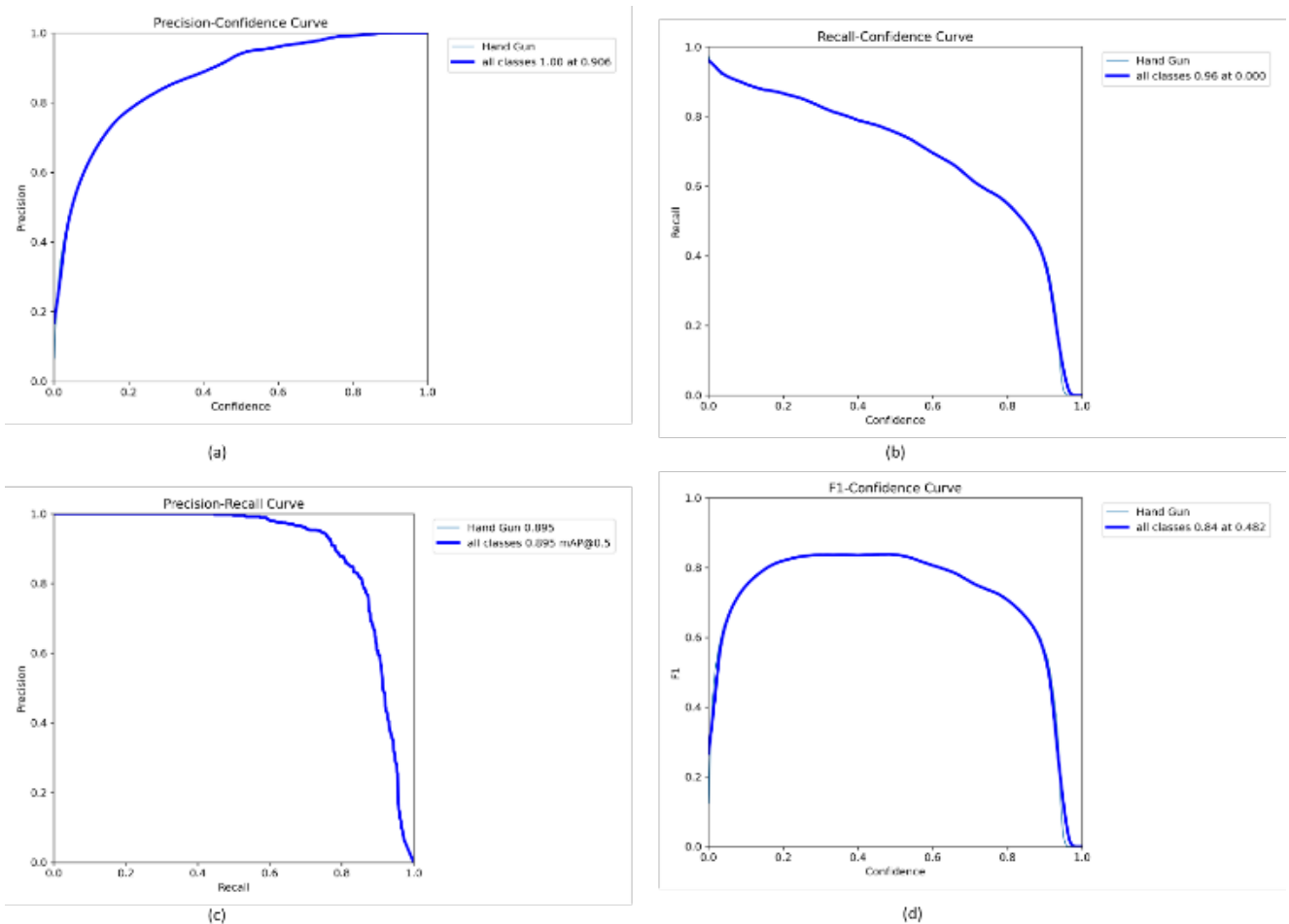**Figure 5.** Detection results in dark images (S1)

Figure 6. Detection results in moderate-dark images (S2)

Figure 7. Detection results in well-lit images (S3)

Figure 8. Model performance curves for the proposed YOLOv11n-based firearm detection system. (a) Precision vs. Confidence, (b) Recall vs. Confidence, (c) Precision vs. Recall, and (d) F1-Score vs. Confidence. These plots illustrate the detection behavior across varying confidence thresholds, highlighting the model's robustness and optimal operating point for firearm identification

5.6 Summary of findings and future directions

The experimental results, supported by both quantitative metrics and visual inspection, demonstrate that the proposed enhancement-integrated YOLOv11n framework significantly improves firearm detection performance in low-light surveillance conditions. This addresses a critical challenge faced by traditional real-time detectors, which tend to suffer from degraded accuracy in visually constrained environments. Compared to the baseline YOLOv11n, the enhanced model achieved:

- A 29.18% absolute increase in F1-score and 35-point mAP@0.5 boost in the darkest test cases (S1),
- A ~10% gain in recall and precision in moderately dark conditions (S2),
- Near-equivalent performance in well-lit settings (S3), validating the adaptive enhancement logic.

These findings are consistent with and build upon recent literature such as Yadav et al. [38], where YOLOv7-DarkVision was used to improve detection in dark environments. However, the current approach takes a step further by:

- Introducing an adaptive, conditional enhancement pipeline,
- Leveraging YOLOv11's architectural efficiencies, and
- Demonstrating robustness across multiple lighting ranges with practical computational cost.

Future Directions: To further enhance the applicability and scalability of this work, the following research avenues are suggested:

- Model Generalization to Other Threat Classes: Extend the system to detect other concealed or handheld threats (e.g., knives, explosives) in dark environments using multi-class training.
- Hardware Optimization and Edge Deployment: Port the model to edge computing devices (e.g., Jetson Nano, Coral TPU) to enable deployment in bandwidth- or power-constrained environments.
- Integration with Tracking and Alert Systems: Combine detection with temporal tracking and automated alert mechanisms to create a complete real-time threat monitoring solution.
- Federated Learning for Privacy-Sensitive Environments: Implement decentralized training strategies for sensitive locations (schools, defense zones) to enable secure, privacy-aware model updates.

By aligning detection robustness with real-time performance, this work contributes toward the development of intelligent, scalable firearm detection systems for next-generation surveillance infrastructures.

6. CONCLUSION

This study presents an enhanced firearm detection framework specifically designed for dark surveillance environments, utilizing adaptive gamma correction, Gaussian noise reduction, and min-max normalization as preprocessing stages. Experimental results demonstrate that these enhancement techniques significantly improve detection accuracy by increasing feature visibility while reducing false positives and false negatives. The comparative analysis between the baseline YOLOv11 and the proposed enhanced model highlights the effectiveness of the preprocessing pipeline, particularly in scenarios with extreme low-light conditions. Future work may focus on integrating deep learning–based denoising methods and adaptive contrast enhancement techniques to further improve detection robustness and real-time performance.

REFERENCES

[1] Gunfire on School Grounds in the United States. Available: https://everytownresearch.org/gunfire-in-school/#ns.

[2] Tessler, R.A., Mooney, S.J., Witt, C.E., O'Connell, K., Jenness, J., Vavilala, M.S., Rivara, F.P. (2017). Use of firearms in terrorist attacks: Differences between the United States, Canada, Europe, Australia, and New Zealand. JAMA Internal Medicine, 177(12): 1865-1868. https://doi.org/10.1001/jamainternmed.2017.5723

[3] Liu, Z.Y., Liu, J.W. (2023). Part-aware attention correctness for video salient object detection. Engineering Applications of Artificial Intelligence, 119: 105733. https://doi.org/10.1016/j.engappai.2022.105733

[4] Ruiz-Santaquiteria, J., Velasco-Mata, A., Vallez, N., Bueno, G., Alvarez-Garcia, J.A., Deniz, O. (2021). Handgun detection using combined human pose and weapon appearance. IEEE Access, 9: 123815-123826. https://doi.org/10.1109/ACCESS.2021.3110335

[5] Olmos, R., Tabik, S., Herrera, F. (2018). Automatic handgun detection alarm in videos using deep learning. Neurocomputing, 275: 66-72. https://doi.org/10.1016/j.neucom.2017.05.012

[6] Bhatti, M.T., Khan, M.G., Aslam, M., Fiaz, M.J. (2021). Weapon detection in real-time CCTV videos using deep learning. IEEE Access, 9: 34366-34382. https://doi.org/10.1109/ACCESS.2021.3059170

[7] Wang, G., Ding, H., Duan, M., Pu, Y., Yang, Z., Li, H. (2023). Fighting against terrorism: A real-time CCTV autonomous weapons detection based on improved YOLO v4. Digital Signal Processing, 132: 103790. https://doi.org/10.1016/j.dsp.2022.103790

[8] Basit, A., Munir, M.A., Ali, M., Werghi, N., Mahmood, A. (2020). Localizing firearm carriers by identifying human-object pairs. In 2020 IEEE International Conference on Image Processing (ICIP), Abu Dhabi, United Arab Emirates, pp. 2031-2035. https://doi.org/10.1109/ICIP40778.2020.9190886

[9] Olmos, R., Tabik, S., Lamas, A., Pérez-Hernández, F., Herrera, F. (2019). A binocular image fusion approach for minimizing false positives in handgun detection with deep learning. Information Fusion, 49: 271-280. https://doi.org/10.1016/j.inffus.2018.11.015

[10] Velasco-Mata, A., Ruiz-Santaquiteria, J., Vallez, N., Deniz, O. (2021). Using human pose information for handgun detection. Neural Computing and Applications, 33(24): 17273-17286. https://doi.org/10.1007/s00521-021-06317-8

[11] Yadav, P., Gupta, N., Sharma, P.K. (2023). A comprehensive study towards high-level approaches for weapon detection using classical machine learning and deep learning methods. Expert Systems with Applications, 212: 118698. https://doi.org/10.1016/j.eswa.2022.118698

[12] Bhatt, A., Ganatra, A. (2022). Deep learning techniques for explosive weapons and arms detection: A comprehensive review. In International Conference on

Advances and Applications of Artificial Intelligence and Machine Learning, pp. 567-583. https://doi.org/10.1007/978-981-99-5974-7_46

[13] Torregrosa-Domínguez, Á., Álvarez-García, J.A., Salazar-González, J.L., Soria-Morillo, L.M. (2024). Effective strategies for enhancing real-time weapons detection in industry. Applied Sciences, 14(18): 8198. https://doi.org/10.3390/app14188198

[14] Nercessian, S., Panetta, K., Agaian, S. (2008). Automatic detection of potential threat objects in X-ray luggage scan images. In 2008 IEEE Conference on Technologies for Homeland Security, Waltham, MA, USA, pp. 504-509. https://doi.org/10.1109/THS.2008.4534504

[15] Xiao, Z., Lu, X., Yan, J., Wu, L., Ren, L. (2015). Automatic detection of concealed pistols using passive millimeter wave imaging. In 2015 IEEE International Conference on Imaging Systems and Techniques (IST), Macau, China, pp. 1-4. https://doi.org/10.1109/IST.2015.7294538

[16] Gaus, Y.F.A., Bhowmik, N., Breckon, T.P. (2019). On the use of deep learning for the detection of firearms in X-ray baggage security imagery. In 2019 IEEE International Symposium on Technologies for Homeland Security (HST), Woburn, MA, USA, pp. 1-7. https://doi.org/10.1109/HST47167.2019.9032917

[17] Flitton, G., Breckon, T.P., Megherbi, N. (2013). A comparison of 3D interest point descriptors with application to airport baggage object detection in complex CT imagery. Pattern Recognition, 46(9): 2420-2436. https://doi.org/10.1016/j.patcog.2013.02.008

[18] Tiwari, R.K., Verma, G.K. (2015). A computer vision based framework for visual gun detection using Harris interest point detector. Procedia Computer Science, 54: 703-712. https://doi.org/10.1016/j.procs.2015.06.083

[19] Castillo, A., Tabik, S., Pérez, F., Olmos, R., Herrera, F. (2019). Brightness guided preprocessing for automatic cold steel weapon detection in surveillance videos with deep learning. Neurocomputing, 330: 151-161. https://doi.org/10.1016/j.neuco.2018.10.076

[20] Wang, C., Yeh, I., Liao, H.M. (2024). YOLOv9: Learning what you want to learn using programmable gradient information. arXiv preprint arXiv:2402.13616. https://doi.org/10.48550/arXiv.2402.13616

[21] Amado-Garfias, A.J., Conant-Pablos, S.E., Ortiz-Bayliss, J.C., Terashima-Marín, H. (2024). Improving armed people detection on video surveillance through heuristics and machine learning models. IEEE Access, 12: 111818-111831. https://doi.org/10.1109/ACCESS.2024.3442728

[22] Lim, J., Al Jobayer, M.I., Baskaran, V.M., Lim, J.M., See, J., Wong, K. (2021). Deep multi-level feature pyramids: Application for non-canonical firearm detection in video surveillance. Engineering applications of artificial intelligence, 97: 104094. https://doi.org/10.1016/j.engappai.2020.104094

[23] Yadav, P., Gupta, N., Sharma, P.K. (2024). Robust weapon detection in dark environments using Yolov7-DarkVision. Digital Signal Processing, 145: 104342. https://doi.org/10.1016/j.dsp.2023.104342

[24] Manikandan, V.P., Rahamathunnisa, U. (2022). A neural network aided attuned scheme for gun detection in video surveillance images. Image and Vision Computing, 120: 104406. https://doi.org/10.1016/j.imavis.2022.104406

[25] Chatterjee, R., Chatterjee, A. (2024). Pose4gun: A pose-based machine learning approach to detect small firearms from visual media. Multimedia Tools and Applications, 83(22): 62209-62235. https://doi.org/10.1007/s11042-023-16441-3

[26] [Dataset] Bohemia Interactive, "Virtual Desktop Training & Simulation Host". Available: https://bisimulations.com/news/press-releases/vbs3-201-release-available-download-now.

[27] [Dataset] IMFDB, Internet, Movies, Firearms database. http://www.imfdb.org/wiki/Main.

[28] Soft Computing and Intelligent Information Systems, A University of Granada research group. https://sci2s.ugr.es/weapons-detection#RP.

[29] Qi, D., Tan, W., Liu, Z., Yao, Q., Liu, J. (2021). A dataset and system for real-time gun detection in surveillance video using deep learning. In 2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Melbourne, Australia, pp. 667-672. https://doi.org/10.1109/SMC52423.2021.9659207

[30] [Dataset] Kaggle Database. "The repository contains labelled images of guns taken from Google Images., Guns Object Detection, Available: https://www.kaggle.com/datasets/issaisasank/guns-object-detection.

[31] Salido, J., Lomas, V., Ruiz-Santaquiteria, J., Deniz, O. (2021). Automatic handgun detection with deep learning in video surveillance images. Applied Sciences, 11(13): 6085. https://doi.org/10.3390/app11136085

[32] Kumar, R., Bhandari, A.K. (2023). Computationally efficient scaled clustering for perceptually invisible image intensity tuning. IEEE Transactions on Emerging Topics in Computational Intelligence, 7(5): 1584-1594. https://doi.org/10.1109/TETCI.2023.3235378

[33] Khanam, R., Hussain, M. (2024). Yolov11: An overview of the key architectural enhancements. arXiv preprint arXiv:2410.17725. https://doi.org/10.48550/arXiv.2410.17725

[34] Sowmya, T., Anita, E.M. (2023). A comprehensive review of AI based intrusion detection system. Measurement: Sensors, 28: 100827. https://doi.org/10.1016/j.measen.2023.100827

[35] Santos, T., Oliveira, H., Cunha, A. (2024). Systematic review on weapon detection in surveillance footage through deep learning. Computer Science Review, 51: 100612. https://doi.org/10.1016/j.cosrev.2023.100612

[36] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2015). You only look once: Unified, real-time object detection. arXiv preprint arXiv:1506.02640. https://doi.org/10.48550/arXiv.1506.02640

[37] YOLO Vision 2024 (YV24), powered by Ultralytics https://www.ultralytics.com/events/yolovision.

[38] Prakash, I.V., Palanivelan, M. (2024). A Study of YOLO (You Only Look Once) to YOLOv8. In Algorithms in Advanced Artificial Intelligence, pp. 257-266. CRC Press.

[39] Nesa, N. (2024). Ranks of 1-motives as dimensions of Ext^1 vector spaces. arXiv preprint arXiv:2403.04271. https://doi.org/10.48550/arXiv.2403.04271

[40] Adem, K. (2022). Impact of activation functions and number of layers on detection of exudates using circular Hough transform and convolutional neural networks. Expert Systems with Applications, 203: 117583.