Digit. Signal Process. 145 (2024) 104342

Contents lists available at ScienceDirect

Digital Signal Processing

journal homepage: www.elsevier.com/locate/dsp

Robust weapon detection in dark environments using Yolov7-DarkVision

Pavinder Yadav [a], Nidhi Gupta [b,*], Pawan Kumar Sharma [a]

[a] Department of Mathematics and Scientific Computing, National Institute of Technology, Hamirpur, Himachal Pradesh, 177005, India
[b] Department of Computer Applications, National Institute of Technology, Kurukshetra, Haryana, 136119, India

ARTICLE INFO

Keywords:
CCTV surveillance
Convolutional neural networks
Dark image enhancement
Deep learning
Weapon detection

ABSTRACT

Currently, weapon detection through video surveillance has been extensively studied using deep learning techniques by researchers. However, limited attention has been given to the detection of weapons in night time or dark scenarios. This paper aims to address this gap by proposing a novel approach for weapon detection specifically modified for low-light conditions. The authors demonstrate accurate and robust detection of weapons in challenging nighttime environments by modifying the deep learning model YOLOv7. The YOLOv7-DarkVision model has been developed by combining a brightening algorithm with the advanced image processing techniques and architecture of YOLOv7. A dataset of 15,367 images were collected for training of the model, along with five dark videos from various sources for performance evaluation. The derived detection model, which has a precision score of 95.50% and an F1-Score of 93.41%, performs absolutely well as a weapon detector.

1. Introduction

The identification of weapons from Closed-Circuit Televisions (CCTVs) is a fundamental and difficult task within the domain of computer vision. Statistics indicate that certain regions of the world, particularly countries where firearm possession is or was prohibited [13], experience a significant number of gun-related crimes. The timely identification of potential weapon-related incidents is crucial for maintaining public safety. Human operators may develop video blindness and miss up to 95% of on-screen activity after 20 to 40 minutes of monitoring [28], leading to poor productivity and up to 83% inaccurate detection.

As a result of the development of high-performance computers, researchers have recently developed a variety of deep learning based automated weapon detection systems. Object detection, in particular, has been a demanding research subject in the recent decade [18], and it has been implemented in a wide range of applications such as autonomous driving [31], stock market prediction [17], intelligent healthcare monitoring [1], smart video surveillance [33], facial recognition [25,16], and many more.

Detecting weapons in night hours or dark videos using deep learning presents numerous issues. One of the most common difficulties is low light conditions, which can significantly degrade image quality and make it difficult for deep learning models to distinguish between a weapon and other objects in the scene. Another problem is a shortage of training data, which requires customised tools and a controlled environment for gathering and annotating huge amounts of nighttime data. Furthermore, variations in weapon shapes, sizes, and colours can make it difficult to train deep learning models to recognise all types of weapons in low-light scenarios. The significant advantages of using automated weapon detection systems at night are as follows,

- **Improved public safety:** Night hours automatic weapon detection CCTV can improve public safety by detecting suspected criminal actions using guns and immediately informing law enforcement officials.
- **Reduced manual monitoring:** Automatic weapon detection can greatly reduce the requirement for human monitoring of CCTV video in night.
- **Enhanced situational awareness:** Security officers can enhance their night hours situational awareness using automatic weapon detection.

This study is an innovative effort in the field of weapon detection, specifically targeting night time scenarios. The existing literature indicates a noticeable gap in comprehensive investigations concerning

The code (and data) in this article has been certified as Reproducible by Code Ocean: https://codeocean.com/. More information on the Reproducibility Badge Initiative is available at https://www.elsevier.com/physical-sciences-and-engineering/computer-science/journals.

* Corresponding author.
E-mail addresses: pavinder_phdmath@nith.ac.in (P. Yadav), nidhi.gupta@nitkkr.ac.in (N. Gupta), psharma@nith.ac.in (P.K. Sharma).

https://doi.org/10.1016/j.dsp.2023.104342

Available online 7 December 2023

the detection of weapons in low-light or night time environments. This work introduces novel insights and methodologies to address the real challenges of nighttime weapon detection by focusing on this previously unexplored area. The findings advance security and surveillance by improving understanding of weapon detection in the variety of environmental conditions.

The presented paper contributes significantly in three focused areas. First, collection of 15,367 weapon images and carefully annotation. Second, a brightening algorithm that improves the quality of low-light videos in real time was developed. Finally, using modified YOLOv7, called YOLOv7-DarkVision a robust model for handgun detection has been proposed. The model performed well in recognising guns in a variety of circumstances, and its performance was assessed using precision, recall, and F1-Score criteria. This study offers useful insights into the design of effective deep learning-based real-time weapon identification.

The main contributions of this work can be summarized as follows:

- Developed a completely new categorised handgun dataset for the nighttime scenario, which includes 15,367 images of weapons from various sources.
- Presented an algorithm for brightening, with the goal of enhancing and restoring dark images of weapons in real-time.
- Using deep learning, a real-time weapon detection system for surveillance videos has been developed. The system demonstrates excellent performance on the validation dataset.

The subsequent sections of this article are organized as follows: Section 2 provides a comprehensive review of previous studies on handgun detection, encompassing both traditional and contemporary methodologies. Section 3 goes through the datasets and resources utilised in this study in detail. The proposed method is elucidated in Section 4, outlining the key principles and techniques employed. In Section 5, we present a summary of the conducted experiments and report the obtained results. Finally, Section 6 represents the conclusions drawn from our study along with potential avenues for future research.

2. Related work

Handgun detection traditionally focuses on concealed handguns in X-ray or millimeter wave images, commonly used for luggage control in airports. Existing methods employ various techniques such as density descriptors [5], single analysing and diagnosis [7,20,21,6], and cascade classifiers with boosting [32] to achieve high accuracy. However, these methods have limitations: they cannot detect non-metallic guns, require costly equipment like X-ray scanners, and react to all metallic objects, resulting in false positives.

The study of objects detection systems has given rise to the fascinating area of weapon detection. For the purpose of detecting weapons like handguns, knives, and assault rifles, several systems have been created, each with a unique set of shortcomings and drawbacks. The detection of pistols, knives, and other weapons is challenging due to a number of factors, including intra-class recognition, mis-classification, a dynamic backdrop, occlusion, and variable lighting conditions. This section is organized in two parts, weapon detection using classical machine learning techniques and deep learning techniques.

2.1. Weapon detection using classical machine learning techniques

The Haar Cascades were used by Żywicki et al. (2011) [34] to identify dangerous weapons such as knives. The training employed a total of 6,518 negative 1,560 positive images of knives. The real-time detection performance achieved a relatively low True Positive Rate (TPR) as 46%, hence the findings were unsatisfactory. Active Appearance Models (AAMs) [4] based knife detection was introduced by Glowacz et al. (2015) [8]. They identified tip-of-the-knife using Harris corner detection [11]. The model achieved classification accuracy of 92.50%. However, AAMs were not rotation-invariant, hence the approach only worked if the tip of knife would visible in the images.

A method using the Harris corner detection methodology and an initialised AAMs with shape-specific interest points was developed by Kmieć et al. (2012) [14]. Out of 40 positive samples, the model misidentified the knife three times. Only when the knife's tip was clearly seen in the image, the claimed technique worked. In order to create a novel method for locating weapons, Tiwari and Verma (2015) [26] employed the Harris Interest Point Detector (HIPD) and Fast Retina Keypoint (FREAK). The total accuracy of this approach is 84.26%. These technologies, however, were too time-consuming and intricate to be employed for real-time firearm identification.

2.2. Weapon detection using deep learning techniques

In Olmos et al. (2018) [19], an automated method was designed for handgun detection using the Faster R-CNN [22] model with VGG-16 as the feature extractor. The model achieved appealing results with a precision rate of 84.21% and an F1-Score of 91.43%. Castillo et al. (2019) [3] developed an automated model for identifying cold steel weapons from video surveillance. They also proposed preprocessing technique named Darkening and Contrast at Learning and Test stages (DaCoLT), which was directed towards enhancing brightness, to enhance detection accuracy. The Faster R-CNN model with InceptionResNetV2 achieved the highest accuracy with an F1-Score of 95%. However, the detection speed of the model at 1.3 frames per second limited its suitability for near-real-time operations. Vallez et al. (2019) [27] trained a handgun detector using a dataset from the University of Seville, Spain. They utilized Faster R-CNN as the CNN architecture and demonstrated that incorporating an autoencoder reduced False Positives (FP) by up to 37.9%. However, using the autoencoder along with the object detector to reduce the FP may increase the computational cost.

In a study by González et al. (2020) [9], the Faster R-CNN model with FPN and ResNet-50 was applied to a newly collected dataset from a genuine CCTV system deployed in a university campus. With an accuracy score of 88.12%, the FPN architecture demonstrated its effectiveness. However, the synthetic dataset created for training and testing purposes did not yield satisfactory results, rendering the developed model unsuitable for further utilization.

Singh et al. (2021) [23] introduced a computer vision-based approach using YOLOv4 for firearm identification. Their dataset consisted of various weapon images, including pistols, knives, machine guns, shotguns, swords, and others, which were combined into a single class called 'weapon' during training. An average loss of 1.314 and mean Average Precision (mAP) of 77.75% were attained by the model. For assessing real-time weapon detection performance, however, mAP was discovered to be insufficient. On a dataset of 8327 images (pistols and non-pistol classes), Bhatti et al. (2021) [2] employed Sliding window and region proposal/object recognition approaches with several algorithms such as Faster R-CNN, Inception-ResnetV2, SSD, and YOLOv4. With an F1_score of 91% and a mean average precision of 91.73%, YOLOv4 performed the best.

Lamas et al. (2022) [15] proposed a reproducible and traceable top-down weapon detection and pose estimation methodology. Using the Sohas weapon dataset, deep learning architectures achieved the precision score of 94.40%, with EfficientDet showing the best performance. However, the method focused solely on detecting weapons carried by individuals.

A YOLOv4-based CCTV weapon detection system is presented by Wang et al. (2023) [30] to identify small objects more effectively. The technology improves performance by training with a blend of synthetic and real images. Tamboli et al. (2023) [24] proposed a weapon identification system based on YOLOv5, SSD, and RCNN algorithms. The models were trained on 3042 weapon images from the University of Granada research group's dataset. However, the achieved mAP values were rather low: 56.20% for YOLOv5, 47.10% for RCNN, and 36.70%

Table 1
Comprehensive analysis of existing methods for weapon detection, including method applied, highlights, and shortcomings.

Publication	Approach	Highlights	Shortcomings	Suitable for nighttime detection
Żywicki et al. (2011) [34]	Haar Cascades	With more positive and negative sample images, the cascade's accuracy increases.	Because of the low true positive rate achieved, the results were unsatisfying.	No
Kmieć et al. (2012) [14] and Glowacz et al. (2015) [8]	AAMs and Harris corner detector	This method beats standard machine learning algorithms with a 92.50% TRP.	Rotation invariance does not apply to this approach. The knife tip must be visible in the images for the procedure to work.	No
Tiwari and Verma (2015) [26]	HIPD and FREAK	HIPD, FREAK, and the K-means clustering approach are all used to increase the accuracy of the colour-based segmentation.	This only works if the gun is fully visible. Guns with partly visible or blurry images fail the approach.	No
Olmos et al. (2018) [19]	Faster R-CNN	Using a two-stage deep learning methodology, this method enhances the accuracy of small weapon identification.	Despite having a greater level of accuracy, the approach is time-consuming, difficult, and costly to compute.	No
Castillo et al. (2019) [3]	InceptionResNetv2 + DaCoLT	The use of a brightness-directed reprocessing approach improves the detection quality of cold-steal weapons.	The detection speed of the model at 1.3 frames per second made it unsuitable for activities that required near-real-time performance.	Only for cold steel weapons
Vallez et al. (2019) [27]	Faster R-CNN	The utilisation of an autoencoder with Faster R-CNN as the CNN architecture reduced FPs by up to 37.9%.	Employing the autoencoder in conjunction with the object detector to mitigate false positives may result in an increase in computational cost.	No
González et al. (2020) [9]	Faster R-CNN with ResNet-50 and FPN	For the dataset that was produced, the FPN architecture showed a higher level of accuracy.	The synthetic dataset developed for training and testing reasons did not provide sufficient results.	No
Singh et al. (2021) [23]	YOLOv4	In terms of speed, YOLOv4 is superior to the other deep learning models.	The model's mAP was very low, making it unsuitable for real-time weapon recognition.	No
Bhatti et al. (2021) [2]	Faster RCNN - InceptionResnetV2, SSD, and YOLOv4	YOLOv4 is faster than the other deep learning architectures applied in this study.	Despite a recall rate of 88% using YOLOv4, the model produces a considerable number of false negatives.	No
Lamas et al. (2022) [15]	EfficientDet	The combination of weapon identification and posture estimation increases the performance of the model.	Only when a person is carrying a weapon can the technique detect it.	No
Wang et al. (2023) [30]	YOLOv4	Optimization techniques such as spatial attention module augmentation and multi-scale dilation convolution are employed to enhance object detection against complex backgrounds.	Less effective with the images with complex background and on synthetic data	No
Tamboli et al. (2023) [24]	YOLOv5, SSD, and RCNN.	The experimental results demonstrate that YOLOv5 outperforms other models in terms of its performance.	The obtained mAP values were notably low, with YOLOv5 achieving 56.20%; consequently, this model is not be suitable for the specific application of weapon detection.	No

for SSD. As a result, these models could not be adequate for this application of weapon detection.

Further, Table 1 provides a comprehensive review of existing approaches for weapon detection. It summarizes the various techniques employed, highlights the key contributions and advancements made by each method, and discusses their limitations. It is clearly noticeable that none of the work is yet suitable for the night time detection of weapons.

To work on these observed research gaps from the above mentioned Table 1, the present work introduces a novel deep learning technique specifically designed for the detection of firearms in low-light or dark scenarios, which has not been extensively explored in the existing literature. The study focuses on the development of a robust model capable of accurately identifying firearms under challenging lighting conditions. The findings would fill the gap in the literature and open up new research avenues in this important field by revealing the challenges and possible solutions for detecting firearms at night.

3. Dataset and materials

The dataset utilized for weapon detection in this research consists of a diverse compilation of 15,367 images obtained from multiple sources. These sources include real-world scenarios, publicly available datasets, named Internet Movie Firearms Database [12], and other online platforms. Each image has undergone meticulous manual annotation to identify the presence or absence of a handgun. Out of the total images, 12,060 were allocated for the training phase, while the remaining 3,307 were utilized for the validation purposes. To diversify the collection, it also includes 1,719 negative samples, which are the images that contain no weapons. This extensive dataset is helpful for testing and training machine learning algorithms, enabling them to identify and categorise various types of weapons in real-world images. Table 2 provides extensive information and description of the dataset. Additionally, Fig. 1 shows representative images from the gathered dataset for weapon detection. The images depict various scenarios and point of different views, offering a wide range of examples for training and evaluation purpose.

Table 2
Description of the collected dataset, including the number of images in training and validation subsets, and number of appearance of handguns.

Class Name	Number of Images	Train	Intense in Train	Validation	Intense in Validation
Handgun	15,367	12,060	11,384	3307	3188

Fig. 1. Sample images from collected dataset used for weapon detection.

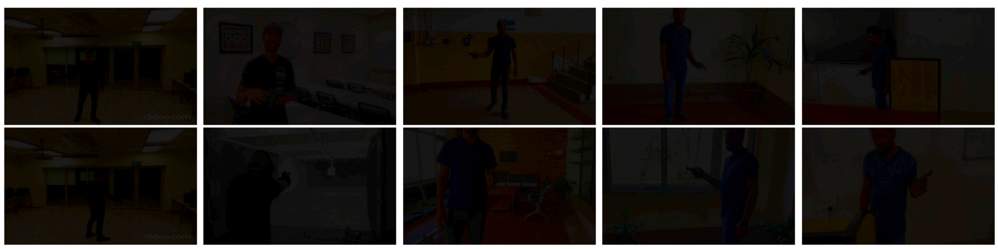

Fig. 2. Samples of dark frames that were extracted from the collected videos in order to test the performance of the model.

For evaluation purposes, a set of five videos was collected, sourced from various origins and annotated to assess the performance of the models. Each video was meticulously annotated to specifically evaluate the performance of the models in handgun detection. Notably, every video within this collection prominently features instances of handguns, aligning with the central focus of evaluation. These videos are composed of self-recorded footage from a mobile camera and recordings from platforms like YouTube and other scholarly studies. To assess the functionality and performance of YOLOv7-DarkVision, the collected videos were manually adjusted to simulate dark conditions. This allowed for thorough testing and evaluation of the performance of model to detect weapons in nighttime or low-light scenarios. Fig. 2 shows the samples of dark frames that were extracted from the collected videos in order to test the performance of the model. The description of each collected video is provided below as,

Vid-1: Dataset sourced from the study conducted by Grega et al. (2016) [10].

Vid-2: Retrieved from YouTube.

Vid-3: Self-recorded in the lobby area of the campus.

Vid-4: Self-captured in an outdoor environment.

Vid-5: Self-captured in an indoor setting.

4. Proposed methodology

The proposed approach combines the YOLOv7 deep learning model with a brightening algorithm to detect weapons effectively at night or in low-light conditions. This improves the accuracy and robustness of weapon detection in the adverse lighting circumstances by integrating modern image processing techniques and exploiting the characteristics of YOLOv7, giving useful insights for surveillance and security applications.

4.1. Deep learning model YOLOv7

YOLOv7 purportedly enhances speed as well as accuracy through various architectural reforms [29]. The model YOLOv7 supposedly relies exclusively on training with the Common Objects in Context (COCO) dataset. The architecture of YOLOv7 is shown in Fig. 3. The alleged major changes has been introduced in the YOLOv7 model, which includes the following as,

Architectural Reforms

1. Model Scaling for Concatenation based Models
2. Extended Efficient Layer Aggregation Network (E-ELAN)

Trainable Bag of Freebies (BoF)

1. Coarse for auxiliary and fine for lead loss
2. Planned re-parameterized convolution

The E-ELAN module is a computational block found within the YOLOv7 backbone. It draws inspiration from prior studies on network efficiency. The design of this module takes into consideration several factors that can affect both speed and accuracy, i.e.,

- I/O channel ratio
- Memory access cost
- Gradient path
- Activations
- Element-wise operations

YOLO-V7 employs a loss function with three parts: bounding box, objectness, and class loss. The bounding box loss evaluates coordinate positioning, objectness loss measures confidence error, and class loss accounts for category prediction discrepancies. The objectness loss is implemented as the CIoU loss, which incorporates ground truth box distance, overlap, box scale, and penalty terms. This enhances the stability of bounding box regression.

4.2. Dark image enhancement

The technique of dark image enhancement is used to improve the visibility and clarity of images acquired in low-light or dark scenarios. It involves adjusting the brightness, contrast, and other image prop-

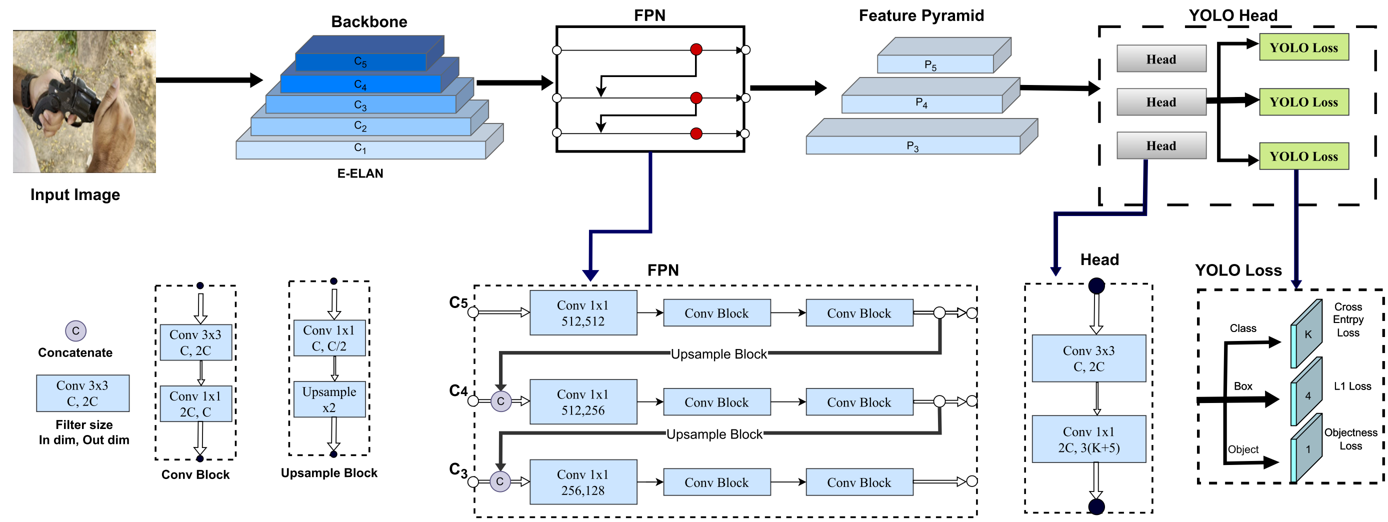

Fig. 3. Architecture of YOLOv7 [29].

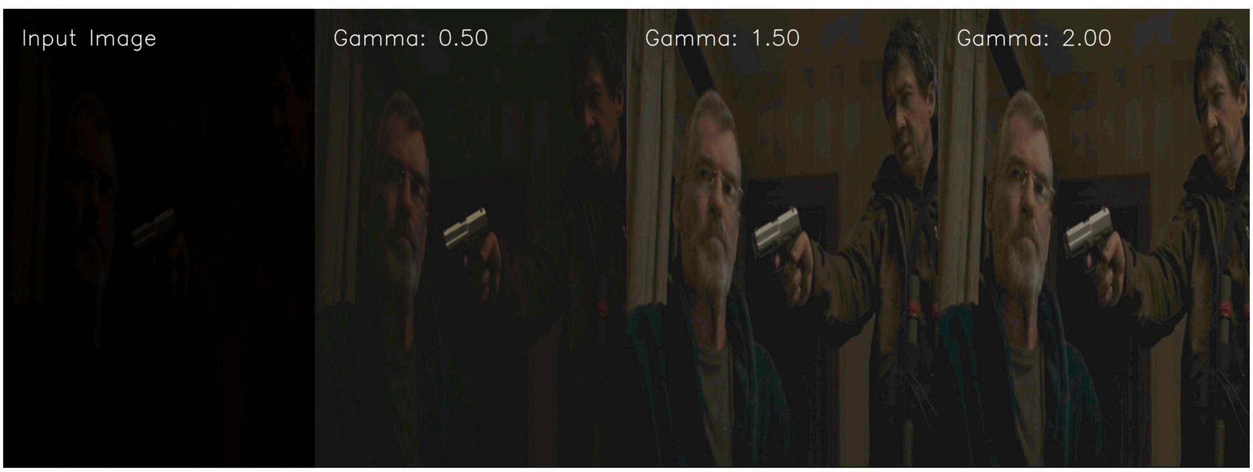

Fig. 4. Comparison of gamma correction with different gamma values. The image shows the original image and its variations with different gamma values (0.30, 0.50, 1.50, and 2.0)

erties in order to improve details, minimise noise, and reveal hidden information in dark areas.

4.2.1. Gamma correction

Gamma correction involves adjusting the brightness and contrast of an image using a non-linear transformation of its pixel values. To achieve this, the pixel intensities of the input image are first scaled from the range [0, 255] to [0, 1.0]. The output image is then obtained by applying the following equation:

$$O = I^{\frac{1}{G}} \tag{1}$$

where, O is the gamma-corrected output image, I is the input image, and G is the gamma value. The resulting image is then rescaled back to the range $[0, 255]$.

Fig. 4 illustrates the effect of gamma correction on the image with different values of gamma. The gamma values used in the correction process are 0.30, 0.50, 1.50, and 2.0.

4.2.2. Contrast and brightness

The difference in intensity between the darkest and brightest parts of an image is referred to as contrast. The brightness of an image corresponds to its overall lightness or darkness. Adjusting the brightness entails adding or subtracting a constant value from all of the image's pixel intensities.

$$\begin{aligned} I_c &= I_n \times \alpha \\ I_b &= I_c + \beta \end{aligned} \tag{2}$$

where, I_n represents the input pixel intensity, I_c represents the contrast-adjusted pixel intensity, I_b is brightness-adjusted pixel intensity, α is the contrast factor, and β is the brightness offset.

4.2.3. Gaussian blur

Gaussian blur is a popular method used in image processing to reduce noise and blur an image. This involves the application of a Gaussian filter to convolve the image, resulting in a smoothed version of the original image. The convolution of the Gaussian function with the input image is defined as,

$$I_b(i,j) = (I * G)(i,j) = \iint I(m,n) * G(i-m, j-n)\,dm\,dn \tag{3}$$

where I_b is the blurred image, I is the original image, and G is the Gaussian kernel, defined as,

$$G(m,n) = \frac{1}{2\pi\sigma^2} e^{-\frac{\left(m^2+n^2\right)}{2\sigma^2}} \tag{4}$$

where, σ is the standard deviation of the Gaussian function, m and n are the coordinates of the image.

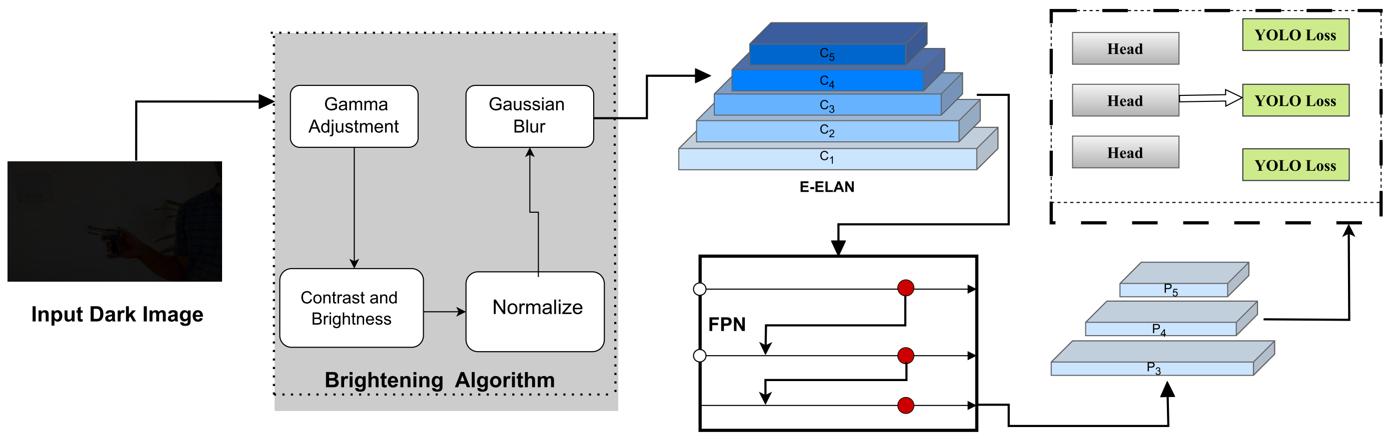

Fig. 5. Architecture of YOLOv7-DarkVision.

4.2.4. Normalization

The purpose of normalization is to enhance the contrast and dynamic range of the image, making it easier to analyse or display. The formula for min-max normalization is given by,

$$I_{norm} = \frac{I - I_{min}}{I_{max} - I_{min}} \tag{5}$$

where, I is the input image, I_{norm} is the normalized image, I_{min} is the minimum pixel value in the image, I_{max} is the maximum pixel value in the image.

4.3. Proposed architecture YOLOv7-DarkVision

A brightness algorithm has been included alongside the YOLOv7 deep learning model in order to improve its compatibility with the detection of firearms at night hours. The adoption of this algorithm successfully tackles the obstacle created by low-light environments. As a result, the YOLOv7 model is now able to reliably recognise and categorise firearms even while operating in the areas with low levels of illumination. The overall performance and reliability of the weapon identification system are greatly improved in nighttime circumstances when YOLOv7 is combined with the brightening algorithm. The architecture of YOLOv7-DarkVision is shown in Fig. 5.

4.4. Detection procedure: dark frame to handgun identification

In this section, the end-to-end procedure of handgun detection is presented, highlighting the integration of the YOLOv7 and YOLOv7-DarkVision models. The detection process begins with the user providing the model with an input, which can be an image, video, or web-cam feed.

To ensure optimal detection performance, the algorithm first determines whether the input frame is classified as dark or not. I be the input frame with H rows and W columns representing the pixel intensity values. The mean pixel value M of the frame I is calculated as follows:

$$M = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} I(i,j) \tag{6}$$

$M < 80$ dark frame
$M >= 80$ bright frame

This condition serves as the criterion to determine whether the input frame exhibits relatively low overall brightness, indicating a darker image. The threshold value of 80 was chosen to distinguish "dark" frames from non-dark frames based on empirical observations. Frames with a mean pixel value below 80 were found to generally appear darker, justifying its selection as the criterion for identifying dark frames. If the frame is deemed adequately illuminated, the YOLOv7 model is employed for the detection. This model is well-suited for handling images or videos with sufficient lighting conditions, providing accurate and efficient handgun detection.

However, in cases where the input frame is identified as dark, the YOLOv7-DarkVision model takes charge of the detection process. The YOLOv7-DarkVision model is specifically designed to handle dark frames and low-light scenarios. It begins by enhancing the input dark frame using advanced image processing techniques, effectively improving the visibility and quality of the image. Once the enhancement is applied, the YOLOv7-DarkVision model performs the detection, identifying handguns within the enhanced dark frame. Fig. 6 illustrates the step-by-step process of handgun detection procedure.

By integrating both the YOLOv7 and YOLOv7-DarkVision models, this end-to-end procedure ensures robust and accurate handgun detection across a wide range of lighting conditions. Whether the frame is well-illuminated or dark, the system adapts accordingly to deliver reliable results, enhancing the overall effectiveness of firearm detection in various scenarios.

Fig. 7 depicts a detailed visualisation of the image processing pipeline for the YOLOv7-DarkVision model. The illustration depicts the step-by-step operation of the brightening algorithm, followed by the detection of handguns in the processed frames. The visualisation shows how the image is adjusted for contrast and brightness, gamma correction, and other preprocessing steps before it is transmitted to the YOLOv7 model for handgun detection.

5. Results and discussion

In computer vision, commonly used measurements include True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). TP represents the number of images correctly labeled as positive when using classifiers to detect weapons, indicating the presence of a weapon in the input image. FP refers to instances that are incorrectly classified as positive by the classifier. TN represents the accurately classified negative images, while FN denotes the wrongly classified negative images. Performance parameters such as accuracy, precision, recall, F1-Score, and mAP are calculated using Eq. (7) to Eq. (10), respectively. Average Precision is the area under the Precision-Recall curve, and it is calculated by taking the average of Precision values at different Recall levels.

$$precision = \frac{TP}{TP + FP} \tag{7}$$

$$recall = \frac{TP}{TP + FN} \tag{8}$$

$$F1 - Score = 2 \times \frac{precision \times recall}{precision + recall} \tag{9}$$

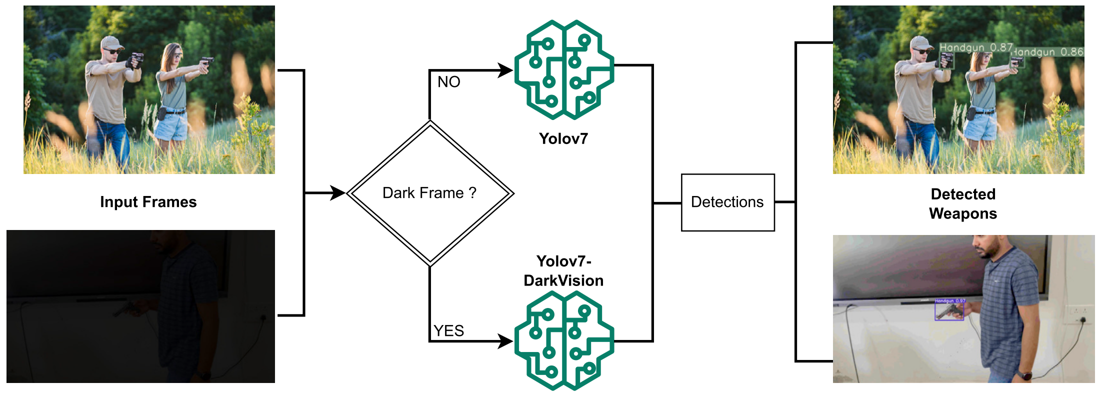

Fig. 6. End-to-end architecture of proposed methodology.

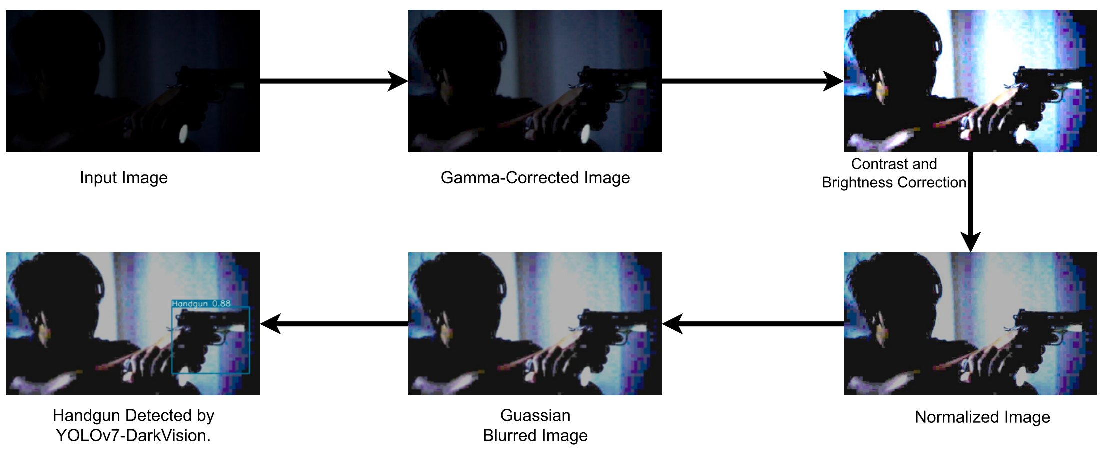

Fig. 7. Visualization of the YOLOv7-DarkVision working model. The figure showcases the image processing pipeline, including the brightening algorithm and handgun detection.

Table 3
Values of variables used in different parameters applied in the experiments E1, E2, and E3, respectively.

Parameters	Functioning	Variable Notation	Values
Gamma Correction	Applies gamma correction to adjust image brightness using a specified gamma value.	G	E1- 1.5, E2- 2.0, E3- 2.5
Contrast and Brightness	Performs linear blending of images to enhance contrast and brightness.	α and β	E1 - 3 and 10, E2 - 4 and 15, E3 - 4 and 15
Normalize	Scales pixel values of an image within a specified range to enhance visibility.	I_{min} and I_{max}	E1 - 30-250, E2 - 30-200, E3 - 30-200
Gaussian Blur	Applies Gaussian blur to reduce noise and smooth the image.	(m,n)	E1, E2 and E3 - (3,3)

$$\mathrm{mAP} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{AP}_i \tag{10}$$

where, N represents the number of classes. In this study, a single class, specifically "handgun," is the primary focus for the object detection task. Intersection over Union (IoU) is the ratio of the intersection area between the predicted bounding box and the ground truth bounding box to the union area (Eq. (11)). The evaluation metrics, mAP@0.50, represents the Average Precision at an IoU threshold of 0.50. It measures the performance of the model in detecting the target class with an IoU threshold.

$$\mathrm{IoU} = \frac{Area\ of\ Intersection}{Area\ of\ Union} \tag{11}$$

By testing the models exclusively on dark frame videos, their ability to detect and classify weapons under low-light conditions is assessed thoroughly. A number of experiments were carried out to evaluate the performance of the YOLOv7-DarkVision model at various degrees of darkness. The goal was to examine the detection accuracy of model in various low-light circumstances and establish the effect of specific parameter adjustments on its performance. Three separate experiments, labelled as experiment-1 (E1), experiment-2 (E2), and experiment-3 (E3), were created in order to examine the model's behaviour throughout a range of darkness levels. Table 3 summarizes the value of several variables of parameters used in the experimentation. In experiment E1, particular parameter values were chosen, such as a gamma of 2.0, brightness and contrast levels of 3 and 10, and normalisation values of 30 (min) and 250 (max). Similarly, for experiment E2, a gamma of 1.5 was used, as were brightness and contrast levels of 4 and 15,

Table 4
Performance comparison of YOLOv7 and YOLOv7-DarkVision on five dark frame videos.

Source	#Frames	Model	Experiment Name	Precision (%)	Recall (%)	F1-Score (%)
Vid-1	517	YOLOv7		97.72	25.28	40.17
		YOLOv7-DarkVision	E1	97.30	48.31	64.56
			E2	96.52	42.21	58.73
			E3	82.68	21.17	33.71
Vid-2	219	YOLOv7		93.10	43.42	59.22
		YOLOv7-DarkVision	E1	95.65	90.18	92.83
			E2	95.65	88.68	92.03
			E3	94.75	89.56	92.08
Vid-3	334	YOLOv7		96.20	76.68	85.34
		YOLOv7-DarkVision	E1	98.87	95.80	97.31
			E2	98.73	94.92	96.79
			E3	98.43	91.95	95.08
Vid-4	157	YOLOv7		100.00	70.15	82.46
		YOLOv7-DarkVision	E1	97.38	91.75	94.48
			E2	97.32	92.46	94.83
			E3	99.31	86.63	92.54
Vid-5	159	YOLOv7		97.12	83.68	89.90
		YOLOv7-DarkVision	E1	100.00	97.56	98.76
			E2	100.00	96.25	98.09
			E3	100.00	93.18	96.47

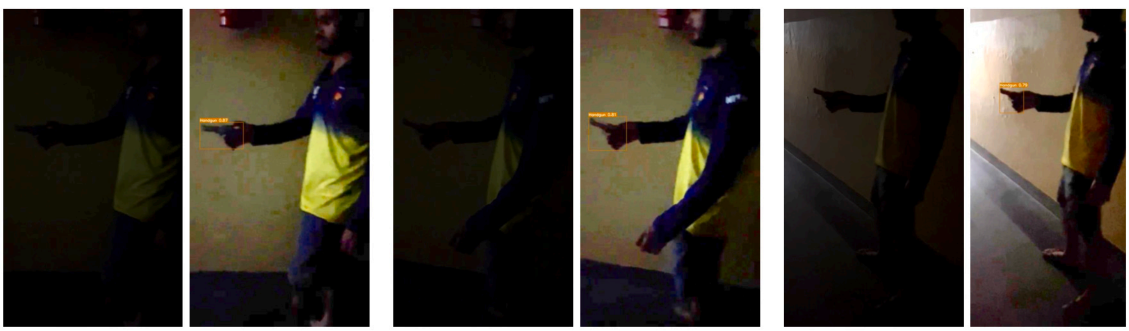

Fig. 8. Handgun detection comparison in challenging low-light conditions using the models YOLOv7 and YOLOv7-DarkVision.

Table 5
Training Results of YOLOv7-DarkVision Model

Model	Images	Labels	Precision (%)	Recall (%)	F1-Score (%)	mAP@0.5 (%)
YOLOv7-DarkVision	3307	3188	95.50	91.41	93.41	95.74

respectively, with normalisation values of 30 (min) and 200 (max). Experiment E3 used a gamma of 2.5, brightness and contrast levels of 3 and 10, and normalisation values of 30 (min) and 200 (max).

The results of these studies were rigorously summarised and are shown in Table 4, demonstrating the detection accuracy of the YOLOv7-DarkVision model at various degrees of darkness. The table also presents the performance comparison between YOLOv7 and YOLOv7-DarkVision models on five videos consisting solely of dark frames. The results provide insights into the comparative effectiveness of both models for weapon detection in dark scenarios.

Fig. 8 depicts the noticeable augmentation and adjustment done to a dark frame, demonstrating the handgun detection capabilities of both the YOLOv7 and YOLOv7-DarkVision versions. The graphic shows a clear visual comparison of how each model identifies handguns in the difficult context of low-light conditions. This figure is significant because it allows for a side-by-side comparison of detection performance, revealing light on the improved capabilities introduced by the YOLOv7-DarkVision model.

Table 5 shows the training performance of the YOLOv7-DarkVision model in terms of precision, recall, F1-Score, and mAP. It provides a comprehensive overview of the model's accuracy and effectiveness in detecting weapons in night hours or from dark scenarios. The precision, recall, and F1-Score metrics measure the model's ability to accurately identify positive instances (weapons), while mAP quantifies the overall precision across different confidence thresholds.

Fig. 9 displays multiple performance curves for the proposed weapon detection model. The curves include the precision curve, recall curve, F1-curve, and precision-recall curve. The precision curve demonstrates the precision values achieved by the model at various threshold as levels shown by Fig. 9(a). The recall curve showcases the recall values obtained by the model at different threshold levels in Fig. 9(b). The precision-recall curve illustrates the trade-off between precision and recall in Fig. 9(c). Later, in Fig. 9(d), F1-curve presents the F1-Scores achieved by the model at various threshold levels. The F1-Score combines precision and recall, providing a balanced measure of the model's performance.

The model takes dark frames as input, which typically exhibit low visibility due to inadequate lighting conditions. The YOLOv7-DarkVision model enhances the dark images using brightening algorithm, improving their visibility and quality. Subsequently, the model performs accurate and reliable handgun detection, highlighting the detected handguns in the enhanced images. The Fig. 10 illustrates the process of handgun detection in nighttime or dark scenarios using the YOLOv7-DarkVision model. This approach enables effective weapon detection even in challenging low-light environments, enhancing overall surveillance and security systems.

Gamma correction serves as crucial for changing the overall brightness and contrast of the images and ensuring that the key details and features are effectively represented. Furthermore, denoising is an important step in the brightening process since it helps to minimise the

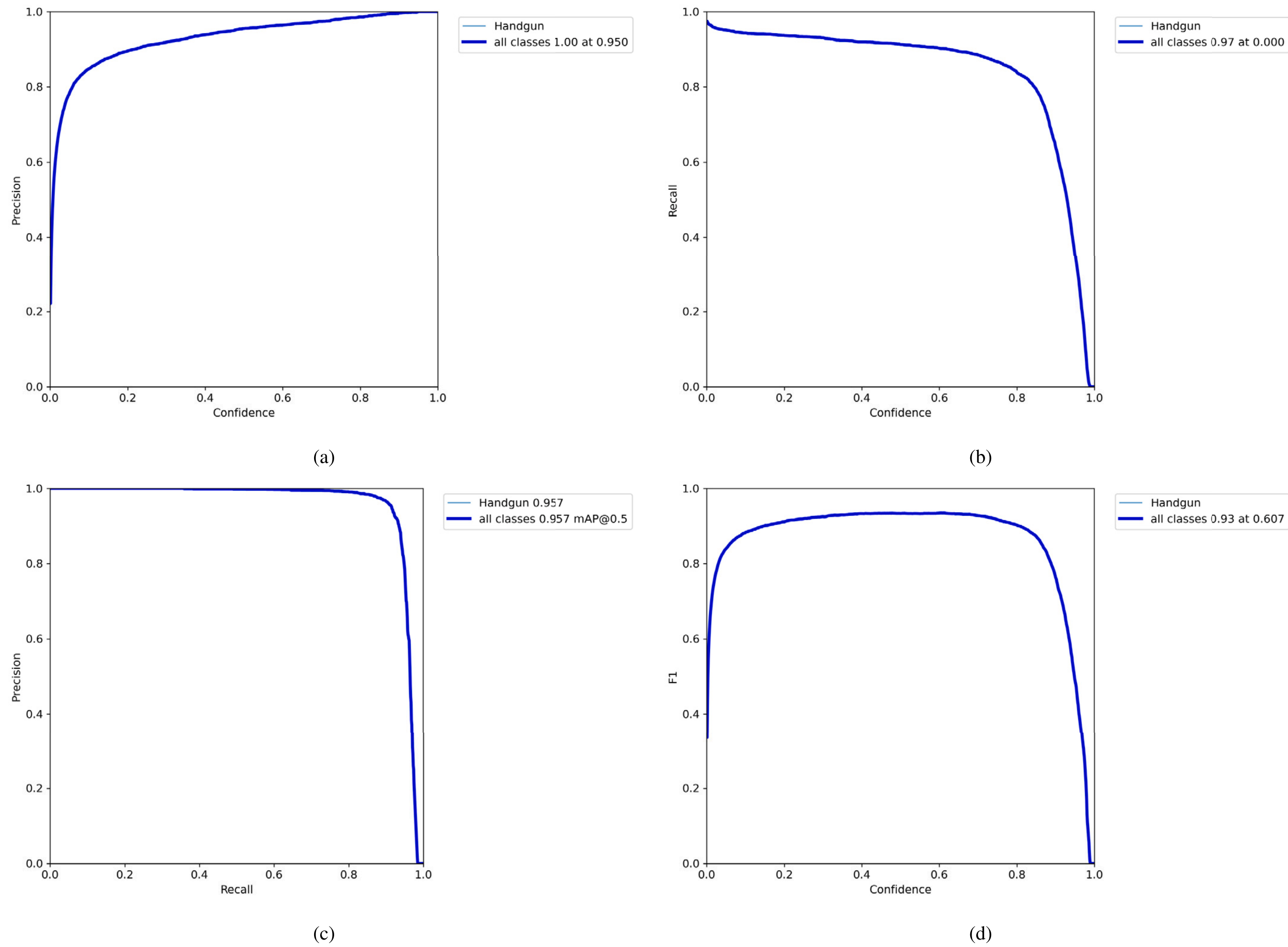

Fig. 9. Graphs depicting the various performing metrics during training: a) Precision Curve, b) Recall Curve, c) Precision-Recall Curve, and d) F1-Curve.

Fig. 10. Visualization of input dark frames and weapon detection in enhanced images. (a) Original Dark images (b) The weapon was detected in the night scene or dark images using YOLOv7-DarkVision model

amount of noise and distortions in the images. In low-light or dark environments as shown in Fig. 10(a), noise can dramatically reduce the quality and clarity of the obtained frames, making it difficult to identify weapons correctly. The combination of gamma correction and denoising in the brightening algorithm is critical for overcoming the constraints imposed by low-light or dark conditions. These procedures ensure that the enhanced images not only have increased brightness, but also adequate clarity and quality. Henceforth, after the application of Gamma correction, the model able to detect weapons from the images with higher accuracy as illustrated in Fig. 10(b).

6. Conclusion

In this study, a novel approach for accurate and robust weapon detection in night hours or from dark scenarios has been presented. The deep learning model YOLOv7 was modified and combined with a brightening algorithm, resulting in the development of the YOLOv7-DarkVision model. The extensive experiments and evaluations were conducted on a large dataset comprising 15,367 annotated images and five dark videos. The results demonstrate the effectiveness and efficiency of the proposed model in detecting handguns in challenging low-light conditions. The YOLOv7-DarkVision model achieved a high accuracy of 92.50% and exhibited superior performance in terms of precision, recall, F1-Score, and mAP. The collected dataset proved valuable in evaluating the model's performance and enhancing its capabilities.

The implications of this research are significant, addressing a crucial gap in weapon detection for night time or dark scenarios. The YOLOv7-DarkVision model can enhance security measures in applications such as video surveillance, law enforcement, and most importantly in public safety. This work contributes to the advancement of weapon detection technology, providing an effective solution for identifying handguns in challenging low-light environments. Further research could explore enhancements to the YOLOv7-DarkVision model and investigate its applicability in real-world scenarios as well. The study opens up new avenues for improved safety and security in diverse domains.

Funding information

This study received no particular funds from government, commercial, or non-profit funding entities.

CRediT authorship contribution statement

Pavinder Yadav: Conceptualization, Data curation, Methodology, Software, Writing – original draft, Writing – review & editing. **Nidhi Gupta:** Methodology, Project administration, Supervision, Writing – review & editing. **Pawan Kumar Sharma:** Project administration, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The dataset created and analysed during the present research are accessible upon reasonable request from the corresponding author through this link: https://forms.gle/B88mZqGC87vVNqZdA.

Refer to GitHub https://github.com/PavinderYadav0/DarkVision for review and access the respective codes used in this study or publication.

Acknowledgment

The presented research work has been carried out in the Computational Lab of the Department of Mathematics and Scientific Computing at the National Institute of Technology in Hamirpur, Himachal Pradesh, India.

References

[1] F. Ali, S. El-Sappagh, S.R. Islam, D. Kwak, A. Ali, M. Imran, K.S. Kwak, A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion, Inf. Fusion 63 (2020) 208–222, https://doi.org/10.1016/j.inffus.2020.06.008.

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

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

[4] T.F. Cootes, G.J. Edwards, C.J. Taylor, Active appearance models, in: H. Burkhardt, B. Neumann (Eds.), Computer Vision — ECCV'98, Springer Berlin Heidelberg, Berlin, Heidelberg, 1998, pp. 484–498.

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

[6] H. Geng, Y. Peng, L. Ye, Y. Guo, Fault identification of rolling bearing with variable speed based on generalized broadband mode decomposition and distance evaluation technique, Digit. Signal Process. 129 (2022) 103662, https://doi.org/10.1016/j.dsp.2022.103662.

[7] H. Geng, Y. Peng, L. Ye, Y. Guo, Generalized broadband mode decomposition method and its application in fault diagnosis of variable speed spherical roller bearing, Measurement 208 (2023) 112450, https://doi.org/10.1016/j.measurement.2023.112450.

[8] A. Glowacz, M. Kmieć, A. Dziech, Visual detection of knives in security applications using active appearance models, Multimed. Tools Appl. 74 (2015) 4253–4267, https://doi.org/10.1007/s11042-013-1537-2.

[9] J.L.S. González, C. Zaccaro, J.A. Álvarez-García, L.M.S. Morillo, F.S. Caparrini, Real-time gun detection in cctv: an open problem, Neural Netw. 132 (2020) 297–308, https://doi.org/10.1016/j.neunet.2020.09.013.

[10] M. Grega, A. Matiolański, P. Guzik, M. Leszczuk, Automated detection of firearms and knives in a cctv image, Sensors 16 (2016), https://doi.org/10.3390/s16010047.

[11] C. Harris, M. Stephens, et al., A combined corner and edge detector, in: Alvey Vision Conference, Citeseer, 1988, pp. 10–5244.

[12] IMFDBs, Imfdbs [Online]. http://www.imfdb.org/wiki/Main_Page. (Accessed 24 May 2023).

[13] A. Karp, Estimating global civilian-held firearms numbers, https://www.smallarmssurvey.org/resource/estimating-global-civilian-held-firearms-numbers, 2018.

[14] M. Kmieć, A. Głowacz, A. Dziech, Towards robust visual knife detection in images: active appearance models initialised with shape-specific interest points, in: Multimedia Communications, Services and Security: 5th International Conference, MCSS 2012, Krakow, Poland, May 31–June 1, 2012. Proceedings 5, Springer, 2012, pp. 148–158.

[15] A. Lamas, S. Tabik, A.C. Montes, F. Pérez-Hernández, J. García, R. Olmos, F. Herrera, Human pose estimation for mitigating false negatives in weapon detection in video-surveillance, Neurocomputing 489 (2022) 488–503, https://doi.org/10.1016/j.neucom.2021.12.059.

[16] F. Liu, D. Chen, F. Wang, Z. Li, F. Xu, Deep learning based single sample face recognition: a survey, Artif. Intell. Rev. 56 (2023) 2723–2748, https://doi.org/10.1007/s10462-022-10240-2.

[17] H. Liu, Z. Long, An improved deep learning model for predicting stock market price time series, Digit. Signal Process. 102 (2020) 102741, https://doi.org/10.1016/j.dsp.2020.102741.

[18] Z.A. Lone, A.R. Pais, Object detection in hyperspectral images, Digit. Signal Process. 131 (2022) 103752, https://doi.org/10.1016/j.dsp.2022.103752.

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

[20] Y. Peng, Z. Li, K. He, Y. Liu, Q. Li, L. Liu, Broadband mode decomposition and its application to the quality evaluation of welding inverter power source signals, IEEE Trans. Ind. Electron. 67 (2020) 9734–9746, https://doi.org/10.1109/TIE.2019.2955429.

[21] Y. Peng, Z. Li, K. He, Y. Liu, Q. Lu, Q. Li, L. Liu, R. Luo, Quality monitoring of aluminum alloy dpmig welding based on broadband mode decomposition and mmc-fch, Measurement 158 (2020) 107683, https://doi.org/10.1016/j.measurement.2020.107683.

[22] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: towards real-time object detection with region proposal networks, IEEE Trans. Pattern Anal. Mach. Intell. 39 (2017) 1137–1149, https://doi.org/10.1109/TPAMI.2016.2577031.

[23] A. Singh, T. Anand, S. Sharma, P. Singh, Iot based weapons detection system for surveillance and security using yolov4, in: 2021 6th International Conference on Communication and Electronics Systems (ICCES), IEEE, 2021, pp. 488–493.

[24] S. Tamboli, K. Jagadale, S. Mandavkar, N. Katkade, T.S. Ruprah, A comparative analysis of weapons detection using various deep learning techniques, in: 2023 7th International Conference on Trends in Electronics and Informatics (ICOEI), 2023, pp. 1141–1147.

[25] M. Taskiran, N. Kahraman, C.E. Erdem, Face recognition: past, present and future (a review), Digit. Signal Process. 106 (2020) 102809, https://doi.org/10.1016/j.dsp.2020.102809.

[26] R.K. Tiwari, G.K. Verma, A computer vision based framework for visual gun detection using Harris interest point detector, Proc. Comput. Sci. 54 (2015) 703–712, https://doi.org/10.1016/j.procs.2015.06.083.

[27] N. Vallez, A. Velasco-Mata, J.J. Corroto, O. Deniz, Weapon detection for particular scenarios using deep learning, in: Pattern Recognition and Image Analysis: 9th Iberian Conference, IbPRIA 2019, Madrid, Spain, July 1–4, 2019, Proceedings, Part II 9, Springer, 2019, pp. 371–382.

[28] S.A. Velastin, B.A. Boghossian, M.A. Vicencio-Silva, A motion-based image processing system for detecting potentially dangerous situations in underground railway stations, Transp. Res., Part C, Emerg. Technol. 14 (2006) 96–113, https://doi.org/10.1016/j.trc.2006.05.006.

[29] C.Y. Wang, A. Bochkovskiy, H.Y.M. Liao, Yolov7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors, arXiv preprint, arXiv:2207.02696, 2022.

[30] G. Wang, H. Ding, M. Duan, Y. Pu, Z. Yang, H. Li, Fighting against terrorism: a real-time cctv autonomous weapons detection based on improved yolo v4, Digit. Signal Process. 132 (2023) 103790, https://doi.org/10.1016/j.dsp.2022.103790.

[31] L. Wang, Z. Song, X. Zhang, C. Wang, G. Zhang, L. Zhu, J. Li, H. Liu, Sat-gcn: self-attention graph convolutional network-based 3d object detection for autonomous driving, Knowl.-Based Syst. 259 (2023) 110080, https://doi.org/10.1016/j.knosys.2022.110080.

[32] Z. Xiao, X. Lu, J. Yan, L. Wu, L. Ren, Automatic detection of concealed pistols using passive millimeter wave imaging, in: 2015 IEEE International Conference on Imaging Systems and Techniques (IST), IEEE, 2015, pp. 1–4.

[33] P. Yadav, N. Gupta, P.K. Sharma, A comprehensive study towards high-level approaches for weapon detection using classical machine learning and deep learning methods, Expert Syst. Appl. (2022) 118698, https://doi.org/10.1016/j.eswa.2022.118698.

[34] M. Żywicki, A. Matiolański, T.M. Orzechowski, A. Dziech, Knife detection as a subset of object detection approach based on Haar cascades, in: Proceedings of 11th International Conference "Pattern Recognition and Information Processing", 2011, pp. 139–142.

Dr. Nidhi Gupta is working as an Assistant Professor in the department of Computer Applications, at National Institute of Technology, Kurukshetra. She obtained her BSc and MSc degrees from Central University of Allahabad, Allahabad, in the year 2008 and 2010 respectively. She obtained her MTech and PhD degree (under dual degree) in the year 2018 from Indian Institute of Information Technology, Design and Manufacturing, Jabalpur. She has worked as a Post-Doctoral Fellow at National Laboratory of Pattern Recognition, Chinese Academy of Sciences, Beijing, China, under CAS Presidential International Fellowship Initiative (PiFi). She has worked as an assistant professor in the department of mathematics and scientific computing at the National Institute of Technology, Hamirpur, for three years. She has qualified National level examination like GATE and UGC NET in Computer Science. She has published more than 12 research papers in journal of repute and international conferences. She is an outstanding reviewer of several reputed SCI journals and Scopus conferences. Her research interest includes image processing, machine learning, computer vision, and pattern recognition algorithms.

Dr. Pawan Kumar Sharma is currently working as an Associate Professor and heading the department of Mathematics and Scientific Computing, in National Institute of Technology, Hamirpur. He obtained his BSc degree from Himachal Pradesh University, Shimla in 1989 and MSc degree in Mathematics Himachal Pradesh University, India, in 1993. In the year 2003, he has received Ph.D. in Mathematics from Himachal Pradesh University, Shimla, India. He has published more than 20 research publications in the field of thermo elasticity and fluid mechanics. He has more than 15 years of teaching experience and technical skills.

Pavinder Yadav received his B.Sc. degree from Govt. College, Hisar, Haryana, India, in 2015 and M.Sc. in Mathematics from DCRUST, Sonipat, Haryana, India in 2017. Currently, he is pursuing Ph.D. in the Department of Mathematics and Scientific Computing, National Institute of Technology, Hamirpur, India. He has qualified CSIR-UGC JRF in Mathematics in the year 2017 and 2018. He also has qualified GATE in the year 2019. His research topics of interest include the field of Computer Vision, Deep Learning, Machine Learning, and Image Processing.