AUTOMATED SURVEILLANCE SYSTEM FOR MILITARY BASES THROUGH INTEGRATED FACIAL RECOGNITION AND WEAPON DETECTION

Madhan S (Asst. Prof.)[1], **Subash K[2], Madhanraja R[3], Manimaran M[4]**

Department of Computer Science and Engineering, University College of Engineering, Thirukkuvalai
(A constituent College of Anna University::Chennai and Approved by AICTE, New Delhi)

---***---

Abstract - Modern military bases need the most sophisticated safety measures possible. This paper proposes a software solution that makes use of cutting-edge technology like face recognition and weapon detection to improve perimeter security. Using face recognition algorithms, the system seeks to identify possible invaders while concurrently identifying and categorizing any weapons they could be carrying. The technology is designed to detect threats and promptly notify those in authority in order to enable prompt and decisive responses. Key considerations include ensuring the accuracy, reliability, and scalability of the system, as well as addressing ethical and regulatory concerns surrounding data privacy and algorithmic biases. By integrating seamlessly with existing surveillance infrastructure, this software offers a proactive approach to safeguarding military installations, ultimately enhancing overall security and operational readiness.

Key Words: Face recognition, Weapon detection, Real-time monitoring, Military base security, Alarm system

1.INTRODUCTION

In recent years, the security landscape surrounding military installations has become increasingly complex and challenging. The safety and integrity of military bases cannot be guaranteed by traditional security measures alone due to the emergence of asymmetric threats. Innovative solutions that make use of state-of-the-art technologies to improve perimeter security and threat detection capabilities are becoming more and more necessary in response to these continuously evolving security threats.

The foundation of military base security has always been conventional security measures like man-powered patrols, surveillance cameras, and physical obstacles. These precautions have inherent limits even if they offer some safety. Human error and exhaustion are common in surveillance, and resolute opponents can overpower or get beyond static measures. Furthermore, because contemporary dangers are ever-changing, security must adopt a proactive and flexible strategy that can recognize and neutralize threats instantly.

The proposed research seeks to explore the feasibility and effectiveness of integrating facial recognition and weapon detection technologies into a unified software solution for military base security. This research is motivated by the need to address the shortcomings of existing security measures and provide military forces with the tools they need to meet the evolving security challenges of the 21st century. By harnessing the power of artificial intelligence and machine learning, the proposed software solution aims to enhance the ability of military installations to detect and deter potential threats, thereby increasing overall security and operational readiness.

Overall, the proposed research represents a timely and critical effort to advance the state-of-the-art in military base security and contribute to the broader goal of ensuring the safety and security of military personnel and assets. By doing so, it aims to empower military forces to effectively confront the threats of today and tomorrow, thereby safeguarding national security and promoting peace and stability in an increasingly uncertain world. The novelty, detailed design, implementation of the proposed system is described in detail in the subsequent sections.

2. LITERATURE SURVEY

A novel approach that combines Principal Component Analysis (PCA) with Convolutional Neural Networks (CNN) for robust face recognition provided in [1]. By employing PCA for data dimensionality reduction and CNN as a classifier, the method achieves high recognition accuracy while optimizing computational resources. Dropout regularization techniques are utilized to simplify the CNN architecture, and GPU implementation accelerates the classification process. Experimental results validate the method's effectiveness, demonstrating improved recognition accuracy, faster inference times, and reduced memory consumption compared to standard CNN implementations.

The paper [2] explores the integration of machine learning (ML) techniques in military surveillance. It investigates the limitations of current threat detection methods and examines various ML algorithms, including neural networks and deep learning, for their potential to improve accuracy and efficiency. Experimental models demonstrate the effectiveness of ML in enhancing threat detection, offering adaptive and robust surveillance systems. Ethical considerations are addressed, highlighting responsible deployment of ML in military contexts.

There are two main approaches for extracting facial features: appearance-based and model-based approaches. Rather than focusing on individual facial traits, appearance-based approaches use global representations based on the complete image to attempt to identify faces. Many techniques for computer graphics and object detection are based solely on photographs, with no need for intermediary three-dimensional models. The majority of these techniques rely on image representation, which creates a vector space structure and, in theory, necessitates dense correspondence. The goal of model-based face recognition techniques is to create a human face model that can capture facial expressions. When designing the model, prior understanding of the human face is heavily relied upon. For instance, model-based matching uses the internal facial element placement to determine the relative position and distance features [4].

Face recognition is a type of task pattern that involves classifying a face as known or unknown based on a comparison with photos of known individuals that are kept in a database. Face recognition can be challenging due to information variability resulting from random variation across individuals and systematic differences caused by diverse parameters like stance and lighting conditions [3].

The research [5] focuses on implementing automatic gun or weapon detection using convolutional neural network (CNN) based SSD (Single Shot Multibox Detector) and Faster R-CNN algorithms. Two types of datasets are utilized: one with pre-labelled images and another with manually labelled images. The results demonstrate good accuracy for both algorithms, although their application in real-world scenarios may require consideration of the trade-off between speed and accuracy. The research focuses on implementing automatic gun or weapon detection using convolutional neural network (CNN) based SSD (Single Shot Multibox Detector) and Faster R-CNN algorithms. Two types of datasets are utilized: one with pre-labelled images and another with manually labelled images. The results demonstrate good accuracy for both algorithms, although their application in real-world scenarios may require consideration of the trade-off between speed and accuracy.

[6] addresses the critical need for security and safety in modern society, emphasizing the importance of automatic weapon detection systems for ensuring a secure environment. Despite advancements in deep learning algorithms and hardware, real-time weapon detection remains a challenge due to factors such as angle differences and occlusions. Among the algorithms tested, YOLOv4 stands out as the most successful, achieving a mean average precision (mAP) of 91.73% and an F1-score of 91%. The results demonstrate the effectiveness of the proposed approach in real-time weapon detection, with a significant improvement in performance.

3. SYSTEM OVERVIEW

The proposed system integrates deep learning algorithms into existing surveillance infrastructure within military bases to enhance security measures through facial recognition and weapon detection capabilities.

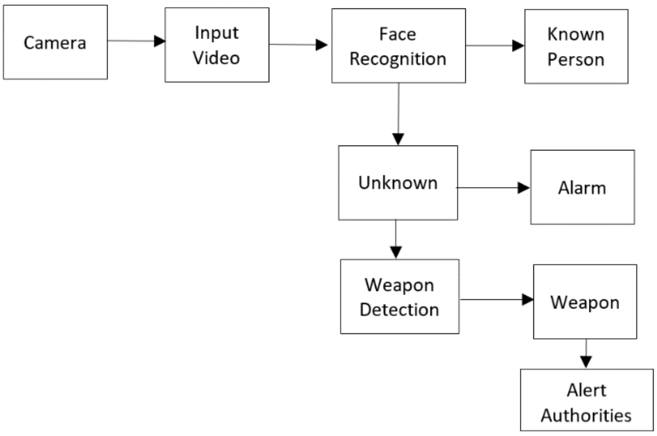


Fig-1: Block Diagram

At its core, the system comprises a network of surveillance cameras strategically positioned throughout the installation, feeding real-time video streams to a centralized processing unit. Here's a breakdown of the key components and their functionalities:

1. Surveillance Cameras: Positioned strategically throughout the military base, surveillance cameras capture footage of the surrounding area on a continuous basis. These cameras provide high-resolution imagery for analysis and are the main source of information for the algorithms used in facial recognition and weapon detection.

2. Facial Recognition Module: Upon receiving video feeds from the surveillance cameras, the facial recognition module employs deep learning algorithms to detect and identify individuals within the frame. Pre-registered facial images of authorized personnel are stored in a database for comparison against real-time video frames. Using Convolutional Neural Networks (CNNs) trained on vast datasets of facial images, the system extracts facial features, such as the arrangement of eyes, nose, and mouth, to generate unique facial embeddings. These embeddings are then compared against the database to determine potential matches, allowing the system to authenticate personnel and grant access privileges accordingly. Upon receiving video feeds, the facial recognition module analyzes the frames to detect and identify individuals. If a person is recognized as an authorized personnel based on the comparison with pre-registered facial images, the system proceeds without initiating weapon detection. However, if the individual's identity cannot be verified or is unknown, the system triggers the weapon detection module for further analysis

3. Weapon Detection Module: When this module activated, the system analyzes the video streams for the presence of weapons using deep learning-based object detection algorithms. Trained on annotated datasets containing diverse examples of firearms, explosives, and other weapons, the weapon detection module employs convolutional neural networks (CNNs) to identify and classify objects within the video frames. Then this system gives security personnel timely, actionable notifications.

4. Alerting and Response Mechanism: Upon detecting a potential security threat, such as an unauthorized individual and the presence of a weapon, the system triggers an immediate alert to designated security personnel. Security personnel receive real-time information about the nature and location of the threat, enabling them to initiate appropriate response measures swiftly and effectively.

4. ALGORITHMS IMPLEMENTED

Deep Learning:

Deep learning is a subset of machine learning that utilizes artificial neural networks with multiple layers to learn from data. It has revolutionized various fields, including computer vision, natural language processing, and speech recognition, by enabling machines to automatically learn representations of data directly from raw inputs.

In deep learning, a computer model learns to perform classification tasks directly from images, text, or sound. Deep learning models can achieve state-of-the-art accuracy, sometimes exceeding human-level performance. Models are trained by using a large set of labeled data and neural network architectures that contain many layers.

At the heart of deep learning are artificial neural networks, which are inspired by the structure and function of the human brain. A neural network consists of interconnected nodes (neurons) organized into layers. Each neuron receives input signals, performs a computation, and passes the result to the next layer of neurons.

Convolutional Neural Networks (CNNs):

CNNs are a specialized type of neural network designed for processing structured grid-like data, such as images. They are particularly well-suited for computer vision tasks, including face recognition. CNNs use convolutional layers to extract features from input images hierarchically, capturing spatial patterns and relationships between pixels.

Different layers of CNN

Input Layer:

The input layer receives the raw input data, which consists of images captured by surveillance cameras within the military base. Each image represents a frame of the video feed, containing visual information about the environment. Each image is represented as a grid of pixel values, with each pixel corresponding to a specific color channel (e.g., red, green, blue).

Convolutional Layer:

Convolutional layers apply a set of learnable filters to the input image, performing element-wise multiplication between the filter weights and the pixel values of the input image. This operation produces feature maps that capture spatial patterns and local features present in the input data.

In the facial recognition part, the convolutional layers would analyze the input images to detect facial features such as eyes, nose, and mouth. These layers learn to extract low-level features like edges and textures, which are gradually combined to form higher-level representations of facial features.

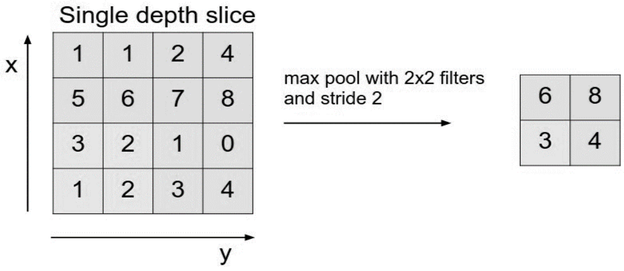
In the weapon detection part, the convolutional layers would analyze the input images to detect the presence and location of weapons. These layers learn to identify patterns associated with weapons, such as their shapes and textures.

Activation Layer (ReLU):

The ReLU activation function introduces non-linearity into the network, allowing it to learn complex relationships in the data. It is applied element-wise to the feature maps generated by the convolutional layers.

Pooling Layer:

The pooling layers downsample the feature maps obtained from the convolutional layers, reducing their spatial dimensions while retaining the most important information. This helps in reducing computation and preventing overfitting. In the below example, we have applied max pooling in single depth slice with stride of 2. You can observe the 4 x 4 dimension input is reduce to 2 x 2 dimension. As a result, the output feature maps from the pooling layers have reduced spatial dimensions compared to the input feature maps. This downsampling process helps in reducing the computational burden of subsequent layers and improving the network's ability to generalize to unseen data. The output feature maps from the pooling layers serve as compact representations of the input images, facilitating subsequent layers in making accurate predictions about the presence or absence of weapons.



1	1	2	4
5	6	7	8
3	2	1	0
1	2	3	4

6	8
3	4

Fig-2: Pooling

Flattening Layer:

The flattening layer reshapes the output from the previous layers into a one-dimensional vector, preparing it for input to the fully connected layers.

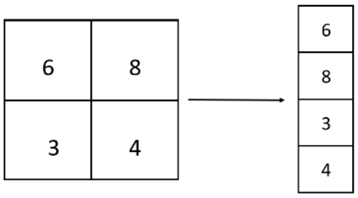


Fig- 3: Flattening

Fully Connected Layer (Dense Layer):

Fully connected layer involves weights, biases, and neurons. It connects neurons in one layer to neurons in another layer. It is used to classify images between different categories by training.

In the facial recognition part, the fully connected layers analyze the flattened feature vector to learn high-level representations of facial identities. Each neuron in the output layer corresponds to a specific individual, and the network learns to assign high probabilities to the correct identities.

In the weapon detection part, the fully connected layers analyze the flattened feature vector to classify whether a weapon is present in the input image. The output layer produces a binary classification indicating the presence or absence of a weapon.

Output Layer:

The output layer produces the final predictions or classifications based on the features learned by the previous layers. The number of neurons in the output layer corresponds to the number of classes in the classification task, with each neuron representing the probability of belonging to a particular class (e.g., softmax activation for multi-class classification). In the facial recognition, the output layer produces the identity of the recognized individual. In the weapon detection part, the output layer produces a binary classification indicating the presence or absence of a weapon.

5. SYSTEM IMPLEMENTATION

Development of the Surveillance System:

The facial recognition and weapon detection algorithms, based on Convolutional Neural Networks (CNNs), have been developed and trained using appropriate datasets. These algorithms have been integrated into a comprehensive surveillance system designed to analyze real-time video streams captured by surveillance cameras deployed within military bases.

Hardware and Software Requirements:

The system implementation involves identifying the hardware requirements, such as computing devices (e.g., servers, GPUs) capable of running the deep learning algorithms efficiently.

The software requirements, including programming languages (e.g., Python), deep learning frameworks and supporting libraries, are specified for developing and deploying the system.

Integration with Existing Infrastructure:

The surveillance system is seamlessly integrated with the existing infrastructure within military bases, including surveillance cameras, networking equipment, and access control mechanisms.

Compatibility with existing security protocols and systems is ensured to facilitate smooth integration and interoperability.

Real-Time Video Analysis:

The deployed system continuously analyzes real-time video streams captured by surveillance cameras, detecting and identifying individuals through facial recognition and detecting weapons through object detection techniques.

The system processes the video data in real-time, ensuring timely detection and response to potential security threats within the military base.

Alerting and Response Mechanisms:

Upon detection of unauthorized individuals or the presence of weapons, the system triggers immediate alerts to designated security personnel.

Alerts may be delivered through various communication channels, including mobile devices, desktop notifications, and centralized command consoles, enabling security personnel to respond swiftly to security incidents.

User Interface and Control Center:

The system is equipped with a user-friendly interface accessible to security personnel for monitoring and managing security operations.

The interface provides real-time visualization of surveillance camera feeds, overlaid with annotations indicating identified individuals and detected weapons, facilitating situational awareness and decision-making.

6. METHODOLOGY

Python was used to implement in our system. Python has a rich ecosystem of libraries and frameworks for various tasks, including deep learning, computer vision, and email communication. Libraries like TensorFlow, OpenCV, NumPy are widely used for tasks relevant to this system.

Library files are collections of functions and small execution codes. This library files will assist us in performing all of the necessary steps of object detection and image processing. We use important library files such as Numpy, Tensor Flow, OpenCV, Keras, and others in this system. These libraries will aid in making our deep learning model more efficient and adaptable for processing real-time images or videos.

Specific libraries that we will use for data preprocessing, which are:

1. Numpy
2. Tensor flow

3. OpenCV
4. Keras

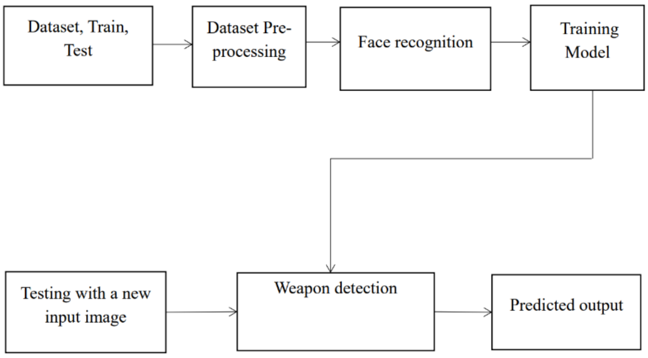


Fig-4: Process of Detection

Dataset Collection and Labeling:

First, we have to gather a diverse set of images containing faces of authorized personnel who are expected to be present in the military base. It's important to capture images of individuals under different lighting conditions, angles, and facial expressions to improve the robustness of the facial recognition model. Label each image with the corresponding identity of the person depicted. This labeling step is crucial for training the facial recognition model to associate each face with its respective identity.

Collect a wide range of pictures and videos showing different kinds of weaponry that could be dangerous if found within the military base. This includes firearms, knives, explosives, and other prohibited things.

Training the Dataset:

During the training of CNN, the neural network is being fed with a large dataset of images being labelled with their corresponding class labels. The CNN network processes each image with its values being assigned randomly and then make comparisons with the class label of the input image.

Face Recognition

First of all, the system access the webcam to capture video frames. This webcam serves as the input source for the face recognition system, allowing it to analyze real-time video streams. The system loads a reference image containing the faces of authorized personnel. This image serves as the template for training the face recognition model to recognize specific individuals. Then it generates facial encodings for the known faces in the reference image. These encodings represent unique numerical representations of facial features extracted from the reference faces. System captures a each and every frame from the video stream obtained from the webcam. Each frame contains potentially detectable faces that the system will analyze for recognition. Captured frames Resized into 1/4 of its original size. This resizing step helps optimize face recognition processing by reducing the computational workload while maintaining sufficient image quality for accurate detection. Color format of the resized frame converted from BGR (used by OpenCV) to RGB (used by the face recognition library). This conversion ensures compatibility between the frame and the face recognition algorithm. Face detection algorithm utilized the to locate and identify faces within the resized frame. The process of extracting facial features from images is facilitated by the ‘face_recognition’ library, which utilizes pre-trained deep learning models to identify and encode facial landmarks and features. These features are then compared to known faces to recognize and identify individuals. This step identifies potential face regions for further analysis. Detected faces compared with the known faces by computing their facial encodings. If a match is found between a detected face and a known face, the system recognizes the individual associated with that face. System determines the recognized individual based on the matching results. If a match is successful, the corresponding name or identity to the recognized face will be assigned. Recognized faces will be visualized by drawing bounding boxes around them and displaying their associated names or identities. Bounding boxes are drawn around detected faces using OpenCV to visually highlight the regions of interest within the image. This provides real-time feedback on the face recognition process to the user or system operator. After completing the face recognition process, the system is then ready to proceed with weapon detection.

Weapon Detection:

If the intruder is found inside the base, then the system looks for if the intruder has any weapons. First of all, the system access the webcam to capture video frames, which will be analyzed for the presence of weapons. It continuously loop over each frame captured from the webcam to detect weapons in real-time. The trained object detection model utilized to identify and localize weapons within the video frames. The object detection model scans each frame and predicts bounding boxes around objects that resemble weapons based on the learned features. It applies a threshold to the detection scores associated with each predicted bounding box to filter out weak detections. Only the bounding boxes with detection scores above a certain threshold as potential weapons will be considered. If a bounding box with a high enough detection score is detected, an alert mechanism will be triggered to notify relevant authorities about the presence of a weapon. This alert mechanism may include activating alarms, sending messages, or any other predefined action to alert officials. Then it continues monitoring the video stream for any additional instances of weapon detection, ensuring continuous surveillance and security.

7. RESULTS

Soldiers whose identity that are already stored in the system are recognized and labeled with their corresponding names displayed directly on their faces. This system draws a bounding box around the recognized face and adds a text label containing the individual's name below the bounding boxes.

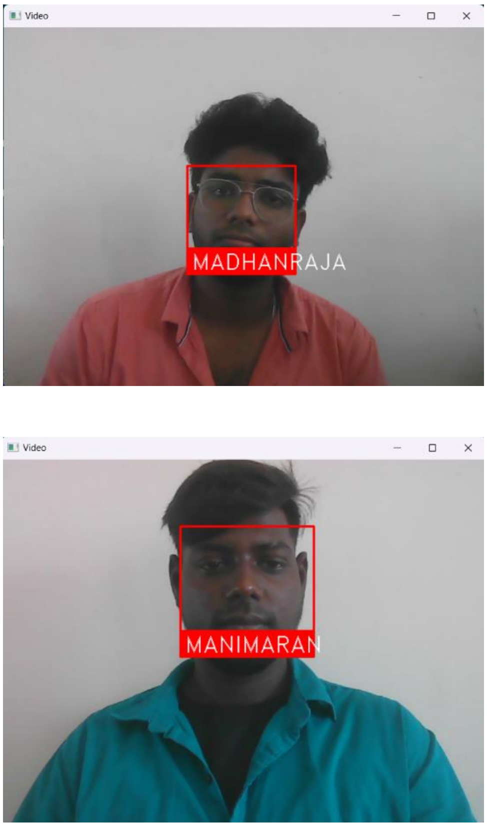


Fig-5: Real-time detection and recognition of individuals who are already stored in the system's database.

If the detected face does not match any of the registered faces in the database, it is classified as an unknown or unrecognized face. The detection of an unrecognized face triggers the system to classify the individual as a potential intruder or unauthorized person.

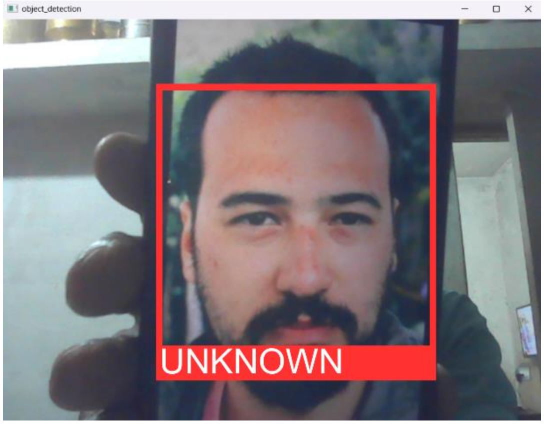


Fig-6: Detection of unknown person

The system looks for weapons if the intruder is discovered within the base. The presence of a weapon will thereafter be reported to the appropriate authorities through an alarm mechanism.

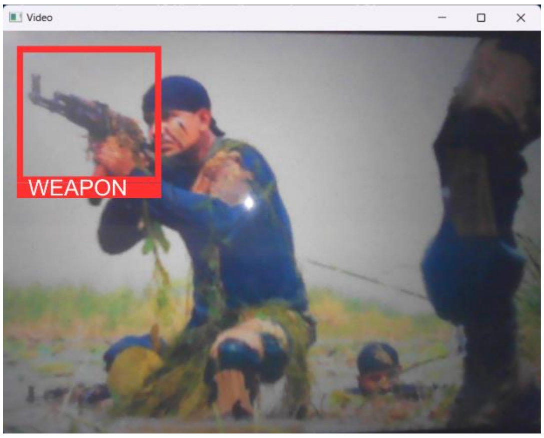


Fig-7: Detection of weapon

8. CONCLUSION

In conclusion, our work proposes a substantial development in security technology, particularly for military bases and high-security facilities. By integrating state-of-the-art computer vision techniques such as face recognition and weapon detection, our system offers a comprehensive solution for enhancing perimeter security, access control, and threat detection.

Through extensive experimentation and implementation, we have demonstrated the effectiveness and reliability of our system in real-world scenarios. The utilization of pre-trained deep learning models for face recognition and object detection ensures high accuracy and efficiency in identifying authorized personnel and detecting potential security threats.

Furthermore, the system's ability to provide real-time monitoring and alerting capabilities helps to improve situational awareness and respond quickly to security events. Our solution provides a proactive approach to security management by utilizing surveillance cameras and sophisticated image processing algorithms, allowing authorities to reduce threats and uphold a secure and safe environment.

REFERENCES

[1] Hana ben Fredj, Souhir Sghaier and Chokri Souani, “An Efficient Face Recognition Method Using CNN” June 30,2021, IEEE.

[2] Maroju Khyathi, “Enhancing Threat Detection in Military Surveillance: A Machine Learning Approach” IJRPR, Vol 5, no 1, pp 1432-1439 January 2024.

[3] Jain, A.K. and Li, S.Z., 2011, “Handbook of face recognition”. New York: Springer.

[4] Nawaf Hazim Barnouti, Sinan Sameer Mahmood, Wael Esam Matti, “Face Recognition: A Literature Review”, Volume 11 – No. 4, September 2016

[5] Harsh Jain, Aditya Vikram, Mohana, Ankit Kashyap, Ayush Jain, "Weapon Detection using Artificial Intelligence and Deep Learning for Security Applications", ICESC 2020.

[6] Muhammad Tahir Bhatti, Muhammad Gufran Khan, Masood Aslam and Muhammad Junaid Fiaz, "Weapon Detection in Real-Time CCTV Videos Using Deep Learning" IEEE, 2021

[7] S. Ren, K. He, R. Girshick, and J. Sun, ''Faster R-CNN: Towards real-time object detection with region proposal networks,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[8] T. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, ''Feature pyramid networks for object detection,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 936–944.

[9] Y. Li, Y. Chen, N. Wang, and Z.-X. Zhang, ''Scale-aware trident networks for object detection,'' in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), Oct. 2019, pp. 6054–6063.

[10] J. Cao, H. Cholakkal, R. M. Anwer, F. S. Khan, Y. Pang, and L. Shao, ''D2Det: Towards high quality object detection and instance segmentation,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Aug. 2020, pp. 11485–11494.

[11] Z.-M. Chen, X. Jin, B.-R. Zhao, X. Zhang, and Y. Guo, ''Hierarchical context embedding for region-based object detection,'' in Proc. IEEE Int. Conf. Comput. Vis., Nov. 2020, pp. 633–648.

[12] Q. Chen, Y. Wang, T. Yang, X. Zhang, J. Cheng, and J. Sun, ''You only look one-level feature,'' in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2021, pp. 13034–13043.

[13] Y. Ma, S. Liu, Z. Li, and J. Sun, ''IQDet: Instance-wise quality distribution sampling for object detection,'' in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2021, pp. 1717–1725.

[14] S. Qiao, L. Chen, and A. Yuille, ''DetectoRS: Detecting objects with recursive feature pyramid and switchable atrous convolution,'' in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2021.