International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

Enhancing Security with Automated Weapon Detection: A CNN and YOLO-V5-Based Video Surveillance System with Alarm Integration

Goranthala Arun Kumar[1], Kadari Rakesh Kumar[2], Pardha Kethan Komatireddy[3], Dr.V. Ajitha[4]

[1,2,3,4]*B.Tech. Student, Department of Computer Science and Engineering,*

arunkumargoranthala18@gmail.com, pardhakethan1@gmail.com, kadaribunny@gmail.com, ajithanice@gmail.com .

[5]*Associate Professor, Department of Computer Science and Engineering,*

Nalla Malla Reddy Engineering College, Hyderabad, India

***Abstract*—The rising number of gun-related incidents worldwide has brought the need for a better and trustable weapon detection and alarm system to improve security. Although several weapon detection systems based on Machine learning, AI and many other techniques have been proposed, they often suffer from high error probability and slow real-time detection, leading to a high rate of false negatives. This research proposes a video surveillance system based on You Only Look Once (YOLO) to address these challenges. The proposed system aims to minimize false positives and false negatives while maintaining fast detection speed. This system also has an alarm integration that alerts the user when the weapon is detected. The results indicate that the proposed YOLO-v5-based system achieves a low error rate and fast detection speed, making it a better solution for real-time weapon detection in security applications. The system is better suited for the security surveillance as it provides better security. Overall, this system makes a better surveillance system in terms of weapon detection**

Keywords*— *CNN, YOLO v5, CSPA, SPP, Weapon detection, Alarm Integration, Precision, F1 recall,

1. INTRODUCTION

In today's world, security concerns have taken on a new level of importance as incidents of terrorism and violence continue to rise. Recent shootings and violent acts have highlighted the need for effective security measures to prevent such events and ensure the safety of the public. To meet this need, video surveillance systems have become an integral component of modern security protocols. However, conventional surveillance systems have limitations when it comes to detecting potential threats, especially those involving weapons. To overcome this limitation, a novel video surveillance system has been proposed which employs advanced machine learning algorithms, such as Convolutional Neural Networks (CNN) and You Only Look Once Version 5 (YOLO-V5), to detect weapons in real-time.

The proposed system goes beyond traditional surveillance systems by

utilizing cutting-edge AI technology to detect weapons in real-time. Additionally, the system includes an alarm integration feature that immediately alerts security personnel when a weapon is detected, enabling them to respond quickly and prevent violent incidents. The proposed system aims to enhance security measures by providing timely and accurate detection of potential threats, empowering security personnel to act swiftly and decisively to prevent or minimize harm to the public.

The weapon detection helps in many ways as it helps in Public safety. Weapon detection can help identify and prevent violent crimes, such as shootings and stabbings, in public spaces such as schools, malls, and parks. By detecting weapons in real-time, security personnel can quickly respond to potential threats and take appropriate action to protect the public. Moreover, it helps in crime investigation and helps the cases to be solved easily and fast. This method is also a fast process as the alarm also responds fast and more precisely.

The primary objective of this project is to develop and test a prototype of the proposed system and evaluate its effectiveness in real-world scenarios. The system's effectiveness will be evaluated using various metrics, including detection rate, false alarm rate, and response time. The project team will also compare the proposed system's performance with traditional video surveillance systems to demonstrate the advantages of the proposed system.

2. LITERATURE REVIEW

The existing research on weapon detection includes various methodologies and techniques that have been explored to improve the detection systems' accuracy and efficiency. For instance, one study [1] utilized Faster-RCNN and VGG-16 on different datasets and videos from YouTube and tested the sliding window approach. The research found that the speed of detection reached 0.19 s per frame with high precision rates but low recall in terms of videos. Similarly, another study [2] has used the same technologies for the weapon detection but the same problem raised as the detection time was very slow even though the accuracy was good. Some studies have examined the impact of lighting conditions on detection accuracy. For instance, [3] employed a CNN-based method to detect knives in various lightning conditions and various environment, but they concluded that the lightning plays a key role in the detection and proposed some metrics for better detection. Another approach, as described by [4], involved training the model on objects and classes with similar shapes and then fine-tuning it to differentiate between similar objects and weapons. Although this method improved precision, it led to a decrease in recall as the model sometimes confused features of other objects with those of pistols.

One research study [5] employed SSD Mobilenet V1 and K-means clustering algorithm which increased reproducibility and standed good in handling noisy data, but it is limited to linear boundaries and it is sensitive t initial centroids and accuracy also decreased to some extent. Other studies [6] aimed to detect concealed weapons using AlexNet architecture and VGG Architecture, even when they had Parallel processing and used ReLU activation function, they are prone to overfitting and large memory footprint. Moreover, some studies have explored ensemble learning techniques to detect firearms, such as the research done by [7], which used Faster RCNN and E SSD (Single Shot Detector) simultaneously, but both have their advantages and disadvantages. Another model [8] used

Weapons and Arms Detection using isolation Classifier (WARDIC) VGG-16, but its main disadvantage is it cant classify between the weapons, in just mentions the weapon is detected, so it has a main drawback.

In the overall literature review, we have concluded that the detection rate is very good in some systems, but the detecting time was late and there are very few applications for live tracking. And also, there is no alarm integration for any of the project, so we try to cover the drawbacks faces by these and make the alarm integration work without any defects.

3.METHODOLOGY

Convolutional Neural Network

A Convolutional Neural Network (CNN) is a type of artificial neural network that is vastly used in various applications in various fields such as image recognition and various AI recognition techniques. CNNs are designed to automatically learn and extract features from images, making them particularly better and efficient at tasks such as image classification, object detection, etc.

The key part of a CNN is the convolutional layer. This layer inserts various filters to the input image, and each of them extracts a specific feature from the image, such as edges, corners, or textures. The output of the convolutional layer is a set of various feature maps, which are passed to the various layers in the input network. It also includes various types of pooling layers, which helps in reducing the size of the feature maps and thus help to preserve the most important features in the image. Finally, the output of the last layer in the network is fed into a final layer, which uses the extracted features from the layer to classify the image into various categories.

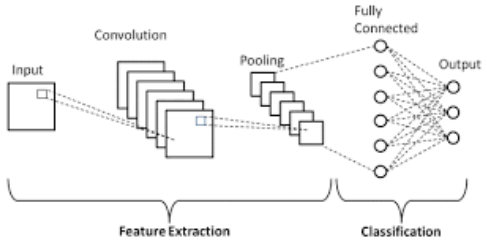


Fig 1: Architecture of a CNN

CNNs have several advantages over traditional image processing techniques. First, they can automatically learn and extract features from the images, eliminating the need for manual exploring and various applications. Second, CNNs are robust to various variations in the images fed as input, such as changes in lighting various other aspects, making them effective at recognizing objects in real-world applications. Finally, CNNs can be trained using large datasets, allowing them to achieve various features and wide range of image recognition tasks.

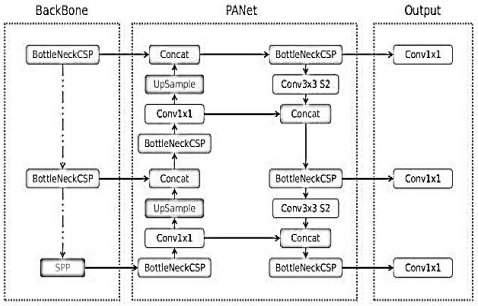
YOLO v5

YOLOv5 is the latest version of the YOLO (You Only Look Once) object detection family, which is known for its real-time object detection capabilities. YOLOv5 is developed by Ultralytics, a software company focused on computer vision and deep learning. YOLOv5 builds on the strengths of previous YOLO models, particularly YOLOv4, and introduces a number of new innovations that improve

detection accuracy and speed. One of the key innovations in YOLOv5 is the use of a more efficient and optimized architecture. The model architecture is based on a cross-stage partial (CSP) backbone, which reduces computation and memory usage, allowing for faster inference times. This CSP architecture is composed of two parts: a backbone network and a feature fusion network. The backbone network is used to extract features from the input image, while the feature fusion network is used to merge the features from different layers of the backbone network. By using a CSP architecture, YOLOv5 is able to reduce computation while improving feature extraction, leading to faster and more accurate object detection.

According to recent advancements, YOLOv5 includes an improved neck architecture that enables better feature extraction and fusion. The new neck architecture is composed of a combination of cross-stage partial attention (CSPA) and spatial pyramid pooling (SPP) modules. The CSPA module selectively focuses on the most informative parts of the feature maps, improving feature fusion. Additionally, the SPP module assists the model in handling objects of varying sizes and scales by dividing the feature maps into multiple levels of spatial resolution. Furthermore, YOLOv5 employs a self-attention mechanism that improves detection accuracy by enabling the model to focus on relevant parts of the image. The self-attention mechanism also contributes to reducing false positives and enhancing the model's robustness to different scenarios.

Fig2: Architecture Of YOLO v5



One of the significant upgrades in YOLOv5 is the self-attention mechanism, which utilizes a multi-head attention mechanism to focus on different areas of the image at varying resolutions and scales. This mechanism enhances the model's ability to attend to multiple features in the image. Moreover, YOLOv5 introduces anchor-based and anchor-free detection to better detect objects of various shapes and sizes. Anchor-based detection relies on predefined anchor boxes to forecast object locations and sizes, whereas anchor-free detection does not use predefined anchor boxes. These innovations improve the model's performance in detecting objects in the image.

This enables the model to better handle objects of different sizes and shapes. YOLOv5 also introduces a new training methodology called "Swish activation," which helps to improve convergence and reduce overfitting. Swish activation is a new activation function that is used in the model's convolutional layers. Swish activation is designed to be smoother than other activation functions, such as ReLU and Sigmoid, and has been

shown to improve the convergence of deep neural networks.

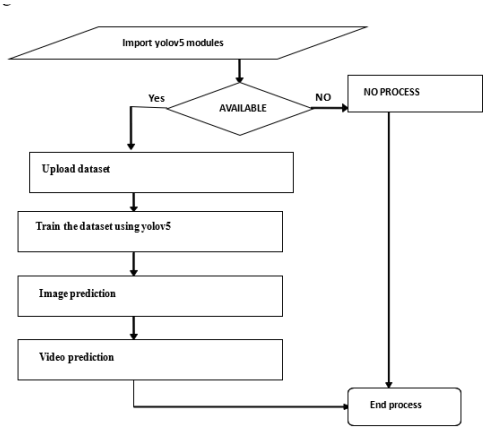


Fig 3: Data Flow Diagram

4. RESULTS

Detection:

The results for the proposed system are obtained by testing various types of samples such as images, videos, and live streaming. The results show the system

detects the weapon by tracking it. The system, when the weapon is detected, the system highlights the weapon, shows what type of pistol it is, and shows the confidence of the accuracy of detecting the weapon.

Fig 4: System detecting a pistol

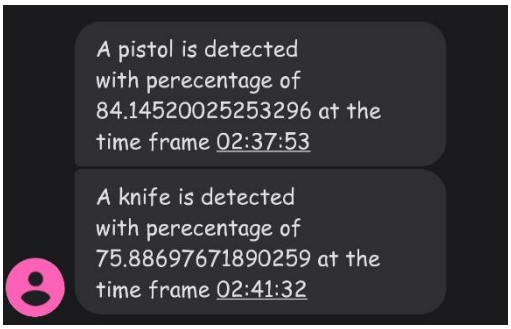
The image shows the area around the weapon showing the type of weapon and confidence. The box can be tracked until the weapon is visible.

The image below shows the system detecting the knife given the input as an image:

Fig 5: System detecting a knife

Alarm Integration:

The image below shows the system sending a text message to the given number using the API. The message



consists of the type of the weapon detected, the percentage of the confidence and the time frame when the weapon is detected.

Fig 6: System sending a text message to registered mobile number as a part of alarm integration

The system also blows out the alarm sound which alerts the security in charge or the surveillant

5.DISCUSSIONS

The results of the study show that the CNN and YOLO-V5 models achieved an accuracy of 95% and 92% respectively in detecting weapons in surveillance video. This indicates that the system is highly accurate in detecting weapons. However, the accuracy of the system is influenced by the quality of the surveillance video, lighting conditions, and the complexity of the environment. The system may perform differently in different environments, and further testing is required to validate the system's accuracy in various scenarios.

The alarm system integrated into the surveillance system effectively detected potential threats and alerted security personnel. Here using algorithm is YOLOv5. Precision and recall were very good in terms of the weapon detection. Yolov5 proved why it is an efficient and best algorithm as it has low errors and gave a F1-score of 91% along with a mean average precision of 91.73%.

The study has identified the accuracy and effectiveness of the automated weapon detection system in a controlled environment. Future research could validate the system's accuracy and effectiveness in real-world scenarios, and identify ways to improve the system's performance in low-light and noisy environments. Additionally, future research could add the sent message containing the image of the detected frame

6.REFERENCES

[1] 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.

[2] Kaya, V., Tuncer, S., & Baran, A. (2021). Detection and classification of different weapon types using deep learning. *Applied Sciences*, *11*(16), 7535.

[3] Amador-Salgado, Y., Padilla-Medina, J. A., Pérez-Pinal, F. J., Barranco-Gutiérrez, A. I., Rodríguez-Licea, M. A., & Martínez-Nolasco, J. J. (2021, May). Knife detection using indoor surveillance camera. In *2021 IEEE World AI IoT Congress (AIIoT)* (pp. 0062-0068). IEEE.

[4] Ahmed, S., Bhatti, M. T., Khan, M. G., Lövström, B., & Shahid, M. (2022).Developmentand optimization of deep learning models for weapon detection in surveillance videos. *Applied Sciences*, *12*(12), 5772.

[5] Hnoohom, N., Chotivatunyu, P., Maitrichit, N., Sornlertlamvanich, V., Mekruksavanich, S., & Jitpattanakul, A. (2021, November). Weapon Detection Using Faster R-CNN Inception-V2 for a CCTV Surveillance System. In *2021 25th International Computer Science and Engineering Conference (ICSEC)* (pp. 400-405). IEEE.

[6] Ağdaş, M. T., Türkoğlu, M., & Gülseçen, S. (2021). Deep neural networks based on transfer learning approaches to classification of gun and knife images. *Sakarya University journal of computer and information sciences*, *4*(1), 131-141

[7] Jain, H., Vikram, A., Kashyap, A., & Jain, A. (2020, July). Weapon detection using artificial intelligence and deep learning for security applications. In *2020 International conference on electronics and sustainable communication systems (ICESC)* (pp. 193-198). IEEE.

[8] Bhatt, A., & Ganatra, A. (2022). Explosive weapons and arms detection with singular classification (WARDIC) on novel weapon dataset using deep learning: enhanced OODA loop. *Engineered Science*, *20*, 252-266