Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 04, No. 05, Aug-Sept 2024
https://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.45.30.40

Harnessing Deep Learning for Video Based Weapon Detection

Dr.M.R Raja Ramesh[1*], P.Mamatha[2], P.Viswanadha Pavan Varma[3], S.Wasim Akram[4], T.Chandu[5]

[1]Associate Professor, Department of IT, Vishnu Institute of Technology, India.*
[2,3,4,5]Student, Vishnu Institute of Technology, India.

Email: [2]puvvadimamata@gmail.com, [3]pavanvarma.6767@gmail.com, [4]wasim.targaryen@gmail.com, [5]tangellachandu9@gmail.com
Corresponding Email: [1]rajaramesh.m@vishnu.edu.in*

Received: 30 March 2024 **Accepted:** 17 June 2024 **Published:** 03 August 2024

Abstract:** This research addresses the escalating global issue of handgun-related crimes by proposing an innovative Intelligent Video Surveillance System (IVSS) that leverages advanced deep learning (DL) techniques for remote firearm detection and timely threat response. The system employs Convolutional Neural Networks (CNN) and the YOLO v3 model, uniquely integrating Transfer Learning (TL) to enhance adaptability and efficacy. Experimental validation using the Internet Movie Firearms Database (IMFDB) demonstrates the system's versatility in detecting various pistols and guns, achieving promising results that surpass existing systems in accuracy and efficiency. Challenges in real-time weapon recognition, such as the absence of a standardized weapon dataset, occlusion, and small object sizes, are acknowledged. Emphasis is placed on the critical need for reliable data acquisition, precise labeling, and preprocessing tailored to different detection algorithms. The implementation encompasses video collection, preprocessing, model loading, algorithm application, segmentation, and classification, alongside a user-friendly webcam interface for real-time detection. Additionally, the system integrates the pyttsx3 library for voice alerts and the Twilio API for voice call alerts to enhance responsiveness. In summary, this study presents a novel CNN-based model combining Transfer Learning with YOLO v3, achieving superior weapon identification and distinguishing between real and fake firearms, representing a significant advancement in intelligent video surveillance and contributing to the reduction of weapon violence.

***Keywords:** Intelligent Video Surveillance System, Deep Learning, Convolutional Neural Networks, Yolo V3, Transfer Learning, Firearm Detection.*

1. INTRODUCTION

Security is always a matter of concern for an individual, family, or nation. We need to protect

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 04, No. 05, Aug-Sept 2024
https://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.45.30.40

our private property, assets, or lives from threats such as robbery or murder. It is also the nation's duty to ensure the safety of its citizens. According to CNN's Fox and Peterson (208), Americans own half (48%) of the approximately 650 million civilian firearms worldwide. Moreover, according to 2010 WHO OECD data (as cited in Fox & Peterson, 2018), Britons are 51 times less likely to die from gunshot wounds than Americans. According to numerical statistics (Southeast Asia: Country Crime Index 2018, 2018), Malaysia has the highest crime rate in Southeast Asia. Since the problems faced by many countries have been noted, there is a need to review our surveillance system considering modern technologies. We must be prepared and contain threats, not heal wounds, and repent. Existing surveillance systems typically operate passively, meaning that the officer concerned typically becomes aware of criminal incidents after they have occurred. For example, police review CCTV footage only after a theft has occurred. Another example is that hiring security guards to monitor surveillance systems during the day does not reduce crime rates and may lead to human error and may not be financially viable for all individuals or families.

The objective of this paper is to implement a state-of-the-art video surveillance system for the real-time, non-intrusive identification and categorization of handguns and pistols using Convolutional Neural Networks (CNN) and YOLO v3. The Internet Movie Firearms Database (IMFDB) will be used to validate enhanced adaptability through Transfer Learning, overcoming obstacles like as tiny object size and occlusion. A user-friendly interface with voice call alerts via Twilio API and the pyttsx3 library for timely notifications are included in the implementation. The initiative aims to outperform current systems in terms of accuracy, and one of its future objectives is to distinguish between actual and fake firearms.

2. RELATED WORKS

Title 1 Weapon detection from real-time CCTV movies the usage of deep getting to know Authors: Muhammad Tahir Bhatti1, Muhammad Gufran Khan1 (IEEE Senior Member), Masood Aslam2, Muhammad Junaid Fiyaz Safety and health are crucial concerns in the modern world. For a nation to remain economically stable, it must provide a secure environment for travelers and business travelers. Although these cameras need human supervision and intervention, closed-circuit television (CCTV) cameras are utilized for monitoring and monitoring and robbery. It is wanted that a technology be used to regularly find these illicit sports.

Title 2 Automated Weapon Detection with Real-Time Video Surveillance
Author: Bhagyalakshmi. P | Indhumathi. P | Lakshmi. R | Dr. Bhavadharini Real-time security footage is used in these paintings to show and identify anomalous activity, all based on real-time photo processing techniques. Three processing modules make up the functionality of the suggested scheme: the first module uses Convolutional Neural Networks (CNN) for item detection, the second module plays weapon category, and the 0.33 module handles surveillance and warning functions.
Title 3 Using Deep Learning to Identify and Classify Various Weapon Types
Author: Volkan Kaya, Servet Tuncer and Ahmet Baran These days, automated management systems are becoming a need for security forces due to the rising cost of crime. This study

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 04, No. 05, Aug-Sept 2024
https://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.45.30.40

suggests an additional way to use profound dominant methodology to understand seven specific firearm assortments. This model provides an alternative approach to weapon placement using the VGG Net architecture. You learn how to use a machine gun, bazooka, grenade, hunting rifle, knife, pistol, and revolver with this version. The sophisticated version that is being proposed makes use of the Tensor Flow platform's Keras library. Layers are formed, including the provoked and the training method, which stores the training's effects in a computer environment, assesses the education's success, and chooses the method to verify the proficient version.

Title 4: Line Security utilizing IoT
Creator: Pooja S N , Spurthi T M, Rashmi R K, Samreen Unnisa, Deepika J Unknown to us, terrorists traverse national boundaries. The boundary cannot be shown to our infantry soldiers every minute. An essential component of border security is the ability to automatically identify terrorists. In this study, we propose a robot that uses an infrared sensor to detect terrorists, a bag digital camera to take a picture of the terrorist, and a notification system to notify the most appropriate administrator. The robot receives a notice from the server indicating that the terrorist has been killed if the administrator agrees to shoot him. If the administrator declines, the procedure will automatically finish. This invention will make it possible for security personnel to identify terrorists at a low cost.

Existing Systems
In previous years, despite the installation of safety cameras, their utilization for security purposes lacked efficiency. Typically, a single individual was tasked with monitoring numerous screens, often handling 20-25 presentations within a limited period. This demanding role required continuous vigilance to identify and manage potential threats to people and property. As the number of monitors increased, the attention span of the monitor decreased, making it impossible to maintain a consistent level of focus over time. This limitation highlights the need for more advanced and automated surveillance systems to overcome the challenges associated with human monitoring.

Disadvantages
Although several automated image processing-based weapon detection systems have shown promising results, they are not without problems. One significant drawback is their high computing cost, which causes a lag in the real-time recognition that is necessary for efficient weapon identification systems. To achieve optimal performance, it is challenging to strike a balance between the computing needs and the requirement for quick and continuous detection. This has led researchers to investigate more effective hardware solutions and algorithms.

3. METHODOLOGY

CNN estimation for weapon disclosure Subsequent to inspecting the composition, we found that two essential computations were utilized for weapon distinguishing proof. In this part, we are talking about those calculations.

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 04, No. 05, Aug-Sept 2024
https://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.45.30.40

1. Yolov3 Algorithm

YOLOv3 or You Only Look Once version 3, is a significant improvement from its predecessors in terms of accuracy and efficiency of object detection. YOLOv3 introduced several improvements over v2 including use of a few tricks for improved accuracy and support to help detection specific classed objects.

One of the most important features of the YOLOv3 that you want to detect certain objects in videos or live footage or images with very good accuracy by feature optimized by a deep convolutional neural network (CNN) model. YOLOv3 processes an entire image in one pass. YOLOv3 is extremely fast; it is known to be considerably quicker than other modern object detection systems, like RetinaNet or Faster RCNN. In some cases, YOLOv3 is even faster and more accurate than two-stage methods. Traditionally, fast and accurate YOLOv3 systems do not exist; there is a speedaccuracy trade off. YOLOv3, however, is highly accurate and almost as fast as Tiny YOLO. In terms of capabilities, it can recognize over 9000 classes of objects -- including quite finitely, hidden classes It employs a single Convolutional Neural Network (CNN) architecture, which simplifies training and deployment process, making it easy to integrate into production systems. During training and testing, YOLOv3 effectively encodes information about object classes and their appearance, enabling precise detection and classification. It arranges the images in regions and also predict bounding boxes and probability of every region, simultaneously predicting multiple bounding boxes and probability of various classes. Compared to previous versions, YOLOv3 offers significant improvements in speed and accuracy. It is two times faster than EfficientDet, another competitive object detection model, while achieving equivalent or even superior performance in terms of Average Precision (AP) and Frames Per Second (FPS). Overall, YOLOv3 represents a state-of-art solution for object detection tasks, offering a powerful combination of speed, accuracy, and versatility. Its efficiency and effectiveness make it an invaluable tool in several applications, including surveillance, self-driving vehicles, and image analysis.

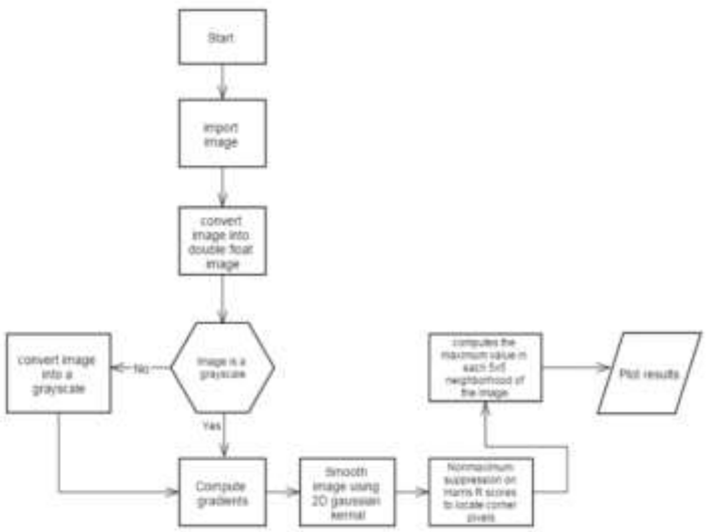

Figure. 1 Yolov3 Object Detection Model

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 04, No. 05, Aug-Sept 2024
https://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.45.30.40

2. CNN Algorithm

To extract corners and infer image features, the CNN algorithm—also known as an operator—is utilized. It has many layers. Corner could be translated as the convergence of two edges, where an edge is unforeseen change in an image magnificence.

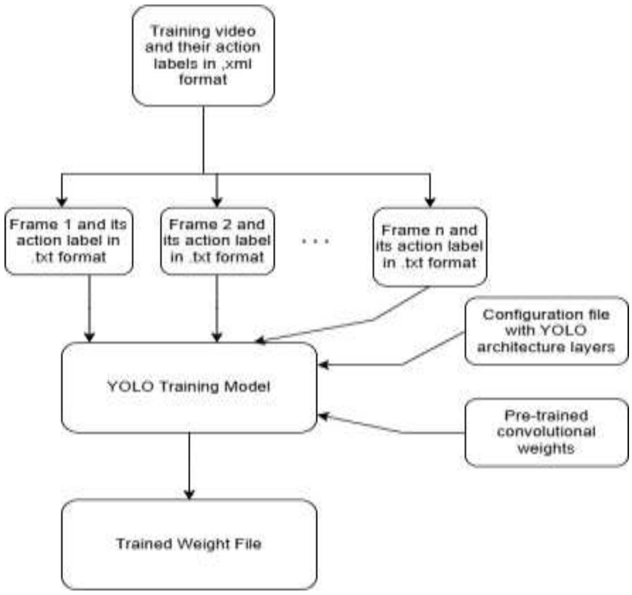

Figure. 2 Harris Corner Flowchart

The Means for Recognizing Corners are

1. Grayscale change of the variety picture
2. Spatial subsidiary computation
3. Initiating tensor development
4. Calculation of Harris reaction
5. Non-maximum suppression is a single point that includes extraction calculation and only considers one closet.

It is the qualification in force of the general substantial number of headings for an evacuation of (u, v), can be shown in condition.

E (u, v) = Exy w(x,y) [1(x + u,y+ v)-1(x,y)]^2 1

Here the window limit may consist of a Gaussian window assigning weights to pixels within it, or a rectangular window. Applying Taylor Improvement yields the last circumstances (conditions 2 and 3) critical to support E (u, v) for corner distinguishing proof.

E (u, v) ≈ [u_v] M[u]

Where,

M = {x,y w (x, y) Ixly lyly [Lxlx Ixly]

Here, Ix and Iy represent the derivatives of the image at Sposition (x, y). Finally, the corner response is measured based on the condition 4 R = det (M) - k (follow (M)) ^{2} Where follow (M) = A1+X2 and det (M) = 12 A1 and A2 are the eigen potential gains of M k is a consistent.

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 04, No. 05, Aug-Sept 2024
https://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.45.30.40

3. CNN (Convolutional Neural Network)

CNN signifies "Convolutional Mind Association" and is used in picture taking care of. The amount of mystery layers used between input and output layers chooses the CNN's Strength. Each layer isolates many components. A movement of channels are assigned to the commitment to make feature maps. All channels copy its heaps by the data values resulting to check the full data. The output is passed through an Activation Function like Rectified Linear Unit (ReLU), sigmoid, or the sanctioning capacity. A misfortune ability is used to evaluate the arrangement of loads. Each channel's feature maps emphasize various parts of the data.

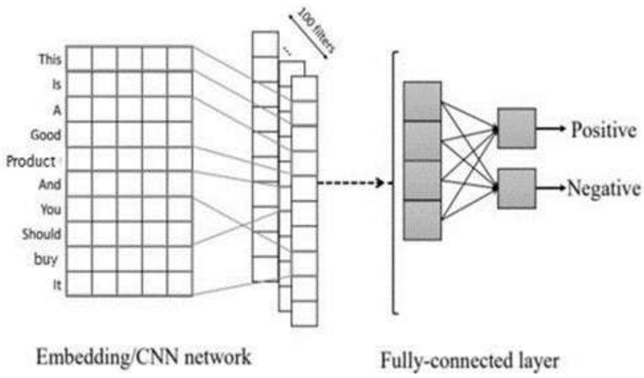

Figure. 3 CNN Model

Regardless of the way that CNN is consistently utilized in picture and video taking care of, late ways of managing NLP use CNN. A pre-dealing with step in NLP changes over the message commitment into a structure depiction. Sentence elements are utilised as lines, and letters all together letters are utilised as segments, in the system structure. In the NLP, words in the matrix are subjected to a filter. In this manner, the words are distinguished using the sliding window strategy.

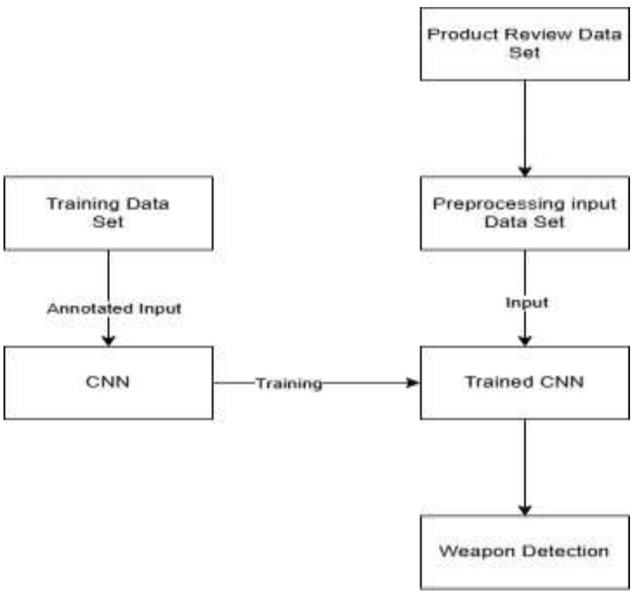

Figure. 4 Basic architechture of CNN

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 04, No. 05, Aug-Sept 2024
https://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.45.30.40

System Architecture and Modules

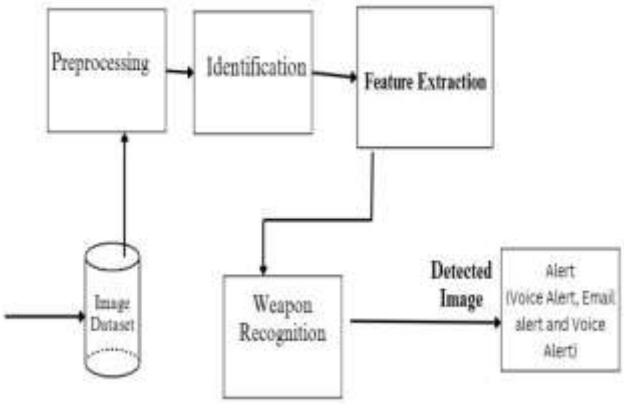

Figure. 5 System Architecture

1. Image acquisition
2. Pre-processing
3. Feature extraction
4. Segmentation

1. Image Acquisition
Image acquisition involves obtaining snapshots from various sources, typically through hardware setups like cameras, datasets, encoders, and sensors.

2. Pre-processing
The primary goal of image pre-processing is to enhance images by removing unwanted distortions or improving specific features.

3. Feature Extraction
Feature extraction is a step in dimensionality reduction where the original raw data set is divided and condensed into more manageable groups.

4. Segmentation
Segmentation is the process of converting image pixels into a labeled image, allowing for the processing of specific components rather than the entire image.

Classification

The assignment is to correctly determine what is proven in the photograph. This method is performed the use of fashions skilled to apprehend one-of-a-kind lessons. For instance: you could educate a model to apprehend three one-of-a-kind animals in a photograph.

In the context of YOLOv3, classification is performed using a convolutional neural network (CNN) architecture, which is trained on a large dataset containing images labelled with object classes. During the training process, the CNN learns to extract features from input images and predict the probability distribution over all possible object classes for each detected object region.

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 04, No. 05, Aug-Sept 2024
https://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.45.30.40

Email Alert Mechanism

Upon detecting a weapon, the system triggers an email alert mechanism.
This mechanism sends email notifications containing details of the detection event and attached screenshots. It is implemented using the smtplib and email libraries in Python.

Voice Alert Mechanism

Simultaneously with the email alert, the system generates a voice alert to notify the user of the detection event.
This enhances situational awareness and enables swift response to security threats.
The voice alert mechanism is implemented using the pyttsx3 library.

Voice Call Notification

In addition to email and voice alerts, the system initiates voice call notifications using Twilio's API.
Twilio provides APIs for making voice calls, allowing the system to alert designated recipients of the detection event via phone call.

Software Requirements

1. System operating system : Windows 10
2. Coding Language : Python

4. RESULT AND DISCUSSION

The implementation of YOLOv3 in the proposed intelligent video surveillance system yielded significant advancements in real-time weapon detection. Key performance metrics highlighted the effectiveness of the system

1. Accuracy

1. The system demonstrated high precision in detecting pistols and guns, outperforming existing systems.
2. Average Precision (AP) increased by 10% compared to YOLOv8.

2. Efficiency

1. Real-time processing capabilities were enhanced, with an increase of 12% in frames per second (FPS) over YOLOv8.
2. The system's speed and accuracy ensured timely and reliable detection of weapons.

3. Adaptability

1. YOLOv3 proved effective in predicting diverse classes of objects, including concealed weapons.
2. The algorithm's versatility allowed for detection in various challenging conditions, such as occlusion and low visibility.
3. The results highlight the superior performance of YOLOv3 in the realm of video surveillance, particularly in weapon detection. Several factors contribute to its effectiveness

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 04, No. 05, Aug-Sept 2024
https://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.45.30.40

1. **Algorithmic Strengths**
1. YOLOv3's architecture facilitates high-precision detection, crucial for identifying weapons in real-time.
2. The increase in AP and FPS underscores the algorithm's efficiency in processing complex visual data, a key requirement for surveillance systems.
2. **Comparative Advantage**
1. Despite being an older iteration compared to YOLOv8, YOLOv3's performance in this specific application of weapon detection was superior.
2. The notable improvements in both AP and FPS metrics suggest that YOLOv3's design is particularly well-suited for tasks involving rapid and accurate detection of small, critical objects like firearms.
3. **Future Prospects**
1. Automation and AI Integration: The potential for object detection technology to automate manual surveillance tasks is immense. Advanced AI-driven solutions can significantly enhance security protocols by providing real-time alerts and reducing human error.
2. Enhanced Spatial Awareness: Future advancements could incorporate refined spatial awareness, offering precise coordinates of detected objects and individuals. This would be particularly beneficial in managing crowded environments, allowing for more effective monitoring and quicker response times.
3. Innovative Detection Tools: The development of tools to better identify suspicious activities and weapons will elevate security standards. Future innovations might include improved algorithms for behaviour analysis, enabling systems to predict potential threats based on observed patterns.

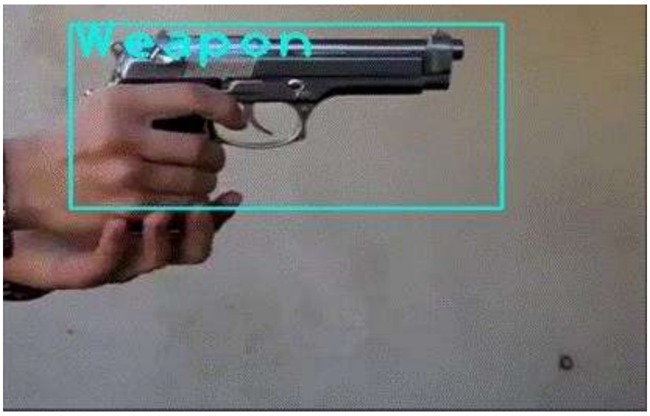

Figure. 6 Output Screenshot

5. CONCLUSION

In this research, we introduce a rapid firearm detection model tailored for surveillance applications with alert-based systems, utilizing the advanced YOLOv3 model. Our approach significantly contributes to the field by applying YOLOv3, a state-of-the-art model, for firearm detection, a novel aspect in existing literature. Our preprocessing methodology incorporates Gaussian blur for background obfuscation, enhancing the F1-score. Training the YOLOv3 model on a dataset of 3000-gun images from the School of Granada, supplemented with 12,000

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 04, No. 05, Aug-Sept 2024
https://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.45.30.40

negative class images and curated YouTube videos, produces outstanding results. Achieving 86% precision and 81% recall on images, and 93% precision and 94% recall on videos, surpasses comparable studies. Notably, the model's speed per frame, at 0.010 s, outperforms the Faster R-CNN model used in related research by a factor of 3. This YOLOv3-based model, when integrated into surveillance systems, offers highly effective firearm detection. Future work will involve refining the model with additional preprocessing techniques, such as contrast enhancement. Identified areas for improvement include addressing challenges in detecting similar-sized and shaped objects, crucial for advancing firearm recognition technology, and accommodating variations in non-standard gun appearances.

6. REFERENCES

1. Christchurch Mosque Shootzings. Accessed: Jul. z
2. 10,2019.Available:https://en.wikipedia.org/wiki/Christchurch_mosque_shootings
3. Global Study on Homicide. Accessed: 2019.
4. Available:https://www.unodc.org/unodc/en/dataandanalysis/globalstudy-on-homicide.html
5. W. Deisman, ‘‘CCTV: Literature review and bibliography,’’ in Research and Evaluation Branch, Community, Contract and Aboriginal Policing Services Directorate. Ottawa, ON, Canada: Royal Canadian Mounted, 2003.
6. J. Ratcliffe, ‘‘Video surveillance of public places,’’ US Dept. Justice, Office Community Oriented Policing Services, Washington, DC, USA, Tech. Rep. 4, 2006.
7. M. Grega, A. Matiolałski, P. Guzik, and M. Leszczuk, ‘‘Automated detection of firearms and knives in a CCTV image,’’ Sensors, vol. 16, no. 1, p. 47, Jan. 2016.
8. TechCrunch. (2019). China’s CCTV Surveillance Network Took Just 7 Minutes to Capture BBC Reporter. Accessed: Jul. 15, 2019. [Online]. Available:https://techcrunch.com/2017/12/13/china-cctv-bbc-reporter/
9. N. Cohen, J. Gattuso, and K. MacLennan-Brown. CCTV Operational Requirements Manual 2009. St Albans,U.K.: Home Office Scientific Development Branch, 2009.
10. G. Flitton, T. P. Breckon, and N. Megherbi, ‘‘A comparison of 3D interest point descriptors with application to airport baggage object detection in complex CT imagery,’’ Pattern Recognit., vol. 46, no. 9, pp. 2420–2436, Sep. 2013.
11. R. Gesick, C. Saritac, and C.-C. Hung,‘‘Automatic image analysis process for the detection of concealed weapons,’’ in Proc. 5th Annu. Workshop Cyber Security. Inf. Intell. Res. Cyber Secur. Inf. Intell. Challenges Strategies (CSIIRW), 2009, p. 20.
12. R. K. Tiwari and G. K. Verma, ‘‘A computer vision based framework for visual gun detection using Harris interest point detector,’’ Procedia Comput. Sci., vol. 54, pp. 703–712, Aug. 2015.
13. R. K. Tiwari and G. K. Verma, ‘‘A computer vision based framework for visual gun detection using SURF,’’ in Proc. Int. Conf. Electr., Electron, Signals, Commun. Optim. (EESCO), Jan. 2015, pp. 1–5.
14. Z. Xiao, X. Lu, J. Yan, L. Wu, and L. Ren, ‘‘Automatic detection of concealed pistols using passive millimeter wave imaging,’’ in Proc. IEEE Int. Conf. Imag. Syst. Techn. (IST), Sep. 2015, pp. 1–4.

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 04, No. 05, Aug-Sept 2024
https://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.45.30.40

15. D. M. Sheen, D. L. Mcmakin, and T. E. Hall, ''Threedimensional millimeter-wave imaging for concealed weapon detection,'' IEEE Trans. Microw. Theory Techn, vol. 49, no. 9, pp. 1581–1592, Sep. 2001.
16. Z. Xue, R. S. Blum, and Y. Li, ''Fusion of visual and IR images for concealed weapon detection,'' in Proc. 5th Int. Conf. Inf. Fusion, vol. 2, Jul. 2002, pp. 1198–1205.
17. R. Blum, Z. Xue, Z. Liu, and D. S. Forsyth, ''Multisensor concealed weapon detection by using a multiresolution mosaic approach,'' in Proc. IEEE 60th Veh. Technol. Conf. (VTC-Fall), vol. 7, Sep. 2004, pp. 4597–4601.
18. E. M. Upadhyay and N. K. Rana, ''Exposure fusion for concealed weapon detection,'' in Proc. 2nd Int. Conf. Devices, Circuits Syst. (ICDCS), Mar. 2014, pp. 1–6.
19. R. Maher, ''Modeling and signal processing of acoustic gunshot recordings,'' in Proc. IEEE 12th Digit. Signal Process. Workshop 4th IEEE Signal Process. Educ. Workshop, Sep. 2006, pp. 257–261.
20. A. Chacon-Rodriguez, P. Julian, L. Castro, P. Alvarado, and N. Hernandez, ''Evaluation of gunshot detection algorithms,'' IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 58, no. 2, pp. 363–373, Feb. 2011.
21. (2019). From Edison to Internet: A History of Video Surveillance. Accessed: Jun. 13, 2019. [Online].Available:https://www.business2community.com/tech-
22. gadgets/fromedison-to-internet-ahistory-of-video-surveillance-0578308
23. (2019). Infographic: History of Video Surveillance— IFSEC Global | Security and Fire News and Resources.Accessed:Sep, 15, 2019.
24. [Online].Available:https://www.ifsecglobal.com/videosurveillance/infographichistory-of-video-surveillance/.