American Journal of Innovation in Science and Engineering (AJISE)

ISSN: 2158-7205 (ONLINE)

VOLUME 4 ISSUE 2 (2025)

PUBLISHED BY
E-PALLI PUBLISHERS, DELAWARE, USA

American Journal of Innovation in Science and Engineering (AJISE)

Volume 4 Issue 2, Year 2025
ISSN: 2158-7205 (Online)
DOI: https://doi.org/10.54536/ajise.v4i2.4685
https://journals.e-palli.com/home/index.php/ajise

Enhancing Digital Forensics with Deep Learning: Applications of CNNs and YOLOv5 in Weapon Detection and Image Forgery Analysis

Anmol Deepak Kumar[1*], Tamizharasan Periyasami[1]

Article Information

Received: March 02, 2025

Accepted: April 11, 2025

Published: June 24, 2025

Keywords

Abnormal Activities, Crime Localization, Weapon Recognition, YOLOv5

ABSTRACT

Security is an essential problem in all domains. The crime rates are increased at crowded events or suspected isolated regions. Computer vision has significant uses in detecting and monitoring abnormalities to address diverse issues. Video surveillance systems are increasingly necessary for safeguarding the safety, security, and personal belongings. The ability of these systems to identify and understand scenes and unusual occurrences is crucial for effective intelligence monitoring. The primary objective of this study was to analyze surveillance films to identify weapons and detect any unusual behaviors or actions. This study applied the advanced cutting-edge framework YOLOv5 to examine and identify abnormalities like weapon recognition and criminal behavior in the public surveillance video dataset. The proposed approach implementation accomplishes an (mAP) mean average precision of 96.1%, outperforming state-of-the- art methods in terms of accuracy and efficiency for recognizing weapons and localization of criminal behavior in challenging surveillance datasets.

INTRODUCTION

The rising incidence of public crimes necessitates implementing a robust monitoring system. The use of special tools including both automatic and manual weapon determines the severity of a crime scene (Casino *et al.,* 2022). Criminal actions are frequently reported worldwide. The perpetrators manage to evade from the scene due the inadequate security measures and incapability of competent authorities i.e. police to apprehend them (Holt *et al.,* 2022). In this regard, the Closed-Circuit Television (CCTV) is the most prevalent method of preventing violence. Even with the surveillance of CCTV cameras, criminals are engaged in various illegal activities such as robbery and snatching. The authorities only become aware of the crime, when subsequently examine the crime scene video. Therefore, surveillance is more essential to identify and halt the illegal activities with proactive actions. CCTV is a network of interconnected cameras, used to record and store visual data of a specific area. A security guard watches live and recorded videos to identify unusual activities. Therefore, the use CCTV has been increased in combating with perpetrators at crime area. Studies have shown that CCTV contributes to decreasing crime rates in specific locations, such as car parks and residential areas (Fakiha, 2023).

Furthermore, the monitoring infrastructure entails a substantial upfront cost as well as significant ongoing maintenance expenses (Pearson & Watson, 2010). The essential gear associated with surveillance and personnel, i.e., guards, are the primary factors contributing to these expenses. Another crucial and sometimes disregarded element of surveillance systems is the surveillance team. These workers are actively monitoring the cameras and reporting any unusual activity. It is an advance system that can proactively detect illegal activity and taking appropriate measures would significantly prevent such crimes and safeguard property (Dunsin *et al.,* 2024; Nastasi, 2021). On the other hand, technology has progressed sufficiently, for example, aiding humans in visual activities (Byrne & Marx, 2011; Debnath & Bhowmik, 2021). The rapid advancement in computer-based technology coupled with exponential expansion of inaccessible data have enabled the development of an autonomous system to identify objects and patterns, for example, such systems are installed at airports to identify the objects inside the cases. Significantly, the ability to detect the presence of a weapon in surveillance film enables the identification of questionable or illegal conduct.

This study aims to introduce a novel approach based on advanced YOLOv5 architecture to overcome state-of- the-art circumstances and ensure rapid security and safeguarding lives. This proposed approach outperformed state-of-the-art methods, achieving a mean average precision (MAP) accuracy of 96.1% even in challenging environments.

LITERATURE REVIEW

Related literature was collected based on the use of artificial intelligence (AI), machine learning (ML), and deep learning (DL) approaches to address major areas of digital multimedia analysis in authenticity verification, deception detection, weapon recognition, and crime detection.

Face Detection

Face detection and alignment are required for a variety of facial applications, including face recognition and emotion analysis (Bartlett *et al.,* 2003). Large visual

[1] Computer Science, Birla Institute of Technology and Science, Pilani Dubai, UAE
[*] Corresponding author's e-mail: f20200221@dubai.bits-pilani.ac.in

variations of faces, such as occur- sions, considerable posture fluctuations, and harsh lighting, offer significant hurdles to these tasks in real-world applications. Viola and Jones introduced a cascade face detector that is efficient and performs well in real time. It trains cascaded classifiers with AdaBoost and Haar-Like features (Viola & Jones, 2004). However, studies have shown that in real-world applications with more visual variances in human faces, therefore, this type of detector does not perform effectively, even with more advanced features and classifiers. Mathias proposed face-detection deformable component models that functioned effectively (Chrysos *et al.,* 2018). Therefore, it is difficult to understand these computationally intensive models with annotations. Convolutional neural networks (CNNs) have made significant advances in a variety of computer vision applications, including face recognition and picture categorization. Deep neural networks are taught to recognize facial features in order to provide a high response in face regions and find potential face windows. However, due to complex CNN structure, this approach takes a long time to apply in practice. The cascaded CNNs are used for face identification, the intrinsic relationship between bounding box regression and facial landmark localizations are ignored, resulting in extra processing costs for bounding box calibration. CNNs have demonstrated exceptional success in image classification and face recognition, and recent breakthroughs in this discipline have the potential to revolutionize a wide range of computer vision applications (Luo *et al.,* 2018). The Localization Transformer (LOTR) framework is a significant advance in face landmark localization, providing a direct coordinate regression technique that takes advantage of transformer networks' capabilities (Watchareeruetai *et al.,* 2022). Unlike previous approaches, LOTR prioritizes the effective use of spatial information inside feature maps, which is accomplished with a three-module design that includes a visual backbone, a Transformer module, and a landmark prediction head.

Forgery Detection

Detecting picture counterfeiting has become increasingly important in the digital (technology) world, as powerful editing tools and procedures sometimes undercut the validity of visual content.

Recent advances in DL have transformed picture forgery detection, allowing for the creation of more robust and automated detection systems (Ghai *et al.,* 2024). CNNs have shown promising success in learning discriminative features directly from image data, allowing them to detect various types of forgeries without the requirement for generated qualities. This technique uses a pre-trained DL model. Next, the network's architecture is fine-tuned using a small training set of Counterfeit Media Files (CMF) pictures. Prior to computer-generated forgeries, this strategy performed exceptionally well. The performance of the CMF approach appears insufficient (Wang *et al.,* 2018).

CNN Methods for Identity Verification and Fraud Preven- tion

Recent developments in face detection have focused on addressing challenges associated with unconstrained en- vironments, including variations in scale, pose, illumination, and occlusion (Ranjan *et al.,* 2019). Techniques such as deep pyramid single shot face detection (DPSSD) have emerged as promising solutions, offering both speed and accuracy in detecting faces with large-scale variations, particularly tiny faces. (Zinjurde & Kamble, 2020) highlights the rising importance of Internet banking for financial transactions, emphasizing the danger of consumer confidentiality during online purchases. In response, a revolutionary two-step authentication solution is devised to prevent third-party intrusion during online transactions, including OTP verification and face recognition. This dual-layer authentication approach improves security by verifying transaction information and user identification. Another study provides a unique offline (hand) signature verification technique based on Python-developed CNN models (Alajrami *et al.,* 2020). In this context, CNN model achieves higher accuracy up-to 99.70%, even after intensive training and validation operations, ultimately proving its ability to check offline signatures with more -reliability. Therefore, the hand signature modifications in the CNN model can speed up and can enhance authentication in multiple domains. It is consistent with the current study. In contrast, another showed a novel detection of fraud in terms of online transactions by using CNN models that reorganize raw transaction features into distinct convolutional patterns (Zhang *et al.,* 2018). The model achieves remarkable precision and recall rates by utilizing low-dimensional and non-derivative data, stabilizing around at 91% and 94%, respectively. Such accuracy highlights the promise of novel neural network architectures to enhance the detection of fraud skills and protecting financial systems.

Deep Learning Approaching for Authenticity Verification

This study provides a defined experimental approach and benchmark to offer a fair and consistent basis for comparing new techniques with established methodologies. Therefore, this study evaluates and adapts Time-Aligned Recurrent Neural Networks (TA-RNNs) for online signature verification, demonstrating its complexity and capability over existing techniques by achieving remarkably low Equal Error Rates (EERs) in spite of skilled forgery imposters and a lack of training data. In this regard, research has focused on identifying phony or fraudulent fingerprints(Yoon *et al.,* 2012).

The DL is now used in forensic analysis to detect criminal activities such as arson, burglary, and vandalism in digital multimedia information (Acharjya *et al.,* 2022). A study provided a comprehensive DL framework for crime scene investigation, including CNNs for image processing and classification tasks (Acharjya *et al.,* 2022). The pre-trained models such as VGGNet, ResNet, and

You Only Look Once (YOLO) are known to increase weapon identification accuracy and may be fine-tuned using domain-specific datasets (Acharjya *et al.,* 2022). Transfer learning from the large-scale picture helps to recognize and address the challenges in the narrow goal of recognizing weapons. The YOLO algorithm has emerged as a game changer in object identification, transforming real-time detection systems with speed and accuracy. YOLO substantially improves over standard object identification algorithms since it frames the work as a single regression problem, predicting bounding boxes and class probabilities directly from input photographs.

Zhang *et al.* created a YOLO-based system for video violence identification, using YOLO's real-time detection capa- bilities to recognize aggressive actions and violent situations (Zhang *et al.,* 2018). As security cameras and video surveillance systems grow more integrated into public safety, intelligent threat detection becomes increasingly important. Manually identifying dangerous objects is labor-intensive, necessitating the development of automated systems for rapid threat detection (Ullah *et al.,* 2023).

MATERIALS AND METHODS

We used two major datasets: one for weapon detection a dataset containing nine classes of weapon and UCF Crime, which includes normal versus abnormal activities, like assault or robbery. Each was then divided into training and validation sections and annotated in a way compatible with YOLO.

The YOLOv5 ancho-based detection heads were based on a feature pyramid network and CSP-Darknet backbone. We fine-tuned pre-trained weights on our labeled data, then optimized a multi-part loss of bounding box, classification, and objectness via stochastic gradient descent. We save the best checkpoint and allow for frequent validation checks to prevent overfitting.

The input of YOLOv5 for the entire inference process will be real-time frames from IP cameras or recorded clips. It will trigger an alert in case of high-confidence detection of a weapon or any suspicious behavior such as fighting.

Overview of YOLOv5

YOLOv5 is a pre-trained object detector; it is a CNN model. A CNN is a deep learning algorithm that can receive an input image in raw form and apply learnable weights and biases to different aspects/objects. The CNN model's convolutional layer extracts high-level features from the input image, such as edges (Khanam & Hussain, 2024). The overview of YOLOv5, is shown in figure 1.

Backbone Network

The main component of the net- work is the backbone. In the case of YOLOv5, the backbone is constructed using the New CSP-Darknet53 structure, a modified version of the Darknet architecture utilized in earlier iterations.

Neck

This section links the main structure and the topmost component. YOLOv5 employs SPPF and New CSP-PAN structures.

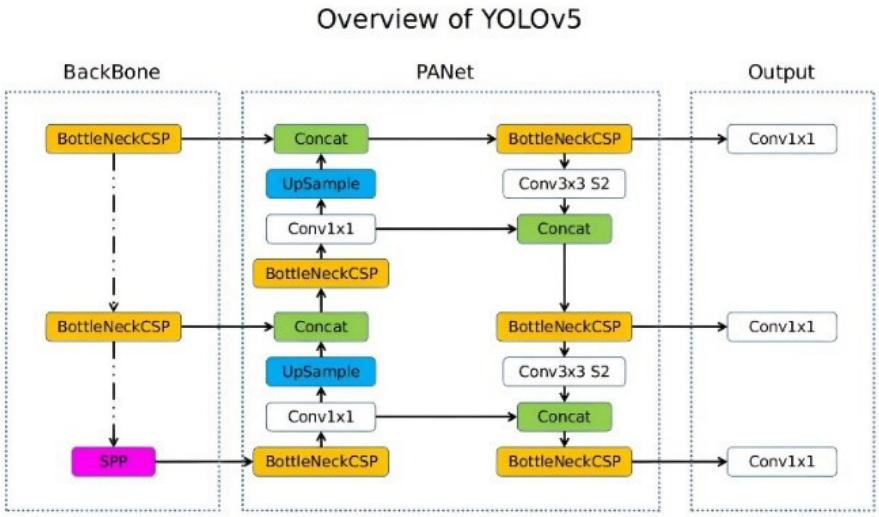

Figure 1: Overview of YOLOv5 Architecture

Detection Head

This section is accountable for producing the ultimate result. YOLOv5 utilizes the YOLOv3 Head for this specific objective.

The YOLOv5 model can recognize objects in real time with the highest accuracy, which makes it a popular option for a variety of computer vision tasks, such as object detection in digital forensics investigations (Kaur & Singh, 2023).

Overview of Proposed Model

In this study, YOLOv5 utilizes a more straightforward structure that allows for highly accurate real-time analysis of images or video frames: it uses a backbone

network composed of convolutional layers to collect hierarchical information from the input image, and then a feature pyramid network (FPN) to communicate these attributes, allowing the collection of multi-scale information needed to identify objects of different sizes within the image. The detection head of the model predicts bounding box coordinates and class probabilities for objects it recognizes, followed by prediction layers that produce predictions at various spatial resolutions over the feature maps. Another essential element of the proposed paradigm that facilitates object localization are anchor boxes, which serve as reference templates for estimating object bounds. During training, the model optimizes a mixture of localization and classification loss functions to measure the difference between expected outputs and ground truth annotations. Customized YOLOv5 uses iterative optimization techniques, such as stochastic gradient descent, to modify its settings in order to reduce the discrepancy and increase detection accuracy. The overall goal of the suggested model structure, as seen in Figure 2, is to quickly evaluate input photographs, accurately identify items of different sizes and classifications, and perform very well across a variety of object identification tasks.

Methodology Workflow

As seen in Figure 2, this approach illustrates how to build machine learning models utilizing crime and weapon statistics. After relevant data has been extracted from the weapon dataset by filtering, it is converted to Pascal format. The purpose of data augmentation is to increase the diversity of this data. The information is then supplied into YOLOv5 for object detection training after being converted to a YOLO-compatible format. At the same time, a sequential model for classification is trained using the crime dataset. Figure 3 illustrates how both models are trained and maintained using YOLOv5, which focuses on weapon detection and sequential models on crime-related classification.

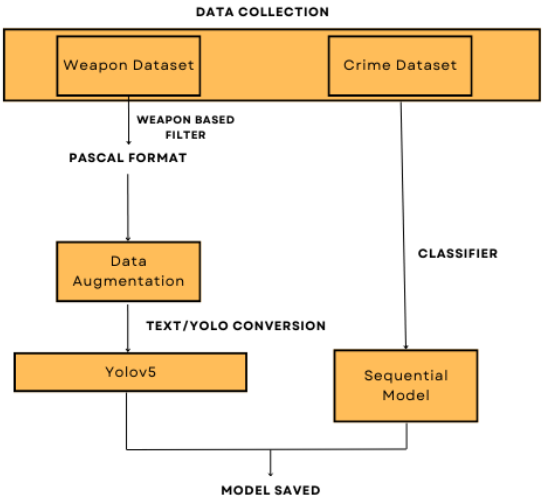

Figure 2: Workflow for Weapon Detection and Crime Classification

Dataset Selection and Preprocessing

The data from the many datasets utilized in this study, along with the various preparation techniques used to improve the data's visibility, are provided in this part.

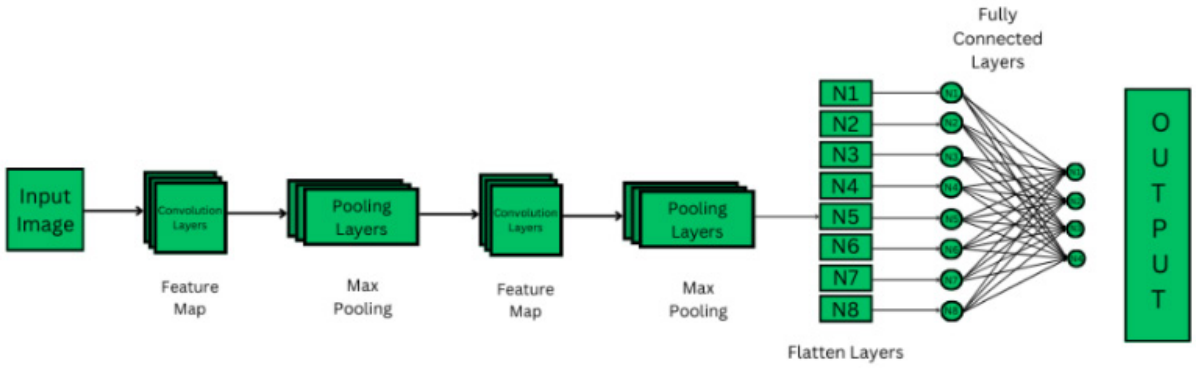

Figure 3: Proposed Advanced YOLOv5 Model for Weapon Detection and Crime Classification

UFC Crime Dataset

One popular dataset in computer vision and video analysis is the UCF Crime Dataset (Yuan *et al.,* 2023). As seen in Figure 4, it consists of a compilation of surveillance

footage taken in a variety of real-world situations. Thirteen odd behaviors in public settings make up this dataset, which is divided into several types. Anomalies can be seen in any surveillance footage.

UFC Dataset Preprocessing

By enhancing the quality of input data and locating relevant features, preprocessing methods for the UCF Crime Dataset can enhance the functionality of computer vision and video analysis algorithms. In order to facilitate effective processing and analysis, video segmentation first divides the surveillance material into smaller pieces or clips. This method is beneficial for several reasons..

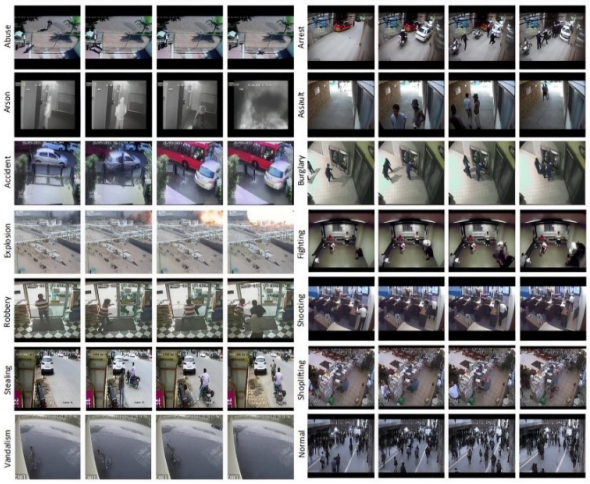

Figure 4: UFC Crime Dataset Sample

in figure 6. This dataset was built using Python's simple image download module, which obtains pictures from the internet. 100 photos from each class were gathered. After the inspection, invalid photographs were eliminated, leaving us with a total of 714 images for all nine classes.

Training Data Distribution

A balanced distribution would imply that each class is equally represented in the training set, ensuring that the model obtains adequate exposure to all sorts of criminal activity while training, is shown in figure 7.

Frame Extraction

Frame extraction is the technique of separating individual frames from a video stream, resulting in a collection of static pictures. This approach is widely utilized in video processing and analysis, including computer vision and machine learning applications, is shown in figure 5.

Weapon Dataset

This dataset includes photos from nine distinct types of firearms. Previously, datasets had just one class: Weapon or Gun. This dataset currently consists of nine classes: automatic rifle, bazooka, handgun, knife, grenade launcher, shotgun, SMG, sniper, and sword, is shown

Figure 5: Frame Extraction of different Classes from Video Frames

Test Data Distribution

An adequate test data distribution is critical for correctly assessing the performance and generalization capabilities of machine learning models trained on the UCF Crime Dataset, is shown in figure 8.

Experimentation and Evaluation Results

The proposed experimental model for digital forensic analysis consists of many essential processes targeted at training, testing, and fine-tuning the object identification model to efficiently analyze multimedia information for forensic applications.

Figure 6: Weapon Detection Datasets Sample

Hyperparameters Tuning

Hyperparameter tuning includes parameters like the number and size of filters, which influence the model's ability to capture various feature types and sizes. A typical setup might use 32 filters with a 3x3 size. The stride determines how the filter moves across the input, with a common default value of 1 for preserving spatial information. Padding ensures output size matches input size, often using the "same" padding. Pooling reduces spatial dimensions to retain essential information, with a 2x2 pooling window and a stride of 2 being typical. ReLU is typically chosen as the activation function due to its efficiency in training deep networks.

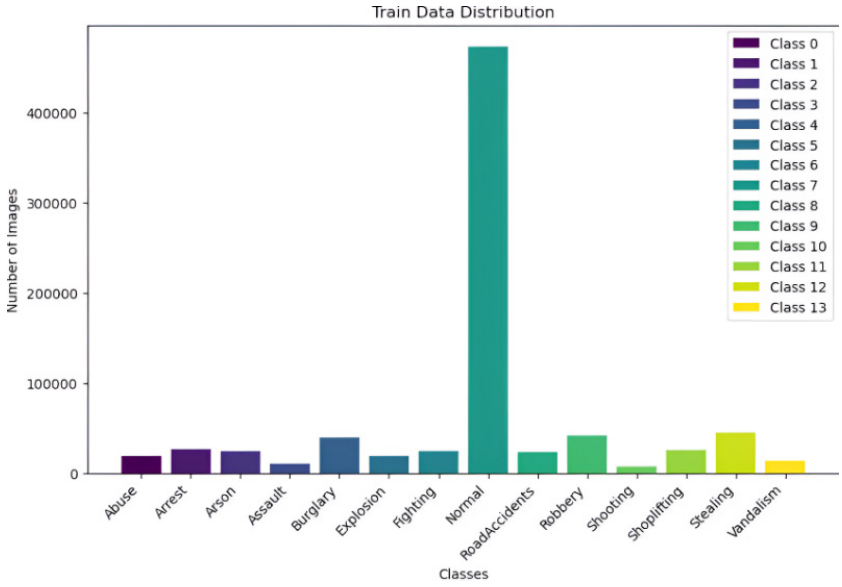

Figure 7: Train Data Distribution of UFC Dataset Classes

Model Fine-Tuning

YOLOv5 fine-tuning involves using the weights of a pre-trained model as an initial reference for further training on specific datasets. This method entails fine-tuning different components of the model to get optimum outcomes. Through the process of fine-tuning the model, you may optimize its performance and achieve the utmost level of efficiency.

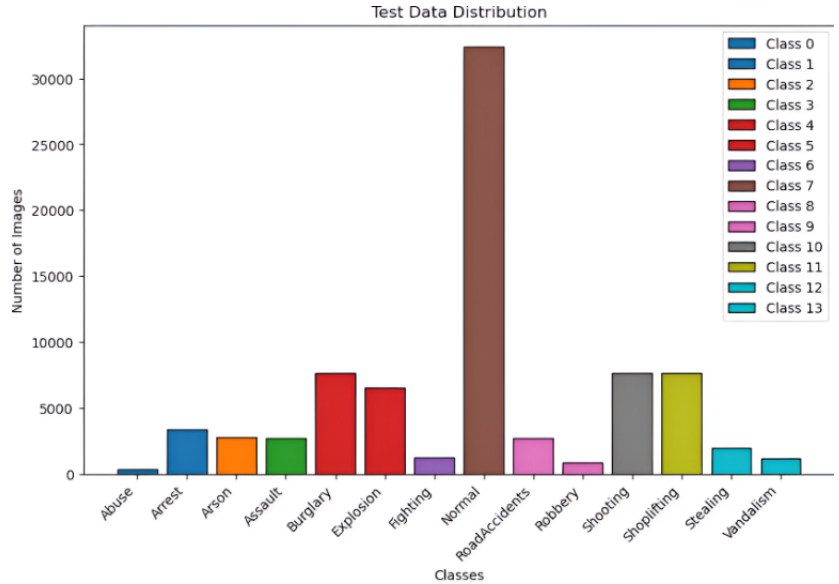

Figure 8: Test Data Distribution of UFC Dataset Classes

Evaluation Results

Various assessment criteria are used to evaluate the efficacy of the suggested model. These assessment indicators demon- strate the effectiveness and dependability of the machine learning models.

Mean Average Precision (mAP)

This statistic is used to assess the quality and precision of the proposed system. The suggested method would provide a detailed analysis of its effectiveness in predicting

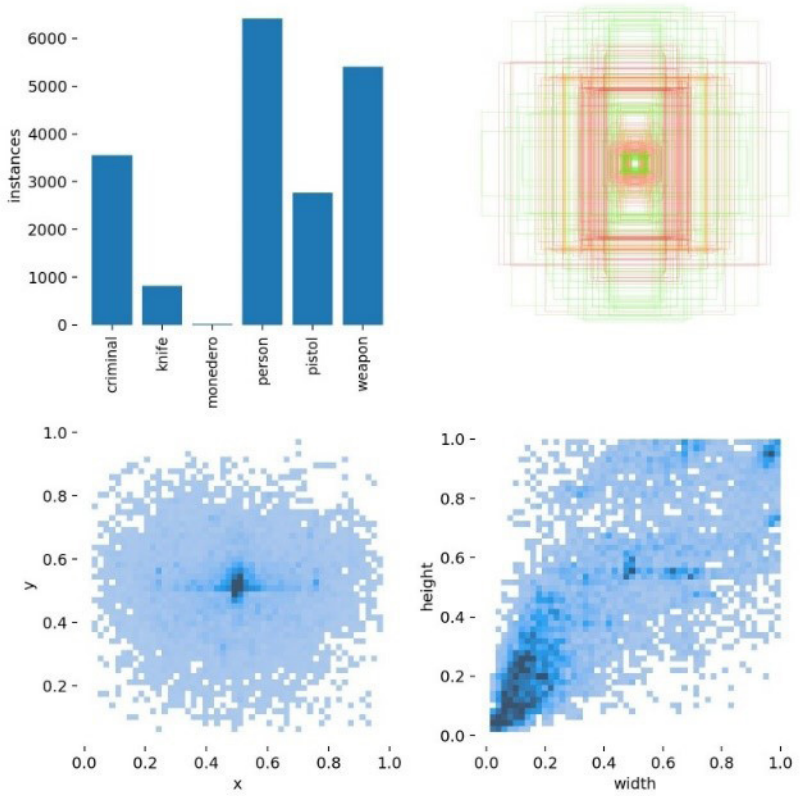

Figure 9: Labelling Classes Heat Maps

weapons and irregularities within the dataset.

Intersection over Union (IoU)

This evaluation approach uses the suggested model to assess the degree to which the predicted bounding box corresponds with the ground truth box.

Precision and Recall

Recall quantifies the model's capacity to accurately detect positive instances, especially true positives, out of all the real positive instances. On the other hand, precision refers to the accuracy with which a model can correctly detect negative occurrences or non-events. The statistic measures the proportion of correct pessimistic predictions to the total number of negative occurrences in the dataset.

Labelling Evaluation

During the training process of the proposed model, our objective is to ensure that the predictions align with the corresponding labels. As the labeling evaluation becomes more rigorous, the model's accuracy increases, as shown

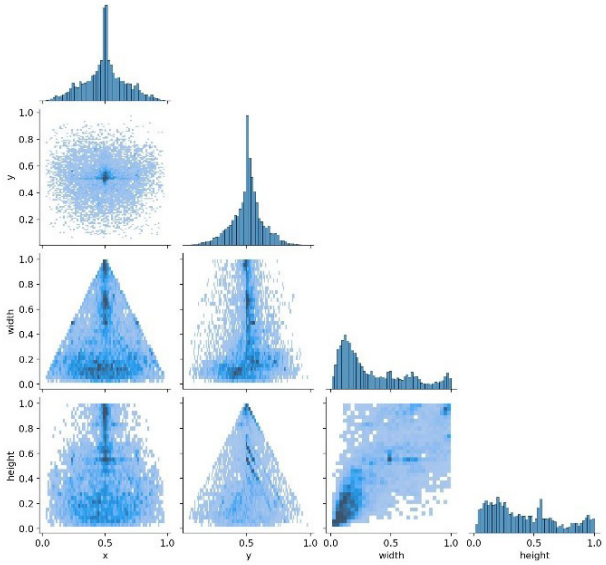

Figure 10: Labelling Classes Heat Maps

in Figure 9 and 10.

Confusion Matrix Evaluation

The matrix of confusion is an essential tool for assessing the performance of a classification model. The visual repre- sentation allows for a direct comparison between the real and predicted labels, providing insights about misclassifications and the overall performance of the model. As shown in Figure 11. Classes with actual classification values near 1, such as Criminal Class, Predicted Person, and Weapon with values of 0.98 and 0.97, exhibit high accuracy, indicating that the proposed

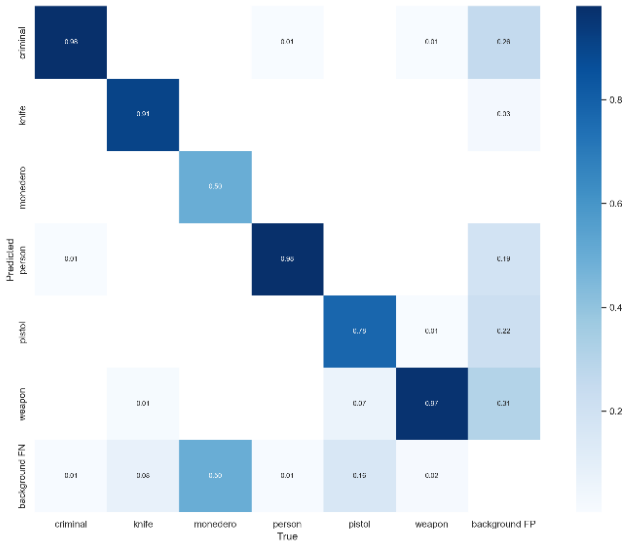

Figure 11: Confusion Matrix Analysis

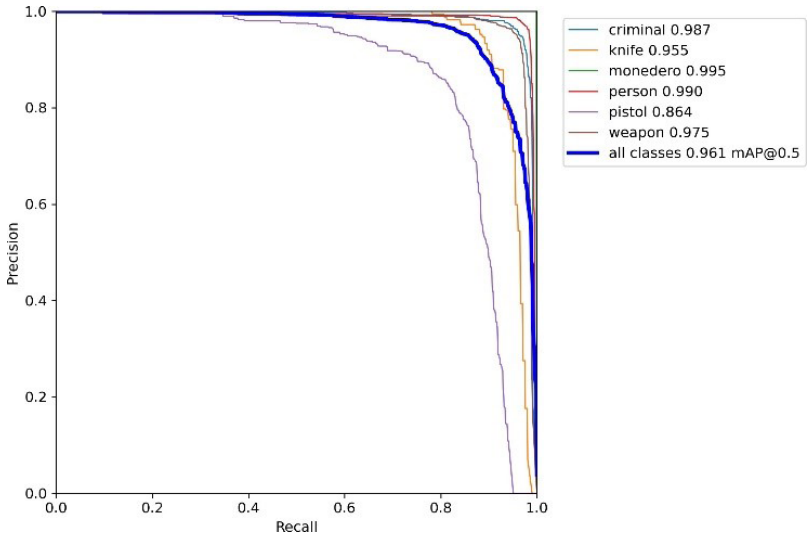

Figure 12: F1-Curve Evaluation Analysis

method correctly identifies instances belonging to these classes.

F1-Curve Analysis

Precision indicates the proportion of true positive predic- tions among all positive predictions generated by the model, whereas recall reflects the proportion of genuine positive predictions among all positive cases in the dataset.

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (1)$$

Figure 12 shows that all the F1 curve values are more significant than 0.93, showing the proposed model's good accuracy.

Figure 13, P-Curve shows that all the predicting classes for the weapon dataset are between 1.00 and 0.95, and the model has a higher Precision value.

PR-Curve Analysis

The precision-Recall curve evaluates that the proposed model predicts how many classes are positive, and the pre- diction quality is explained in these metrics, is shown in figure 14.

ROC-Curve Evaluation Analysis of Crime Abuse Classifi- cation

The ROC curve is an explanation graph that illustrates how

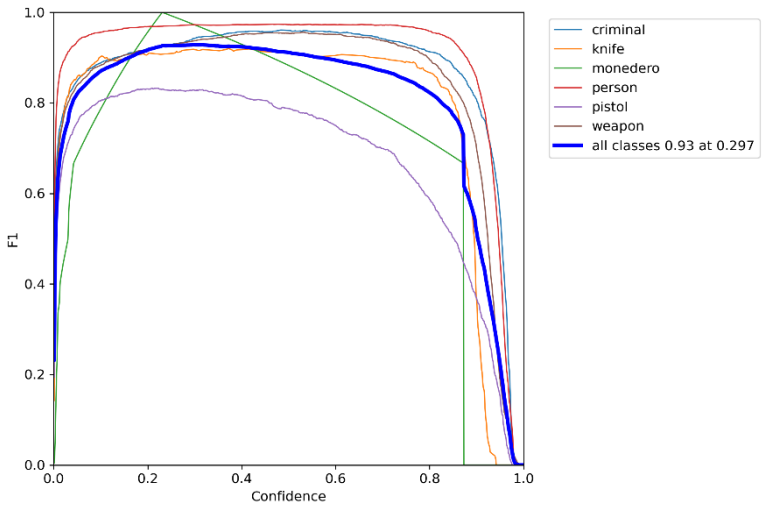

Figure 13: P-Curve Evaluation Analysis

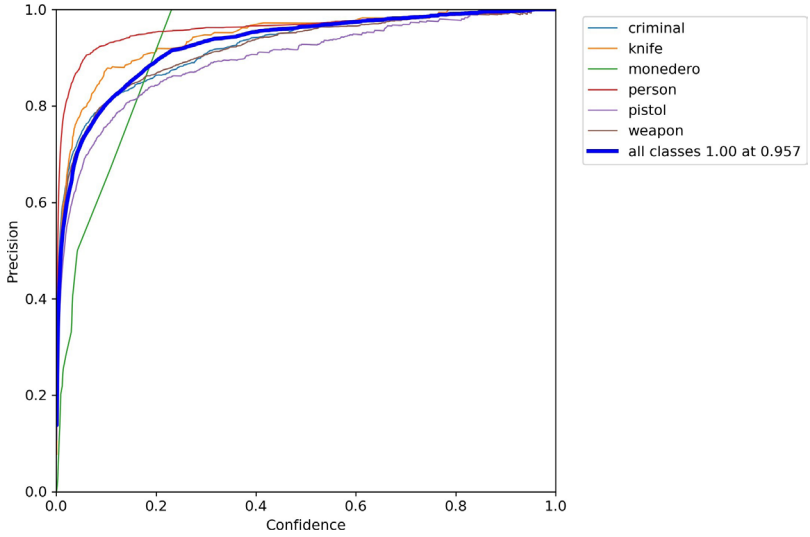

Figure 14: PR-Curve Evaluation Analysis

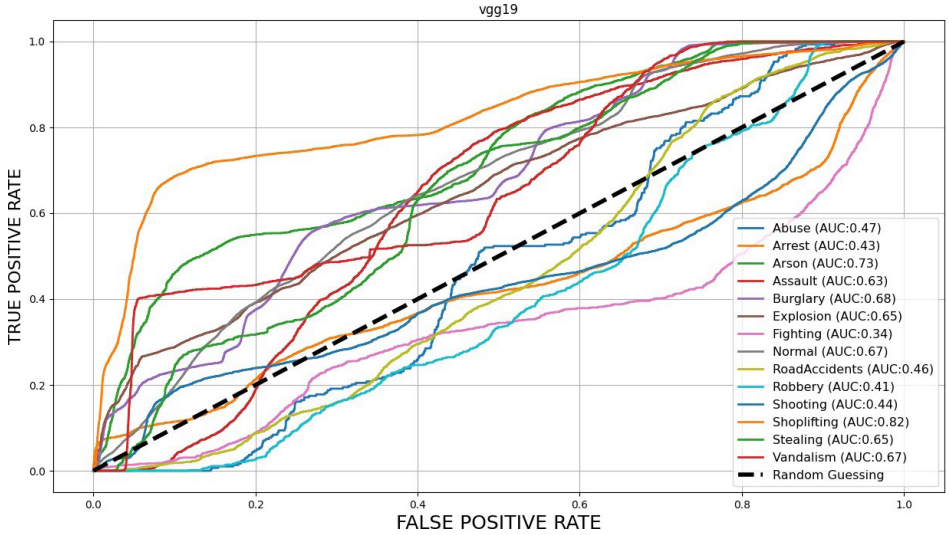

Figure 15: Crime Abuse Classification ROC-Curve Analysis

well the suggested model performs binary classification at various classification thresholds, is shown in figure 15.

Comparing the Proposed Model to Existing Methods

Table 1 compares the performance of various methods

for object detection, showing that the proposed approach, Im- proved YOLOv5, achieves the highest accuracy with a 96.1% mAP (mean Average Precision). This surpasses previous meth- ods, including YOLOv4 (91.73%), VGG-Net19 (92%), and others, demonstrating significant improvement over earlier YOLO models and alternative approaches like multimodal information fusion.

Table 1: Comparison Of The Proposed Model With State-Of-Art Methods

Reference	Approach	Results (mAP%)
(Quyyum & Abdullah, 2022)	YOLOV3- Darknet-53	88.97
(Quyyum & Abdullah, 2022)	Improved YOLOV3- Darknet-53	90.20
(Jiang & Jiang, 2024)	Multimodal Information Fusion	61
(Bhatti *et al.,* 2021)	YOLOV4	91.73
(Navalgund & Priyadharshini, 2018)	VGG-Net19	92
Proposed Model	Improved YoloV5	96.10

CONCLUSION

Concisely, our findings demonstrate the efficacy of deep learning approaches in expanding the area of digital forensic investigation. We effectively handled fundamental difficulties in digital evidence analysis using convolutional neural net- works (CNNs) and other deep learning architectures, such as picture forgery detection, weapon classification, and crime detection in surveillance footage.

The use of Improved YOLOv5 for digital forensics tasks has shown to be an effective method for detecting abnormalities and suspicious activity in surveillance footage. We were able to discover items of interest accurately and in real time by using its efficient design and powerful detection capabilities, hence improving the capabilities of digital forensic investigations.

REFERENCES

Acharjya, P. P., Koley, S., & Barman, S. (2022). A review on forensic science and criminal investigation through a deep learning framework. *Aiding Forensic Investigation Through Deep Learning and Machine Learning Frameworks,* 1-72.

Alajrami, E., Ashqar, B. A., Abu-Nasser, B. S., Khalil, A. J., Musleh, M. M., Barhoom, A. M., & Abu-Naser, S. S. (2020). *Handwritten signature verification using deep learning*.

Bartlett, M. S., Littlewort, G., Fasel, I., & Movellan, J. R. (2003). Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction. 2003 Conference on computer vision and pattern recognition workshop,

Bhatti, M. T., Khan, M. G., Aslam, M., & Fiaz, M. J. (2021). Weapon detection in real-time cctv videos using deep learning. IEEE Access, 9, 34366-34382.

Byrne, J., & Marx, G. (2011). Technological innovations in crime prevention and policing. A review of the research on implementation and impact. *Journal of Police Studies, 20*(3), 17-40.

Casino, F., Dasaklis, T. K., Spathoulas, G. P., Anagnostopoulos, M., Ghosal, A., Borocz, I., Solanas, A., Conti, M., & Patsakis, C. (2022). Research trends, challenges, and emerging topics in digital forensics: A review of reviews. *IEEE Access, 10*, 25464-25493.

Chrysos, G. G., Antonakos, E., Snape, P., Asthana, A., & Zafeiriou, S. (2018). A comprehensive performance evaluation of deformable face tracking "in-the-wild". *International journal of computer vision, 126*, 198-232.

Debnath, R., & Bhowmik, M. K. (2021). A comprehensive survey on computer vision based concepts, methodologies, analysis and applications for automatic gun/knife detection. *Journal of Visual Communication and Image Representation, 78*, 103165.

Dunsin, D., Ghanem, M. C., Ouazzane, K., & Vassilev, V. (2024). A comprehensive analysis of the role of artificial intelligence and machine learning in modern digital forensics and incident response. *Forensic Science International: Digital Investigation, 48*, 301675.

Fakiha, B. (2023). Enhancing Cyber Forensics with AI and Machine Learning: A Study on Automated Threat Analysis and Classification. *International Journal of Safety & Security Engineering, 13*(4).

Ghai, A., Kumar, P., & Gupta, S. (2024). A deep-learning-based image forgery detection framework for controlling the spread of misinformation. *Information Technology & People, 37*(2), 966-997.

Holt, T. J., Bossler, A. M., & Seigfried-Spellar, K. C. (2022). *Cybercrime and digital forensics: An introduction.* Routledge.

Jiang, F., & Jiang, J. (2024). Research on the Application of Object Detection Methods Based on Multimodal Information Fusion in Digital Forensics. In *2024 4th International Conference on Consumer Electronics and Computer Engineering (ICCECE)*

Kaur, R., & Singh, S. (2023). A comprehensive review of object detection with deep learning. *Digital Signal Processing, 132*, 103812.

Khanam, R., & Hussain, M. (2024). What is YOLOv5: A deep look into the internal features of the popular object detector. *arXiv preprint arXiv:2407.20892.*

Luo, D., Wen, G., Li, D., Hu, Y., & Huan, E. (2018). Deep-learning-based face detection using iterative bounding-box regression. *Multimedia Tools and Applications, 77*, 24663-24680.

Nastasi, C. (2021). Multimedia Forensics: From Image manipulation to the Deep Fake. *New Threats in the Social Media Era.*

Navalgund, U. V., & Priyadharshini, K. (2018). Crime

intention detection system using deep learning. 2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET),

Pearson, S., & Watson, R. (2010). *Digital triage forensics: processing the digital crime scene.* Syngress.

Quyyum, M. E. E., & Abdullah, M. H. L. (2022). Weapon Detection in Surveillance Videos Using Deep Neural Networks. *Multimedia University Engineering Conference (MECON 2022)*

Ranjan, R., Bansal, A., Zheng, J., Xu, H., Gleason, J., Lu, B., Nanduri, A., Chen, J.-C., Castillo, C. D., & Chellappa, R. (2019). A fast and accurate system for face detection, identification, and verification. I*EEE Transactions on Biometrics, Behavior, and Identity Science, 1*(2), 82-96.

Ullah, F. U. M., Obaidat, M. S., Ullah, A., Muhammad, K., Hijji, M., & Baik, S. W. (2023). A comprehensive review on vision-based violence detection in surveillance videos. *ACM Computing Surveys, 55*(10), 1-44.

Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International journal of computer vision, 57*, 137-154.

Wang, J., Ma, Y., Zhang, L., Gao, R. X., & Wu, D. (2018). Deep learning for smart manufacturing: Methods and applications. *Journal of manufacturing systems, 48*, 144-156.

Watchareeruetai, U., Sommana, B., Jain, S., Noinongyao, P., Ganguly, A., Samacoits, A., Earp, S. W., & Sritrakool, N. (2022). Lotr: face landmark localization using localization transformer. *IEEE Access, 10*, 16530-16543.

Yoon, S., Feng, J., & Jain, A. K. (2012). Altered fingerprints: Analysis and detection. *IEEE transactions on pattern analysis and machine intelligence, 34*(3), 451-464.

Yuan, T., Zhang, X., Liu, K., Liu, B., Jin, J., & Jiao, Z. (2023). *UCF-Crime Annotation: A Benchmark for Surveillance Video-and-Language Understanding.* arXiv preprint arXiv:2309.13925.

Zhang, Z., Zhou, X., Zhang, X., Wang, L., & Wang, P. (2018). A model based on convolutional neural network for online transaction fraud detection. *Security and Communication Networks, 2018*(1), 5680264.

Zinjurde, A. M., & Kamble, V. B. (2020). Credit card fraud detection and prevention by face recognition. *2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing (ICSIDEMPC).*