IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Enhancing Weapon detection with Pose Analysis: Leveraging Visual and Body pose Features using Open Pose

Sriram C S [a],Dr. G. Prema Arokia Mary[b]
[a]MTech Data Science
[b],Assistant Professor Dept. of Information Technology
[a, b]Kumaraguru College of Technology, Coimbatore
sriram.22mds@kct.ac.in

Abstract: The use of closed-circuit television (CCTV) is essential in modern security systems to proactively identify and address safety threats and potentially dangerous situations, emphasizing the need for early detection. Although deep learning methods have led to the development of automated weapon detection systems with promising results, they often rely heavily on the visual characteristics of weapons. In cases where handguns are barely detectable, the use of body pose information can be beneficial. This paper presents a novel method that integrates weapon appearance with human pose data into a cohesive system. The approach begins by using pose estimation techniques to pinpoint hand regions, which are then cropped and fed as input into the classification model. Achieving an accuracy of 96.39%, this system demonstrates strong capabilities in threat detection. The adaptable nature of this approach allows it to be applied not only in public areas but also in high-security environments, thereby greatly enhancing overall safety and security strategies.

Keywords: Deep learning, Pretrained Models, CNN, open pose, Body pose Estimation, Classification, Weapon.

1.INTRODUCTION

The widespread problem of gun-related violence has emerged as a major global concern, affecting numerous aspects of society, including public health, safety, and economic stability. In countries such as India, the prevalence of firearms is particularly high, presenting an urgent need for effective detection and prevention strategies. With over 71 million firearms in the country—second only to the United States—India grapples with the issue of unlicensed and unregistered weapons, which account for approximately 86% of all firearms.

Surveillance technology has undergone substantial advancements in recent years, leading to the widespread adoption of closed-circuit television (CCTV) systems in various public and private spaces. These systems play a critical role in investigating incidents and aiding security personnel in crowd management, enabling them to oversee multiple areas concurrently. Nevertheless, a key limitation of such systems is the requirement for constant human oversight, which can be hindered by fatigue and diminishing attention over time.

Recent advances in deep learning have shown promise in tasks such as object recognition and classification, leading to improvements in automated threat detection systems. However, challenges persist, particularly when applied to real-world scenarios different from those in training environments, where false positives may hinder effectiveness.

This study examines the potential of integrating human body pose information as an additional data source to improve the performance of existing handgun detection systems. Human pose, which describes the relative positioning of body joints and limbs, is often distinct in shooting scenarios. Considering the limitations of CCTV imagery, including low resolution, poor lighting conditions, and challenges such as occlusion and varying camera perspectives, the combination of pose information may contribute to more robust and reliable detection.

2. LITERATURE REVIEW:

Olmos, R. et al. in Neurocomputing presented a novel approach to firearm detection in videos using deep learning. Their study utilizes CNNs, leveraging pre-trained models like VGG-16 and ResNet for feature extraction. Results demonstrate high accuracy in identifying handguns, highlighting the potential for enhancing security measures. The study offers valuable insights into automated firearm detection systems, contributing to public safety initiatives.

Ineneji C , Kusaf M. present a "Hybrid weapon detection algorithm, using material test and fuzzy logic system". The study integrates material testing and fuzzy logic systems to enhance weapon detection accuracy. The research underscores the potential of hybrid approaches in advancing security measures and suggests avenues for further optimization.

Dwivedi, et al. present "Weapon classification using deep convolutional neural network". The study applies deep learning, particularly CNNs, to accurately classify weapons. By leveraging CNNs, the framework learns discriminative features from weapon images, enhancing classification accuracy. The research showcases the potential of deep learning in advancing weapon detection and classification for improved security measures.

Egiazarov A, et al. introduce "Firearm detection and segmentation using an ensemble of semantic neural networks". The study employs multiple neural networks to accurately identify and segment firearms in images. By leveraging semantic information, the framework achieves promising results, advancing firearm detection and segmentation for enhanced security measures.

Ben Abdallah H, et al. introduce "AMSEP: Automated Multi-level Security Management for Multimedia Event Processing" The framework offers automated security management across multiple levels for multimedia event processing systems. It includes advanced algorithms for threat detection, real-time monitoring, and automated response mechanisms. AMSEP addresses security complexities, enhances efficiency, and ensures robust protection for multimedia data and events.

Grega M, et al. explore "Automated detection of firearms and knives in a CCTV image". Their study introduces a robust framework grounded in computer vision techniques aimed at precisely identifying weapons within surveillance footage. By employing sophisticated algorithms, the system effectively discerns firearms and knives in real-time, thereby bolstering security measures and fostering a safer environment. The research underscores the significance of automated weapon detection systems in enhancing surveillance capabilities and mitigating potential threats, ultimately contributing to the advancement of public safety initiatives.

Tiwari R.K, Verma G.K. present "A Computer Vision based Framework for Visual Gun Detection Using Harris Interest Point Detector". Their study introduces a novel approach to firearm detection, focusing on visually identifying guns in various contexts. Leveraging the Harris Interest Point Detector algorithm, the framework accurately detects potential firearms in images, enhancing security measures. By integrating image processing algorithms and machine learning models, the system offers a reliable solution for firearm detection, contributing to public safety initiatives and advancing firearm detection technologies.

Xu Z, et al. present "Dangerous human event understanding using human-object interaction model". Their study introduces a novel framework for identifying hazardous situations by analysing human-object interactions. By leveraging advanced machine learning algorithms, object detection techniques, and human behaviour analysis methods, the framework proactively detects potential risks in real-time.

Yadav et al. present "A comprehensive study towards high-level approaches for weapon detection using classical machine learning and deep learning methods". This study delves into various methodologies for weapon detection, encompassing classical machine learning and deep learning techniques. Their insights contribute to enhancing security measures and public safety initiatives in weapon detection.

Ahmed, et al. Development and optimization of deep learning models for weapon detection in surveillance videos". Their work focuses on refining deep learning models tailored for weapon detection within surveillance videos. Through optimization strategies, the authors enhance the efficiency and accuracy of these models, addressing challenges associated with identifying weapons in complex visual environments.

Ashraf, Abdul Hanan, et al. present "Weapons detection for security and video surveillance using CNN and YOLO-v5s". Their research focuses on implementing CNN and YOLO-v5s models for weapon detection in security and video surveillance applications. By leveraging these advanced neural network architectures, the authors aim to enhance the accuracy and efficiency of weapon detection systems, thereby contributing to improved security measures. This study underscores the significance of utilizing state-of-the-art deep learning techniques for enhancing surveillance capabilities and promoting public safety initiatives.

Mehmood, Abid, presents "Abnormal behaviour detection in uncrowded videos with two-stream 3D convolutional neural networks". This research focuses on detecting abnormal behaviours in uncrowded videos using two-stream 3D convolutional neural networks (CNNs). By leveraging advanced deep learning architectures, the study aims to enhance the accuracy and efficiency of abnormal behaviour detection systems. This research contributes to the development of more robust surveillance systems for identifying and addressing security threats in various environments.

Alairaji, et al. contribute "Abnormal behaviour detection of students in the examination hall from surveillance videos". This research focuses on detecting abnormal behaviours exhibited by students in examination halls using surveillance videos. The research underscores the significance of leveraging surveillance technology to maintain academic integrity and ensure fairness in examination environments. This study contributes to the development of automated systems for detecting and preventing academic dishonesty, ultimately promoting the integrity of educational assessments.

3.EXISTING SYSTEM:

Numerous research studies have played pivotal roles in advancing firearm detection and surveillance systems, employing a variety of methodologies and techniques. These include deep learning methods tailored for firearm detection in videos, hybrid algorithms merging material testing and fuzzy logic systems, and frameworks utilizing convolutional neural networks (CNNs) for weapon classification. Additionally, there are systems for firearm detection and segmentation leveraging semantic neural networks, automated security management across multimedia event processing systems, and computer vision-based approaches for visual gun detection. Other studies focus on understanding dangerous human events through models analysing human-object interactions, comprehensive assessments of high-level strategies for weapon detection, and optimization of deep learning models for surveillance videos. Furthermore, there are implementations of CNN and YOLO-v5s models for weapons detection, abnormal behaviour detection employing two-stream 3D CNNs, and surveillance-based identification of irregular behaviours in examination halls. These collective research endeavours contribute significantly to the evolution of advanced surveillance systems, bolstering public safety and security measures across various domains.

4. DATA DESCRIPTION:

The dataset offered by the Andalusian Research Institute in Data Science and Computational Intelligence (DaSCI Institute) is dedicated to advancing intelligent video surveillance automatic systems, with a particular focus on weapon detection. It consists of high-quality image datasets [14] tailored for training Deep Learning models, essential for identifying potentially violent scenarios early on. The dataset encompasses various weapons and small objects, including pistols, knives, bills, purses, smartphones, and cards, organized into six distinct classes as shown in figure 1. These classification images are derived from detection images, wherein object bounding boxes are isolated. Moreover, the txt file accompanying the dataset provides detailed information on experiment dataset partitions, contributing essential insights for the associated publication.

Database	# img	Pistol	Knife	Smartphone	Bill	Purse	Card
1	4710	3394	1879	866	134	157	179
2	5454	3523	1879	1022	387	315	387
3	6658	3681	1879	1069	654	710	544
Sohas_weapon	5080	1580	1879	755	345	581	340
Sohas_weapon-Test	1170	294	470	115	123	104	64

Table 1 Dataset Description

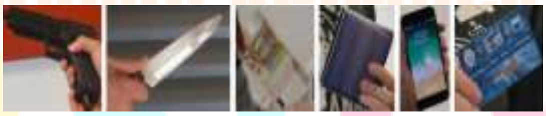

FIG 1. Dataset Sample Images

5. PROPOSED METHODOLOGY:

In this section, we'll explain each step of our proposed methodology, from the starting picture to finding the handguns at the end.

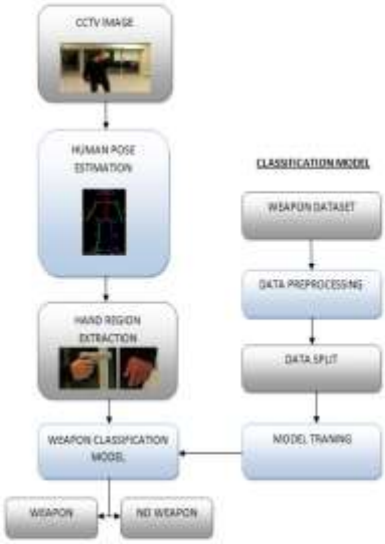

FIG 2. Proposed Workflow

5.1 HUMAN POSE ESTIMATION:

FIG 3. Model pipeline from the paper

The methodology hinges on the integration of the OpenPose framework [15], which serves as a cornerstone in our approach. It commences with an analysis conducted by a CNN initialized with the initial 10 layers of VGG-19², resulting in the derivation of a series of feature maps (F). These feature maps lay the groundwork for subsequent stages of processing. Following this, the multi-stage CNN architecture is employed to perform a forward pass, aiming to predict two vital components: 2D confidence maps denoting the locations of body parts, and 2D vector fields representing Part Affinity Fields (PAFs), encapsulating the relationships between various body parts. This iterative process entails the continuous refinement of predictions for both confidence maps and PAFs, with the introduction of intermediate supervision at each stage to steer the learning process effectively. After the prediction phase, a technique known as non-maximum suppression is applied to the confidence maps to yield discrete part candidate locations, thereby paving the way for further analysis. Concurrently, PAFs play a pivotal role in guiding the association of candidate limbs, facilitating the selection of optimal matches between limbs and individuals depicted in the image. Despite the inherent complexity of the task, characterized by its NP-hard nature, a method proposed by the authors of the paper, based on a greedy relaxation approach, consistently yields high-quality matches, effectively addressing the challenging matching problem. This comprehensive methodology ensures the accurate and robust estimation of human body poses from input images, thus bolstering the effectiveness of our approach in weapon detection.

5.2 HAND REGION EXTRACTION:

FIG 4. Hand region Extraction

In the subsequent phase of our methodology, we leverage the extracted pose data to infer and extract hand regions for each detected individual. Utilizing the positional information of the elbows and wrists obtained from the pose estimation process, alongside calculations of distances and directions between these points, a comprehensive set of bounding boxes encompassing all hand regions within the input image is generated. This pivotal step ensures the precise delineation of potential hand-held objects, as depicted in figure 4 of our framework. The adjustability of the scale parameter in the code provides flexibility in defining the size of the hand bounding boxes, enhancing the adaptability of our approach.

5.3 WEAPON CLASSIFICATION:

After deriving hand regions through the utilization of OpenPose, the subsequent phase revolves around meticulous preprocessing to optimize the hand region images before their integration into a sophisticated classification model. This pivotal segment in the pipeline encompasses the utilization of various Convolutional Neural Network (CNN) architectures, ranging from customized designs to the exploitation of cutting-edge pre-trained models such as ResNet and EfficientNet. Following extensive experimentation and iterative refinement, it was ascertained that ResNet emerged as the frontrunner, showcasing superior accuracy and robustness compared to its counterparts. The classification endeavour is primarily centred on discerning hand-held objects, encompassing a diverse array of items including pistols, knives, bills, purses, smartphones, and cards. These categories were meticulously curated to encapsulate objects frequently held by individuals, thereby mitigating false positives and minimizing detection errors significantly. By homing in on these handheld objects, the classification model not only augments the accuracy of weapon detection but also fortifies the overall system reliability and efficiency in identifying potential security threats within surveillance footage. This symbiotic amalgamation of advanced deep learning methodologies with contextually relevant insights serves as the cornerstone in bolstering the effectiveness of automated surveillance systems, thus upholding public safety and security endeavours. This comprehensive approach, informed by prior methodologies and contextual intricacies, not only empowers the system to identify potential security threats with unparalleled precision but also fosters a safer and more secure environment for all stakeholders involved.

6. RESULTS:

The assessment of model performance encompasses various metrics, including accuracy, precision, and F1 score, each providing distinct insights into its accuracy and efficacy. Accuracy gauges the overall correctness of the model by determining the ratio of correctly predicted instances to the total instances. Precision quantifies the ratio of true positive predictions to the total number of positive predictions.

Precision, denoted as P, is computed using the formula,

$$\text{Precision (P)} = \frac{TP}{TP+FP} \quad (1)$$

representing the proportion of actual positives correctly classified by the model, also referred to as the true positive rate (TPR).

Recall, denoted as R, is calculated as

$$\text{Recall ®} = \frac{TP}{TP+FN} \quad (2)$$

indicating how many of the actual positives are accurately classified by the model. The F1 score, a harmonic mean of precision and recall, offers a balanced assessment by considering false positives and false negatives, particularly valuable for imbalanced datasets. It is computed using the formula ,

$$\text{F1 score} = \frac{2*(P*R)}{(P*R)} \quad (3)$$

	Resnet			Efficientnet		
Class	Precision	Recall	F1-Score	Precision	Recall	F1-Score
Money	0.95	0.96	0.95	0.95	0.90	0.93
Knife	0.98	0.98	0.98	0.94	0.97	0.96
Wallet	0.93	0.93	0.93	0.89	0.90	0.89
Pistol	0.98	0.98	0.98	0.96	0.97	0.97
Smart phone	0.96	0.92	0.94	0.94	0.87	0.91
Card	0.85	0.92	0.89	0.87	0.88	0.87

Table 2 Performance Metrics

FIG 5. Confusion Matrix Resnet

The results obtained from the ResNet and EfficientNet models are shown in table 2. Resnet showcases its outstanding performance in hand region classification. With an accuracy of 96.39%, ResNet demonstrates superior accuracy compared to EfficientNet. The confusion matrix reveals precise classification across various classes, with notable precision rates for pistols (0.98), knives (0.98), and smartphones (0.96). The model exhibits high recall and F1-score values, emphasizing its reliability and effectiveness in detecting weapons.

In contrast, the EfficientNet model achieved an accuracy of 94.24%, slightly lower than ResNet, shown in table 2. While EfficientNet demonstrates satisfactory performance, its confusion matrix indicates slightly lower precision and recall values compared to ResNet. Notably, the model exhibits lower precision rates for certain classes, such as billete and smartphone. However, EfficientNet still maintains high precision and recall for knives and pistols, crucial objects for surveillance. The model's weighted average F1-score of 0.94 reflects its overall effectiveness in hand region classification, although ResNet outperforms it in accuracy and precision.

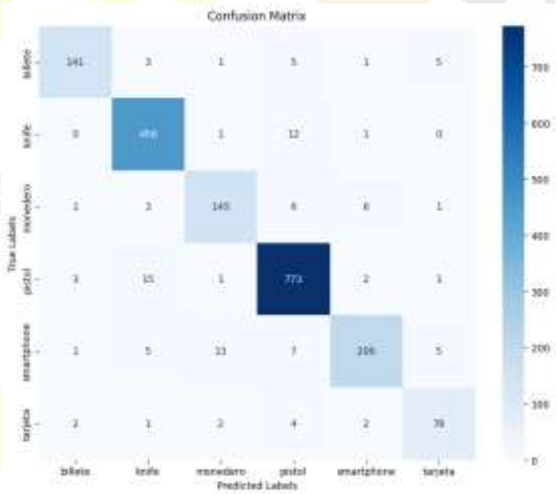

FIG 6. Confusion Matrix Efficientnet

The findings indicate that ResNet stands out as the preferred option for hand region classification in weapon detection systems.

7. Conclusion :

In conclusion, the research demonstrates the effectiveness of leveraging advanced technologies to enhance security measures. Through the integration of computer vision techniques, deep learning models, and body pose analysis, the proposed methodology offers a robust approach to detecting firearms and other dangerous objects in surveillance videos. The utilization of techniques such as Open Pose for body pose estimation, coupled with CNN-based models like ResNet for hand region classification, has shown promising results in accurately identifying handguns, knives, and other potential threats. The findings underscore the significance of automated weapon detection systems in bolstering surveillance capabilities and mitigating security risks. Future research could explore the incorporation of real-time processing and edge computing for enhanced efficiency in weapon detection, as well as the integration of multimodal data sources for comprehensive threat assessment. Additionally, efforts towards developing scalable and adaptable models for diverse environmental conditions and object variations would further advance the field of weapon detection and contribute to the broader goal of ensuring public safety and security.

REFERENCE:

1. Olmos, R., Tabik, S., & Herrera, F. (2018). Automatic handgun detection alarm in videos using deep learning. *Neurocomputing*, *275*, 66-72. https://doi.org/10.1016/j.neucom.2017.05.012
2. Kaya, V., Tuncer, S., & Baran, A. (2020). Detection and Classification of Different Weapon Types Using Deep Learning. *Applied Sciences*, *11*(16), 7535. https://doi.org/10.3390/app11167535
3. Dwivedi, N., Singh, D. K., & Kushwaha, D. S. (2019, December). Weapon classification using deep convolutional neural network. In *2019 IEEE Conference on information and communication technology* (pp. 1-5). IEEE.
4. Egiazarov, A., Mavroeidis, V., Zennaro, F. M., & Vishi, K. (2019, November). Firearm detection and segmentation using an ensemble of semantic neural networks. In *2019 European Intelligence and Security Informatics Conference (EISIC)* (pp. 70-77). IEEE.
5. Abdallah, H. B., Abdellatif, T., & Chekir, F. (2018). AMSEP: automated multi-level security management for multimedia event processing. *Procedia computer science*, *134*, 452-457.
6. Grega, M., Matiolański, A., Guzik, P., & Leszczuk, M. (2016). Automated detection of firearms and knives in a CCTV image. *Sensors*, *16*(1), 47.
7. Tiwari, R. K., & Verma, G. K. (2015). A computer vision based framework for visual gun detection using harris interest point detector. *Procedia Computer Science*, *54*, 703-712.
8. Xu, Z., Tian, Y., Hu, X., & Pu, F. (2015, September). Dangerous human event understanding using human-object interaction model. In *2015 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)* (pp. 1-5). IEEE.
9. Yadav, P., Gupta, N., & Sharma, P. K. (2023). A comprehensive study towards high-level approaches for weapon detection using classical machine learning and deep learning methods. *Expert Systems with Applications*, *212*, 118698.
10. Ahmed, S., Bhatti, M. T., Khan, M. G., Lövström, B., & Shahid, M. (2022). Development and optimization of deep learning models for weapon detection in surveillance videos. *Applied Sciences*, *12*(12), 5772.
11. Ashraf, A. H., Imran, M., Qahtani, A. M., Alsufyani, A., Almutiry, O., Mahmood, A., ... & Habib, M. (2022). Weapons detection for security and video surveillance using cnn and YOLO-v5s. *CMC-Comput. Mater. Contin*, *70*, 2761-2775.
12. Mehmood, A. (2021). Abnormal behavior detection in uncrowded videos with two-stream 3D convolutional neural networks. *Applied Sciences*, *11*(8), 3523.
13. Alairaji, R. A. M., Aljazaery, I. A., & ALRikabi, H. T. S. (2022). Abnormal behavior detection of students in the examination hall from surveillance videos. In *Advanced Computational Paradigms and Hybrid Intelligent Computing: Proceedings of ICACCP 2021* (pp. 113-125). Springer Singapore.
14. Pérez-Hernández, F., Tabik, S., Lamas, A., Olmos, R., Fujita, H., & Herrera, F. (2020). Object detection binary classifiers methodology based on deep learning to identify small objects handled similarly: Application in video surveillance. *Knowledge-Based Systems*, *194*, 105590.
15. Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. (2017). Realtime multi-person 2d pose estimation using part affinity fields. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7291-7299).
16. J. Ruiz-Santaquiteria, A. Velasco-Mata, N. Vallez, G. Bueno, J. A. Álvarez-García and O. Deniz, "Handgun Detection Using Combined Human Pose and Weapon Appearance," in IEEE Access, vol. 9, pp. 123815-123826, 2021, doi: 10.1109/ACCESS.2021.3110335.