# Computer Vision for Weapon Detection in Educational Environments

## A Systematic Literature Review


Maurício Rodrigues Lima
mauricio.rodrigues@discente.ufg.br
Federal University of Goias
Institute of Informatics
Goiânia, Goiás, Brazil

Deller James Ferreira
deller@inf.ufg.br
Federal University of Goiás
Institute of Informatics
Goiânia, Goiás, Brazil
djamesf@email.com

Elisângela Silva Dias
elisangelasd@gmail.com
Federal University of Goiás
Institute of Informatics
Goiânia, Goiás, Brazil
elisangela@inf.ufg.br

Marcos Reges Mota
regesmota@discente.ufg.br
Federal University of Goiás
Institute of Informatics
Goiânia, Goiás, Brazil

Ana Luísa de Bastos Chagas
analuisa23@discente.ufg.br
Federal University of Goiás
Institute of Informatics
Goiânia, Goiás, Brazil

Pedro Lemes Sixel Lobo
pedro_lemes@discente.ufg.br
Federal University of Goiás
Institute of Informatics
Goiânia, Goiás, Brazil



## ABSTRACT

This study presents a systematic review of the literature on the use of computer vision algorithms for weapon detection in educational environments. Through the analysis of 13 selected studies from an initial corpus of 10,519 articles, the results demonstrate that models based on Convolutional Neural Networks, particularly variants of YOLO, are predominantly used due to their high accuracy and real-time efficiency. This work highlights the need for technological advancements to address challenges such as the variability of weapon types and the diverse school scenarios. Furthermore, the practical implications of these technologies in enhancing school security and the importance of ethical and privacy considerations are discussed. The review also reveals significant gaps in current research, such as the lack of studies focused on specific educational environments and the need for more representative and diverse datasets.



## KEYWORDS

deep learning, computer vision, gun detection, educational environments


## 1 INTRODUCTION

Although numerous technological advancements have occurred across various domains, there are still significant threats that people face, one of which is the frequent occurrence of school shootings in a large number of countries (especially in the USA), thereby threatening the lives of young children [14]. In recent years, law enforcement agencies and security professionals have paid much attention to the use of artificial intelligence for criminal detection based on video surveillance. The ability of deep learning models to automatically detect and track potential thrusts saves time and money for law enforcement organizations, allowing them to understand complex data patterns [9]. Currently, many real-world problems require the detection of multiple objects in images or videos. Armed violence imposes a significant impact on public health, psychological well-being, and economic costs. It leads to the loss of many lives each year. Object detection construction is challenging when the objects are small, like handheld weapons [10].

Security in educational settings is a growing concern globally, especially due to the increase in incidents of armed violence in schools. Weapon detection is a very serious and intense issue regarding security and public protection. It is a challenging task, especially when it needs to be done automatically or with some AI model. Various object detection models are available, but weapon detection is complicated due to the size and varied shapes of weapons, along with different background colors [11].

Due to the sensitive context and the need for high accuracy in educational environments, precise and fast detection is critical to minimize false positives and ensure the safety of students. This requires models trained to identify weapons from different angles and under various conditions, avoiding false alarms that could cause undue panic. Consequently, real-time detection is particularly challenging in educational settings due to the need for quick reactions without compromising accuracy. Models like YOLOv4 have proven effective but still face difficulties in dynamic contexts and with small objects [5].


In: I Workshop de Revisões Sistemáticas de Literatura em Sistemas Multimídia e Web (WRSL+ 2024). Anais Estendidos do XXX Simpósio Brasileiro de Sistemas Multimídia e Web (WRSL+'2024). Juiz de Fora/MG, Brazil. Porto Alegre: Brazilian Computer Society, 2024.



ISSN 2596-1683

Educational environments often have crowds and busy situations, which can make weapon detection difficult. Models need to be adapted to handle these complexities, such as variations in the size and visibility of weapons in surveillance videos [18].

Using a systematic literature review method, this study analyzed 13 studies extracted from 10,519 drawn from databases such as ACM Digital Library, EI Compendex, IEEE Digital Library, ISI Web of Science, Science Direct, Capes Periodicals, Scopus, Springer Link, Scielo, and Sol SBC. Results show a need to diversify technological approaches in weapon detection in educational environments.

This study is organized as follows: theoretical foundations and related works are presented in Sections 2 and 3, respectively. Planning and execution of the literature review in Section 4; results and responses to research questions in Section 5; discussions on the findings in Section 6; and conclusions in Section 7.

## 2 BACKGROUND

This section provides the theoretical foundation necessary to understand the methods and technologies used in weapon detection using computer vision. It is divided into subsections detailing different relevant areas, presenting concepts on computer vision, machine learning, deep learning, Object Detection Algorithms, and performance evaluation metrics. Furthermore, it briefly discusses ethical aspects.

### 2.1 Computer Vision

Computer vision is the field of artificial intelligence dedicated to extracting information from digital images [27], aiming to emulate human vision. Therefore, it also takes an image as input, but the output is an interpretation of the image as a whole or in part [21]. It is the science responsible for machine vision, the way a computer perceives its surroundings, extracting meaningful information from images captured by video cameras, sensors, scanners, and other devices. This information allows the recognition, manipulation, and reasoning about the objects that make up an image [4].

### 2.2 Machine Learning

Machine learning, as a branch of artificial intelligence, focuses on the development of techniques and algorithms that enable computers to learn and make predictions or decisions based on data. Instead of being directly programmed to perform a specific task, machine learning systems use algorithms to identify subtle patterns in the provided data, improving their performance over time and optimizing results.

Machine learning methods are generally classified into supervised learning, unsupervised learning, and reinforcement learning. In supervised learning, the model is trained with a labeled dataset, while in unsupervised learning, the algorithm attempts to uncover hidden patterns in unlabeled data. Reinforcement learning involves an agent learning to make decisions through interactions with the environment, receiving rewards or penalties based on its actions.[22].

### 2.3 Deep Learning

Deep learning is a subfield of machine learning that utilizes deep neural networks composed of multiple layers and a large number of parameters. Deep learning, which has been growing in use in recent years, consists of a series of layers of non-linear processing units that enable the automatic extraction of features and understanding of patterns in large volumes of information [28].

Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. These methods have dramatically improved the state-of-the-art in speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics. Deep learning discovers intricate structure in large data sets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer. Deep convolutional nets have brought about breakthroughs in processing images, video, speech and audio, whereas recurrent nets have shone light on sequential data such as text and speech [17].

### 2.4 Object detection algorithms

Object detection algorithms, such as YOLO (You Only Look Once), SSD (Single Shot MultiBox Detector), and R-CNNs (Region-Based Convolutional Neural Networks), are widely used in security applications, such as weapon detection, due to their high effectiveness and precision.

The YOLO algorithm is renowned for its capability to detect objects in real-time by analyzing the entire image in a single evaluation. This is achieved by dividing the image into a grid and making simultaneous predictions of bounding boxes and class probabilities for each cell. YOLO's speed and accuracy make it particularly suitable for surveillance systems where rapid detection is essential.

The SSD combines features of both YOLO and R-CNN algorithms by using a single network to predict bounding boxes and class scores at multiple scales of feature maps. This approach allows SSD to be both fast and accurate, making it suitable for systems requiring immediate response, such as real-time security cameras.

R-CNNs, including variants such as Fast R-CNN, Faster R-CNN, and Mask R-CNN, utilize a two-step approach: initially proposing regions of interest in the image, and then classifying each region and refining the bounding boxes. Although they are slower, R-CNNs are highly precise, making them valuable for forensic analysis of security videos where accuracy is more critical than speed.

These algorithms are chosen for weapon detection due to their ability to identify objects with high precision and speed. The choice of algorithm depends on the specific needs of the system: YOLO and SSD are preferred for real-time applications due to their speed, while R-CNNs are used when detailed accuracy is essential. The combination of these technologies provides a robust solution for weapon detection, ensuring a rapid and precise response in various security situations.

## 2.5 Image processing

Image processing and video analysis are essential areas in the field of computer vision, encompassing a variety of techniques crucial for the preparation and analysis of visual data. Basic image processing techniques, such as filtering, segmentation, and morphological transformations, are frequently employed to enhance image quality and facilitate object detection. Image filtering allows for the removal of noise and the enhancement of important features, which is vital for subsequent data analysis.

Additionally, image segmentation, which involves dividing an image into distinct regions, plays a fundamental role in identifying and classifying objects within a scene. Advanced segmentation techniques using convolutional neural networks have significantly increased accuracy in object detection in videos. On the other hand, morphological transformations, which include operations such as erosion and dilation, are used to modify the geometric structure of images, facilitating the analysis of shapes and contours. Such transformations are essential for manipulating and analyzing complex structures in binary and grayscale images.

When effectively integrated, these techniques form the basis for developing robust and efficient detection algorithms. The combination of filtering, segmentation, and morphological transformations provides appropriate data preparation, which is crucial for applying advanced detection algorithms in various contexts such as security, monitoring, and industrial automation.

## 2.6 Performance evaluation metrics

Performance evaluation metrics are essential for measuring the effectiveness of constructed models. Three significant metrics in the context of this work are precision, recall, F1-score, and mAP (mean Average Precision).

Recall, also known as Sensitivity in psychology, is the proportion of actual positive cases that are correctly predicted as positive. It measures how well the "Predicted Positive" rule covers the actual positive cases. Conversely, Precision, also known as Confidence in data mining, refers to the proportion of predicted positive cases that are actually true positives [23].

The F1 score is a metric used to evaluate the accuracy of a classification model, particularly in cases where the class distribution is imbalanced. It is the harmonic mean of precision and recall, providing a balance between these two aspects. Precision measures the accuracy of the positive predictions, while recall assesses how well the model identifies actual positives. The F1 score reaches its best value at 1 (perfect precision and recall) and its worst at 0. It is particularly useful when you need a single metric to compare models with respect to how they handle both false positives and false negatives [12].

## 2.7 Ethical Concepts

Video surveillance raises significant ethical and privacy issues that warrant careful examination. The primary concern lies in the invasion of individual privacy, as the continuous collection and monitoring of images can compromise individuals' freedom and privacy. Studies indicate that surveillance in public places, although intended for security, can lead to a feeling of constant observation, which affects behavior and personal autonomy [19].

Consent is another crucial aspect. In many cases, surveillance is implemented without the explicit consent of the monitored individuals, raising ethical questions about the legitimacy and transparency of these processes. The lack of informed consent can result in a violation of individuals' privacy and autonomy rights, and can generate distrust toward the institutions employing these technologies.

Additionally, the issue of false positives and false negatives is relevant to the ethical discussion. False positives occur when a surveillance system incorrectly identifies an individual as a threat, while false negatives occur when a real threat is not identified. Both cases have serious implications: false positives can lead to unjust punitive actions, while false negatives can result in failures in public safety protection. The accuracy of detection algorithms must be continuously assessed and improved to minimize these errors and their ethical consequences).

Therefore, video surveillance requires a delicate balance between the need for security and the protection of individual rights. Ethical and privacy considerations must be integrated into the development and implementation of these technologies to ensure that individuals' rights are respected and that surveillance systems are used in a fair and responsible manner.

# 3 RELATED WORKS

The study by [26] performs a systematic review on weapon detection in surveillance footage using deep learning. The main focus is to identify the methods used, the main features of existing datasets, and the major challenges in the area of automatic weapon detection. Additionally, [2] presents a systematic review of published studies on object detection using machine learning in the context of security, covering the period from 2010 to 2019, in which they collected and analyzed 73 articles from relevant journals and conferences.

As of the time this systematic literature review was conducted, no other review has sought to analyze computer vision algorithms applied to the security of educational environments.

# 4 SYSTEMACTIC LITERATURE REVIEW

This paper presents a Systematic Literature Review. Systematic reviews are characterized by a methodical and replicable methodology, involving a comprehensive search to locate all published and unpublished works on a particular topic; a systematic integration of the results of this search; and a critical analysis of the scope, nature, and quality of the evidence in relation to a specific research question. They synthesize studies to draw broad theoretical conclusions about the significance of a literature, connecting theory and evidence, as well as evidence and theory [29].

The protocol, however, was planned in light of the best practices for conducting literature reviews [16], structured in three well-defined stages: (i) planning, (ii) conducting, and (iii) reporting of results, as follows. The communication of the results is presented in Section 5.

## 4.1 Planning

The objective is to analyze publications on deep learning algorithms in computer vision applied to weapon detection in school environments. From this objective, the following **research questions (RQs) were derived:**

***RQ1: What computer vision algorithms are most commonly used for weapon detection in school environments and other public spaces, and what is their accuracy?***

***RQ2: What improvements or models of computer vision algorithms can be implemented to increase efficiency in weapon detection in school environments? Are the latest algorithms and improvements being implemented in these scenarios?***

***RQ3: How do the implementation and results of computer vision algorithms for weapon detection impact the environment in which the algorithm was deployed?***

**Search Strategy and Sources.** This study selection process was carried out using a search string, developed to include works published in both Portuguese and English. The search string used was:

**Portuguese:** *(("visão computacional" OR "processamento de imagens" OR "reconhecimento de padrões") AND ("armas") AND ("ambiente escolar" OR "ambientes educacionais" OR "escolas" OR "universidades") AND ("detecção" OR "reconhecimento"))*;

**English:** (("computer vision" OR "image processing" OR "pattern recognition") AND ("weapons") AND ("educational environments" OR "schools" OR "colleges" OR "universities") AND ("detection" OR "recognition"))

The search was conducted considering studies that contained the search string keywords in their title and abstract. The databases used were ACM Digital Library, EI Compedex, IEEE Digital Library, ISI Web of Science, Science Direct, Periódicos Capes, Scopus, Springer Link, Scielo, and Sol SBC.

**Study Selection Criteria.** The defined inclusion criteria (IC) were as follows:

- IC1: The study explicitly addresses weapon detection;
- IC2: The study includes in its abstract or title a connection with the research objective.

On the other hand, the defined exclusion criteria for studies were:

- EC1: The study is a previous version of a study already selected;
- EC2: The study is duplicated;
- EC3: The study is not a primary study;
- EC4: The study is less than 4 pages long;
- EC5: The study is not fully available;
- EC6: The study is not written in Portuguese or English;
- EC7: The study is gray literature (such as conference preface).

## 4.2 Conducting

The study selection process for this review was conducted as shown in Figure 1.

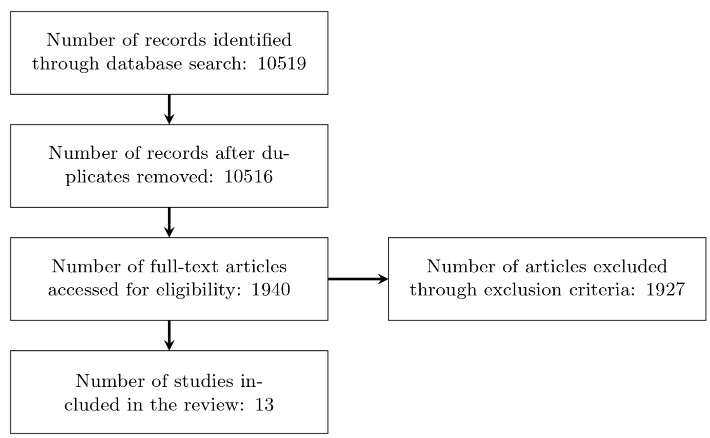


**Figure 1: Study selection process.**

The search was conducted between May 1st and May 30th, 2024. *Parsifal* was used for planning, execution, conduction, and documentation of the research. This tool allows researchers to set goals and objectives, import articles using BibTeX files, eliminate duplicates, establish selection criteria, and generate reports [7].

# 5 RESULTS

This section addresses the results obtained from the systematic literature review. Table 1 presents the selected studies, detailing the method used, the scenario, the metrics employed, and the results achieved based on these metrics.

**Table 1: Analyzed Studies**

| No. | Method | Scenario | Metrics | Results |
|---|---|---|---|---|
| [3] | Yolov4 | Public places | mAP and IOU | mAP: 90.35%, IOU: 73.68% |
| [1] | Scaled-YOLOv4 | Public places monitored by CCTV | mAP | mAP: 92.1%, FPS: 85.7 |
| [8] | CNN, Semantic Segmentation | Public use and camera monitoring | Accuracy | Accuracy: 95.2% in tests |
| [13] | Fuzzy Logic, material testing | Areas monitored by CCTV | False positive rate | Significant reduction in false alarm rates |
| [6] | Deep Learning | Public places | Accuracy, precision | Accuracy: 93%, Precision: 91% |
| [15] | CNN, YOLOv3, HAAR cascade classifier | Schools | Precision, Recall | Precision: 96%, Recall: 94% |
| [20] | Neural networks | Public places | Accuracy, detection speed | Accuracy: 91%, Speed: 0.5s per image |
| [31] | Deep Learning | Schools, hallways | False positive rate (FPR), Precision | Reduction in FPR to 37.9% while maintaining detection capability |
| [24] | YOLOv3, YOLOv5 | Public places | Precision, Recall | Precision: 92.2%, Recall: 85.3% |
| [25] | Deep Learning, YOLOv4 | Urban areas | Accuracy, Recall | Accuracy: 93.7%, Recall: 90.1% |
| [30] | CNN, YOLOv7 | Public places | Precision, Recall | Precision: 95.78%, Recall: 96.92% |
| [33] | Deep Neural Networks | Public spaces | AP | AP improved by 17.5% compared to the baseline detector |
| [34] | YOLOv4, Night vision | Outdoor night areas | Precision, Recall | Precision: 94%, Recall: 92% |

**QP1: What computer vision algorithms are most commonly used for weapon detection in educational environments and other public settings, and what is their accuracy?** Among the selected articles, only 2 explicitly use deep learning algorithms for weapon detection in educational settings. Isaac Ritharson et al. [15] uses YOLOv3 for weapon detection, combining it with facial recognition techniques (HAAR cascade classifier) and color detection to identify individuals and ensure they are not carrying weapons. The work highlights YOLOv3 for its real-time object detection capability, which is effective in identifying weapons in videos and images, crucial for implementation in school environments where rapid detection can prevent violent incidents.

Vallez et al. [31] Proposes an innovative approach using a deep autoencoder to reduce the false positive rate in weapon detection systems. The autoencoder is trained with typical false positives from a specific scenario to learn to filter them out, improving the system's accuracy without compromising its detection capability. This method is tested using a synthetic dataset created with Unreal Engine 4, showing a reduction of up to 37.9% in false positives while maintaining detection capability.

**QP2: What improvements or which models of computer vision algorithms can be implemented to increase efficiency in weapon detection in educational environments? Are the most current algorithms with the latest improvements deployed in these scenarios?**

The development of weapon detection methods in educational settings has evolved significantly, with various studies highlighting the integration of advanced technologies to enhance security. Initially, Ineneji and Kusaf [13] emphasizes the importance of fuzzy logic and material testing systems, which, combined with computer vision algorithms, can significantly improve accuracy and reduce false positives. This method is crucial, especially in school environments, where accuracy is a key factor.

Moving forward with this technology, Isaac Ritharson et al. [15] describes the use of convolutional neural networks (CNN) with the YOLOv3 algorithm, emphasizing the possibility of incorporating more advanced models such as YOLOv4 to enhance accuracy and detection speed. The implementation of complementary technologies, such as facial recognition and color detection, also helps differentiate between students and strangers, thus increasing school security.

Furthermore, Manikandan and Rahamathunnisa [20] and Ahmed et al. [1] highlight the effectiveness of neural networks and the importance of optimizing deep learning models like Scaled-YOLOv4. Such techniques not only improve real-time detection but also reduce false positive rates, contributing to a safer school environment.

The practical implementation of these technologies is illustrated by the SaveLives system, mentioned by Anand and Koshariya [3], which uses YOLOv4 to detect weapons and instantly notify authorities, reducing the response time to threats. Similarly, Bhatti et al.

[6] discusses the use of deep learning to monitor CCTV footage in real-time, suggesting that the inclusion of newer algorithms like YOLOv4 and YOLOv5 can significantly enhance security.

The approach based on semantic segmentation, as discussed by Egiazarov et al. [8], and advancements in object detection algorithms with YOLOv7, mentioned by Sivakumar et al. [30], are other examples of how innovative technologies are being adapted to improve weapon detection in educational settings. These techniques enhance the ability to identify weapons even in challenging conditions, such as poorly lit areas, as suggested in Yadav et al. [34].

Finally, the importance of accuracy and rapid response capability is emphasized in studies like those by Rosales et al. [24] and Ruiz-Santaquiteria et al. [25], which demonstrate how the combination of deep learning techniques can increase both accuracy and recall rate in weapon detection. The implementation of these techniques in schools, as discussed by Velasco-Mata et al. [33], not only improves the effectiveness of weapon detection but also helps prevent violent incidents, reinforcing the safety and protection of students and staff.

**QP3: How does the implementation and outcomes of computer vision algorithms for weapon detection impact the environment in which the algorithm is deployed?**

The integration of technological advancements in school security has been a significant focus of recent studies, revealing remarkable progress in weapon detection. Initially, Ineneji and Kusaf [13] discusses the use of fuzzy logic and material testing in conjunction with computer vision algorithms to enhance accuracy and minimize false positives, a crucial measure in educational settings where precision is vital. Extending this approach, Isaac Ritharson et al. [15] highlights the employment of convolutional neural networks and the YOLOv3 algorithm, supplemented by facial recognition and color detection, to distinguish between students and intruders, thus enhancing safety.

The article by Manikandan and Rahamathunnisa [20] suggests that neural networks specifically tuned for weapon detection are highly effective, a notion corroborated by Ahmed et al. [1] who emphasizes optimizing deep learning models like Scaled-YOLOv4 for this purpose. These techniques not only increase accuracy but also detection speed, as evidenced by the SaveLives system mentioned by Anand and Koshariya [3], which employs YOLOv4 for instant notifications to authorities.

Security can be further reinforced through advanced algorithms like YOLOv4 and YOLOv5, discussed by Bhatti et al. [6], and the innovative approach based on semantic segmentation to detect concealed weapons, presented in Egiazarov et al. [8]. Moreover, Sivakumar et al. [30] introduces improvements in object detection algorithms, applying YOLOv7 to enhance weapon detection efficiency in educational environments.

The potential to reduce false alarms through the use of deep learning is highlighted in Vallez et al. [31]. At the same time, combining weapon detection with human pose information, as discussed by Velasco-Mata et al. [33], can significantly enhance the accuracy of security systems. This approach is complemented by Yadav et al. [34], addressing weapon detection in low-light conditions using YOLOv4, a crucial aspect for less illuminated areas in schools.

Finally, Rosales et al. [24] and Ruiz-Santaquiteria et al. [25] highlight the use of advanced convolutional neural networks to detect threats in real-time, enhancing the capability for quick and accurate responses to potential incidents, which is fundamental for public safety in educational environments. This sequence of innovations demonstrates the continuous progression in school security technology, emphasizing the importance of the constant evolution of techniques to keep educational environments safe.

## 6 DISCUSSION

The results obtained in this systematic review indicate that weapon detection in educational environments through computer vision is an evolving field, with the predominant adoption of Convolutional Neural Networks (CNNs) and YOLO models due to their demonstrated efficacy in real-time. However, these results also underline the complexity associated with developing detection systems that are simultaneously fast, accurate, and capable of operating under variable conditions.

**Comparison of Algorithms and Efficacy:** Studies such as [15] and [32] highlight the superiority of YOLOv3 and YOLOv4 in terms of precision and speed, crucial for the prevention of violent incidents in real-time. However, the effectiveness of these models can be compromised in low-light scenarios or when objects are partially obscured, challenging researchers to develop significant improvements or hybrid approaches.

**Variations in Results Based on Scenarios:** The variability of scenarios, such as the type of educational environment (primary versus university) and population density, directly influences the efficacy of the models. For example, the study [8] suggests that semantic segmentation may provide improvements in detection in environments with complex backgrounds, which is typical in schools with high student activity.

**Practical Implications for School Safety:** The implementation of these detection systems not only enhances physical security but also contributes to the psychological well-being of students and staff. However, the installation of these systems should be accompanied by clear policies on privacy and ethics, ensuring that monitoring is conducted respecting individual rights.

**Practical Implications for School Safety: Challenges and Future Considerations:** One of the main challenges identified is the detection of unconventional or modified weapons, which may not be easily recognizable by models trained on standard datasets. Additionally, the generalization of models to different cultures and legal regulations about weapons requires a more adaptive and customizable approach.

# 7 FUTURE WORKS

For future studies, it is important that concerns are addressed not only with respect to technical aspects such as the development of deep learning algorithms, but also with validations and training of people with simulations of real scenarios.

**Development of Specific Datasets:** There is a critical need to develop and publicly make available more robust and specific datasets that include a wider variety of weapon types, different school settings, and diverse lighting conditions. This would help to improve the accuracy of existing models and to test their effectiveness in scenarios closer to reality.

**Adaptive Algorithms:** Future research should focus on the development of algorithms that dynamically adapt to changes in the environment, such as changes in lighting or in the arrangement of spaces. This is especially important in educational environments, where conditions can vary significantly throughout the day or the academic year.

**Integration of Multimodal Technologies:** Exploring the integration of computer vision with other technologies, such as audio analysis and sensory data, could lead to more comprehensive and effective detection systems that do not rely solely on visual images.

**Ethical and Privacy Considerations:** It is essential that future research also focus on addressing the ethical and privacy implications of implementing surveillance technologies in educational environments. Developing guidelines and frameworks that balance security and privacy will be crucial for the acceptance and effective implementation of these technologies.

**Validation in Real Environments:** We encourage the conduct of field studies to validate weapon detection models in actual school environments. This includes conducting pilot tests in different types of schools to assess the efficacy and practicality of the proposed systems.

**Development of Automated Responses:** In addition to detecting weapons, it is important to develop systems that can automate the response to detected threats. This includes integration with emergency communication systems and school security protocols.

**Training and Empowerment:** Investing in training programs for school administrators and security personnel on how to effectively operate and respond to weapon detection alerts is crucial. This will help ensure that the technology is used effectively and responsibly.

# 8 CONCLUSION

This study conducted a systematic review of the literature on computer vision algorithms applied to weapon detection in educational environments, identifying the main methods used and highlighting the challenges faced in this critical research area. The analysis showed that, although technologies based on Convolutional Neural Networks, especially variants of YOLO, are widely used due to their efficiency and real-time operation capability, there are several significant limitations and challenges that still need to be overcome.

**Technical and Ethical Challenges:** The main challenges include variability in the shapes, sizes, and concealment methods of weapons; variable environmental conditions that can interfere with the accuracy of systems; the need to minimize false positives and negatives; balancing security and privacy; the requirement for real-time performance to effectively prevent incidents; and the adaptation and scalability of systems to keep up with new technologies and emerging threats.

These challenges point to the need for continuous development of computer vision technologies that are not only advanced in technical terms but also sensitive to the complex ethical and social issues surrounding their application in educational environments. Additionally, privacy and ethical concerns are of utmost importance, suggesting the need for more stringent guidelines and discussions on balancing security and individual rights.

**Future Directions:** Future research should focus on improving algorithms to handle the variability of environmental conditions and weapons, and on integrating multimodal technologies that can complement computer vision to create more robust and reliable detection systems. Furthermore, it is important to develop and implement automated response systems to address detected threats efficiently and ethically.

# REFERENCES


[1] Soban Ahmed, Muhammad Tahir Bhatti, Muhammad Gufran Khan, Benny Lövström, and Muhammad Shahid. 2022. Development and Optimization of Deep Learning Models for Weapon Detection in Surveillance Videos. *Applied Sciences* 12, 12 (June 2022), 5772. https://doi.org/10.3390/app12125772

[2] Ali AlZaabi, Manar Abi Talib, Ali Bou Nassif, Ali Sajwani, and Omar Einea. 2020. A Systematic Literature Review on Machine Learning in Object Detection Security. In *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*. 136–139. https://doi.org/10.1109/ICCCA49541.2020.9250836

[3] Nishit Anand and Rupesh Kumar Koshariya. 2022. SaveLives - A Real-Time Threat Detection System. In *2022 1st International Conference on Informatics (ICI)*. IEEE, Noida, India, 114–119. https://doi.org/10.1109/ICI53355.2022.9786892

[4] D.H. Ballard and C.M. Brown. 1982. *Computer Vision.* Prentice-Hall. https://books.google.com.br/books?id=EfRRAAAAMAAJ

[5] Muhammad Tahir Bhatti, Muhammad Gufran Khan, Masood Aslam, and Muhammad Junaid Fiaz. 2021. Weapon Detection in Real-Time CCTV Videos Using Deep Learning. *IEEE Access* 9 (2021), 34366–34382. https://doi.org/10.1109/ACCESS.2021.3059170

[6] Muhammad Tahir Bhatti, Muhammad Gufran Khan, Masood Aslam, and Muhammad Junaid Fiaz. 2021. Weapon Detection in Real-Time CCTV Videos Using Deep Learning. *IEEE Access* 9 (2021), 34366–34382. https://doi.org/10.1109/ACCESS.2021.3059170

[7] Angela Carrera-Rivera, William Ochoa, Felix Larrinaga, and Ganix Lasa. 2022. How-to conduct a systematic literature review: A quick guide for computer science research. *MethodsX* 9 (2022), 101895. https://doi.org/10.1016/j.mex.2022.101895

[8] A. Egiazarov, F. Zennaro, and V. Mavroeidis. 2020. Firearm Detection via Convolutional Neural Networks: Comparing a Semantic Segmentation Model Against End-to-End Solutions. In *2020 IEEE International Conference on Big Data (Big Data)*. IEEE Computer Society, Los Alamitos, CA, USA, 1796–1804. https://doi.org/10.1109/BigData50022.2020.9377745

[9] U. G, I. Sudha, V. Lakshmanan, F. Shadrach, P. Krishnammal, and N. T. J. 2023. Crime Scene Object Detection from Surveillance Video by using Tiny YOLO Algorithm. In *2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN)*. IEEE Computer Society, Los Alamitos, CA, USA, 654–659. https://doi.org/10.1109/ICPCSN58827.2023.00114

[10] Rakesh Garg and Someet Singh. 2021. Intelligent Video Surveillance Based on YOLO: A Comparative Study. In *2021 International Conference on Advances in Computing, Communication, and Control (ICAC3)*. 1–6. https://doi.org/10.1109/ICAC353642.2021.9697321

[11] Tufail Sajjad Shah Hashmi, Nazeef Ul Haq, Muhammad Moazam Fraz, and Muhammad Shahzad. 2021. Application of Deep Learning for Weapons Detection in Surveillance Videos. In *2021 International Conference on Digital Futures and Transformative Technologies (ICoDT2)*. 1–6. https://doi.org/10.1109/ICoDT252288.2021.9441523

[12] Hao Huang, Haihua Xu, Xianhui Wang, and Wushour Silamu. 2015. Maximum F1-score discriminative training criterion for automatic mispronunciation detection. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 23, 4 (2015), 787–797.

[13] Collins Ineneji and Mehmet Kusaf. 2019. Hybrid weapon detection algorithm, using material test and fuzzy logic system. *Computers & Electrical Engineering* 78 (Sept. 2019), 437–448. https://doi.org/10.1016/j.compeleceng.2019.08.005

[14] P Isaac Ritharson, G Madhavan, M. Rajeswari, and D. Brindha. 2022. Prevention of School Shooting using Neural Networks and Computer Vision. In *2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT)*. 1703–1709. https://doi.org/10.1109/ICICICT54557.2022.9917656

[15] P Isaac Ritharson, G Madhavan, M. Rajeswari, and D. Brindha. 2022. Prevention of School Shooting using Neural Networks and Computer Vision. In *2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT)*. IEEE, Kannur, India, 1703–1709. https://doi.org/10.1109/ICICICT54557.2022.9917656

[16] B. Kitchenham and S. Charters. 2007. *Guidelines for performing systematic literature reviews in software engineering*. Technical Report EBSE 2007-001. Keele University and Durham University Joint Report.

[17] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. 2015. Deep learning. *nature* 521, 7553 (2015), 436–444.

[18] JunYi Lim, Md Istiaque Al Jobayer, Vishnu Monn Baskaran, Joanne Mun-Yee Lim, John See, and Koksheik Wong. 2021. Deep multi-level feature pyramids: Application for non-canonical firearm detection in video surveillance. *Eng. Appl. Artif. Intell.* 97 (2021), 104094. https://doi.org/10.1016/j.engappai.2020.104094

[19] David Lyon. 2018. *The Culture of Surveillance: Watching as a Way of Life.*

[20] V.P. Manikandan and U. Rahamathunnisa. 2022. A neural network aided attuned scheme for gun detection in video surveillance images. *Image and Vision Computing* 120 (April 2022), 104406. https://doi.org/10.1016/j.imavis.2022.104406

[21] Maurício Marengoni and Stringhini Stringhini. 2009. Tutorial: Introdução à visão computacional usando opencv. *Revista de Informática Teórica e Aplicada* 16, 1 (2009), 125–160.

[22] T. M. Mitchell. 1997. Does Machine Learning Really Work? *AI Magazine* 18, 3 (1997), 11–20. https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1303

[23] David Martin W. Powers. 2011. Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness and Correlation. *Journal of Machine Learning Technologies* 2, 1 (2011), 37–63.

[24] Maverick Poma Rosales, Ciro Rodriguez, Yuri Pomachagua, and Carlos Navarro. 2021. Firearm Detection in Images of Video Surveillance Cameras with Convolutional Neural Networks. In *2021 13th International Conference on Computational Intelligence and Communication Networks (CICN)*. IEEE, Lima, Peru, 61–65. https://doi.org/10.1109/CICN51697.2021.9574690

[25] Jesus Ruiz-Santaquiteria, Alberto Velasco-Mata, Noelia Vallez, Oscar Deniz, and Gloria Bueno. 2023. Improving handgun detection through a combination of visual features and body pose-based data. *Pattern Recognition* 136 (April 2023), 109252. https://doi.org/10.1016/j.patcog.2022.109252

[26] Tomás Santos, Hélder Oliveira, and António Cunha. 2024. Systematic review on weapon detection in surveillance footage through deep learning. *Computer Science Review* 51 (2024), 100612. https://doi.org/10.1016/j.cosrev.2023.100612

[27] Thiago Teixeira Santos, Jayme Garcia Arnal Barbedo, Sônia Ternes, J Camargo Neto, Luciano Vieira Koenigkan, and Kleber Xavier Sampaio de Souza. 2020. Visão computacional aplicada na agricultura. (2020).

[28] Pramila P. Shinde and Seema Shah. 2018. A Review of Machine Learning and Deep Learning Applications. In *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*. 1–6. https://doi.org/10.1109/ICCUBEA.2018.8697857

[29] Andy P Siddaway, A. Wood, and L. Hedges. 2019. How to Do a Systematic Review: A Best Practice Guide for Conducting and Reporting Narrative Reviews, Meta-Analyses, and Meta-Syntheses. *Annual review of psychology* 70 (2019), 747–770. https://doi.org/10.1146/annurev-psych-010418-102803

[30] H. Sivakumar, V. Arvind.R, P. V, and G. Balamurugan. 2023. Novel Deep Learning Pipeline for Automatic Weapon Detection. In *2023 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*. IEEE Computer Society, Los Alamitos, CA, USA, 1–5. https://doi.org/10.1109/CSDE59766.2023.10487762

[31] Noelia Vallez, Alberto Velasco-Mata, Juan Jose Corroto, and Oscar Deniz. 2019. Weapon Detection for Particular Scenarios Using Deep Learning. In *Pattern Recognition and Image Analysis*, Aythami Morales, Julian Fierrez, José Salvador Sánchez, and Bernardete Ribeiro (Eds.). Springer International Publishing, Cham, 371–382.

[32] Noelia Vallez, Alberto Velasco-Mata, Juan Jose Corroto, and Oscar Deniz. 2019. Weapon Detection for Particular Scenarios Using Deep Learning. In *Pattern Recognition and Image Analysis*, Aythami Morales, Julian Fierrez, José Salvador Sánchez, and Bernardete Ribeiro (Eds.). Vol. 11868. Springer International Publishing, Cham, 371–382. https://doi.org/10.1007/978-3-030-31321-0_32 Series Title: Lecture Notes in Computer Science.

[33] Alberto Velasco-Mata, Jesus Ruiz-Santaquiteria, Noelia Vallez, and Oscar Deniz. 2021. Using human pose information for handgun detection. *Neural Computing and Applications* 33, 24 (Dec. 2021), 17273–17286. https://doi.org/10.1007/s00521-021-06317-8

[34] Pavinder Yadav, Nidhi Gupta, and Pawan Kumar Sharma. 2024. Robust weapon detection in dark environments using Yolov7-DarkVision. *Digital Signal Processing* 145 (Feb. 2024), 104342. https://doi.org/10.1016/j.dsp.2023.104342