DLIB, YOLO, and Deep Learning-Based Motion and Anomaly Detection Systems for School Security Education

Belma Karanlık Tuna[1] and Egemen Aslan[2]

[1] İzmit Science and Art Center, Kocaeli, Türkiye
[2] İzmit Science and Art Center, Kocaeli, Türkiye
* Corresponding author: Belma Karanlık Tuna (belmakt@gmail.com)

Abstract: In recent years, the increase in unwanted incidents, especially at the high school level, has raised concerns about the adequacy of school security. This study presents the design and implementation of an artificial intelligence-based system developed to enhance safety in educational institutions. Developed using the Python programming language, the system utilizes OpenCV, YOLO, and the face-recognition libraries to identify students' faces and monitor movements within the school environment. Each student's facial data is registered in the system through five photos taken from different angles; subsequently, students' locations and activities are tracked in real-time via multiple camera streams. The system detects recognized faces and stores them in categorized data files for analysis purposes. Additionally, using a YOLO-based AI model, the system detects prohibited items such as phones, cigarettes, and knives within the school environment, and records these detections along with time and location information in dedicated files. The developed system offers a data-driven approach to addressing incidents at school and contributes to creating a safer and more controllable educational environment.

Keywords: School security, Face recognition, Object detection, Artificial intelligence, Image processing

Received: 14 November 2025 | Revised: 27 February 2026, 30 May 2026 | Accepted: 31 May 2026

Cite : Karanlık Tuna, B. & Aslan, E. (2026). DLIB, YOLO, and Deep Learning-Based Motion and Anomaly Detection Systems for School Security Education. *Journal of Artificial Intelligence and Human Sciences*, *3*(1), 16-25. Doi: 10.5281/zenodo.21462052

1. Introduction

Security is one of the fundamental requirements for maintaining social order and protecting individuals' rights within society (Zedner, 2009). Although technological developments have significantly improved modern security systems, crime and safety-related problems continue to remain major social concerns in many countries. Ensuring secure environments is especially important in places where large groups of people interact on a daily basis, such as schools, public institutions, and transportation areas.

Among these environments, schools have a particularly important role because children spend a substantial part of their developmental years in educational settings. For this reason, school safety is not only an institutional responsibility but also a social necessity. Unsafe school environments may negatively affect students both academically and psychologically. According to data published by the Turkish Statistical Institute (TÜİK, 2024), a considerable number of children in Türkiye have been involved in incidents requiring law enforcement intervention in recent years. Assault, theft, threats, and substance-related offenses constitute a significant proportion of these incidents. Previous studies have also shown that exposure to crime or involvement in criminal behavior at an early age may increase the likelihood of criminal activities later in life (Farrington, 1992; Moffitt, 1993). Therefore, improving safety in educational environments has become an increasingly important issue for both school administrators and policymakers.

One of the most widely used technologies for maintaining safety in schools is Closed-Circuit Television (CCTV) systems. Security cameras make it possible to record incidents, review events afterward, and identify individuals involved in problematic situations. Studies have shown that CCTV systems can contribute to reducing certain types of crimes, especially property-related offenses (Welsh & Farrington, 2009; Braga et al., 2014). However, their effectiveness in preventing violent incidents or improving students' perceptions of safety remains limited in some cases. Traditional CCTV systems mainly function as passive monitoring tools that record events after they occur rather than actively preventing them. As a result, there is a growing interest in intelligent surveillance systems capable of analyzing events in real time and detecting potential threats more effectively.

Recent developments in artificial intelligence, computer vision, and deep learning technologies have created new opportunities for improving surveillance systems. AI-supported applications are now being used in various fields such as facial recognition, object detection, public safety monitoring, and intelligent surveillance systems (Chui et al., 2020; Zhang et al., 2021). In particular, deep learning-based object detection models such as YOLO (You Only Look Once) provide fast and accurate detection performance suitable for real-time environments (Redmon et al., 2016; Bochkovskiy et al., 2020). Similarly, facial recognition technologies allow individuals to be identified and tracked across multiple camera streams. Integrating these technologies into school security systems may help not only with recording incidents but also with detecting suspicious situations and prohibited objects at an earlier stage. Although AI-supported surveillance technologies have attracted increasing attention in recent years, studies focusing specifically on integrated school-based surveillance systems are still limited (Chui et al., 2020; Zhang et al., 2021). Many existing systems focus only on single tasks such as facial recognition or object detection. In contrast, systems capable of combining real-time monitoring, behavioral tracking, and structured data recording within educational environments remain relatively uncommon. In addition, practical implementations of multi-camera AI-supported monitoring systems in schools have not been sufficiently examined in the literature.

In this context, the present study proposes an AI-supported school security platform that integrates facial recognition, prohibited object detection, and real-time movement monitoring within a multi-camera architecture. Developed using Python-based deep learning frameworks, the system utilizes YOLOv8, OpenCV, and face-recognition libraries to monitor school environments and record structured behavioral data. The proposed system aims to support proactive surveillance and contribute to creating safer and more controllable educational environments through intelligent monitoring technologies.

2. Methodology

In this study, an artificial intelligence-supported school security platform was developed to monitor school environments and detect potential security threats in real time. The system was implemented using the Python programming language due to its extensive support for artificial intelligence, computer vision, and image processing applications. Throughout the development process, several widely used libraries and frameworks, including OpenCV, YOLOv8, face_recognition, and Tkinter, were utilized (Redmon et al., 2016; Bochkovskiy et al., 2020; Zhang et al., 2021).

The developed platform consists of multiple interconnected modules responsible for camera management, facial recognition, object detection, data recording, and data visualization. The overall architecture of the proposed system is presented in Figure 1. The graphical user interface was developed using the Tkinter library to provide a simple and accessible interaction environment for users. Through the main interface, users are able to start the monitoring system, register new individuals, and access previously recorded data.

The monitoring process is managed through a multi-camera streaming module that enables multiple cameras to operate simultaneously within the system. Cameras can be dynamically added to or removed from the monitoring interface during operation. While configuring a camera, the user can define the monitored location and optionally activate notification settings for specific areas. This structure allows the system to perform location-based monitoring within the school environment.

Facial recognition operations were performed using the face_recognition library. During the registration process, five facial images of each individual were collected from different angles in order to improve recognition accuracy. These images were locally stored and converted into facial encodings. During live monitoring, detected faces within the camera streams were compared with the previously stored facial encodings. When a match was identified, the system displayed the individual’s identity on the screen and simultaneously recorded the related location and timestamp information.

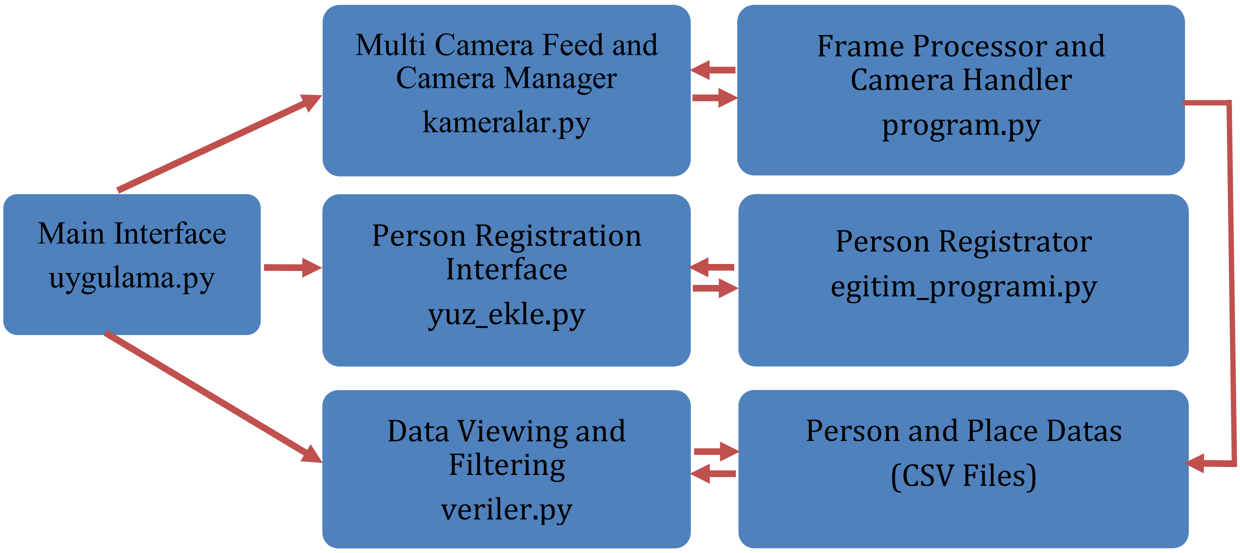

Figure 1. System Architecture

In addition to facial recognition, the platform performs real-time object detection using a YOLOv8-based deep learning model. YOLO (You Only Look Once) is a single-stage object detection architecture designed for fast and efficient real-time detection tasks (Redmon et al., 2016). The object detection component of the proposed system was trained to identify prohibited objects such as mobile phones, cigarettes, and knives within school environments. Publicly available datasets obtained from Kaggle and Roboflow Universe platforms were utilized during the training process (Bhatia, 2024; objectdetection-lryar, 2022). The datasets were reorganized, relabeled, and combined into a unified dataset compatible with the objectives of the study. The final dataset contained 9,191 training samples, including images of mobile phones, knives, and cigarettes.

The YOLOv8 model was trained for 108 epochs and evaluated using common object detection performance metrics, including mean Average Precision (mAP), precision, and recall values. The obtained results demonstrated strong detection performance, particularly for mobile phones and knives. During system operation, detected objects were recorded together with their timestamps and monitored locations in dedicated CSV files for later analysis.

To improve computational efficiency and enable real-time performance, GPU acceleration technologies developed by NVIDIA, namely CUDA and cuDNN, were integrated into the system. These technologies allowed computationally intensive image processing and deep learning operations to be executed on the GPU rather than the CPU, thereby significantly reducing processing latency and improving overall system performance.

The developed platform also includes a structured data recording and visualization infrastructure. Monitoring data generated by the system are automatically stored in CSV format under different categories, including locations, students, and potential violation elements. These records include movement histories, interaction durations, location usage information, and detected prohibited objects. In addition, a dedicated data visualization interface enables users to search, filter, and display recorded data in table format for further examination.

Additional implementation details, graphical user interfaces, data structures, and example outputs are presented in the Supplementary Material (Supplementary File A). Overall, the methodology of this study combines facial recognition, deep learning-based object detection, multi-camera monitoring, and structured behavioral data recording within a single integrated school security platform. The proposed framework aims to support proactive monitoring processes and contribute to the development of safer educational environments through AI-assisted surveillance technologies.

3. Findings

The developed AI-supported school security platform was evaluated under real-time monitoring conditions using multiple camera streams. The obtained findings indicate that the system successfully performed facial recognition, object detection, behavioral monitoring, and structured data recording within the school environment.

3.1. Facial Recognition Performance

The facial recognition module successfully identified registered individuals appearing in live camera streams. During testing, students enrolled in the system using facial images captured from different angles were correctly recognized and labeled in real time. In addition to identity recognition, the system successfully recorded location and timestamp information associated with each detection.

The generated records enabled the monitoring of student movements across different locations within the school environment. The collected data demonstrated that the system could effectively create individual movement histories and support retrospective investigation of events when necessary.

The recognition process generally performed reliably under normal lighting conditions. However, decreases in recognition accuracy were observed when faces were partially occluded or when lighting conditions were insufficient. Despite these limitations, the overall performance indicated that the facial recognition module was suitable for practical school monitoring applications.

3.2. Object Detection Results

The object detection module was evaluated using a YOLOv8-based deep learning model trained to detect mobile phones, knives, and cigarettes. The quantitative performance results demonstrated that the model achieved an mAP50 score of 84% and an mAP50-95 score of 59% on the test dataset. Furthermore, the mean precision and recall values were calculated as 0.86 and 0.75, respectively.

Performance varied among object classes. Mobile phones and knives achieved higher detection accuracy, whereas cigarette detection produced lower performance values. This difference can largely be attributed to the relatively limited number of cigarette images available during the training process.

As shown in Figure 2, the confusion matrix demonstrates strong classification performance for mobile phones and knives. However, some background regions were occasionally misclassified as cigarettes, leading to false-positive detections. This issue was partially mitigated by adjusting the confidence threshold used during deployment.

The F1-Confidence Curve presented in Figure 3 further illustrates the relationship between confidence thresholds and detection performance. The curve indicates that the selected threshold values provide a balanced trade-off between precision and recall, allowing the system to operate effectively in real-time surveillance scenarios.

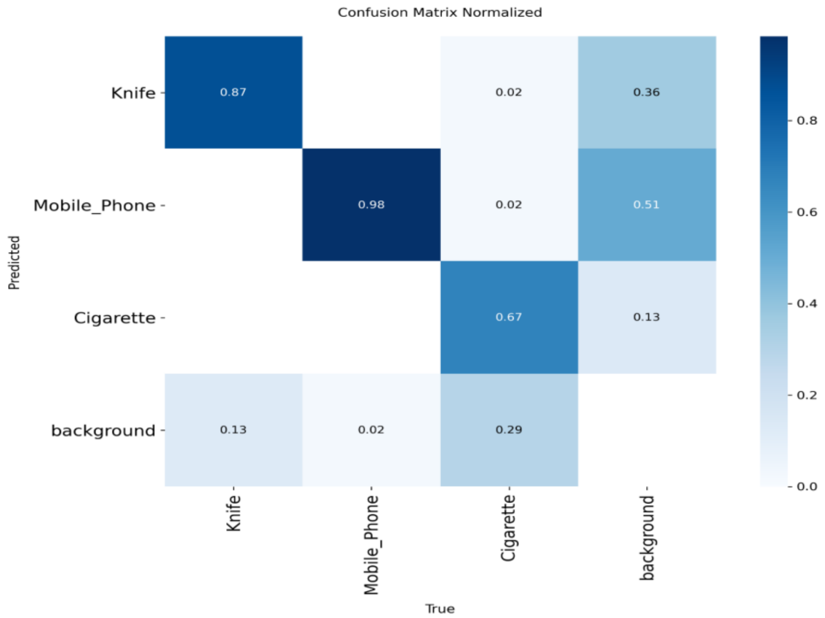

Figure 2. Confusion Matrix of the YOLOv8 Detection Model

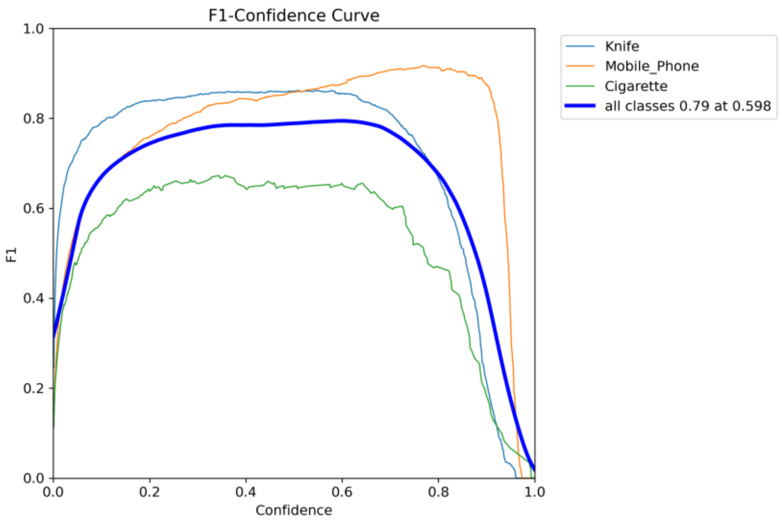

Figure 3. F1-Confidence Curve of the YOLOv8 Detection Model

Overall, the obtained results indicate that the object detection module is capable of reliably identifying prohibited objects within school environments and generating meaningful security-related alerts.

3.3. Behavioral Monitoring and Data Recording

One of the major outcomes of the developed platform is its ability to generate structured behavioral data. During operation, the system automatically recorded movement histories, location information, interaction durations, and prohibited object detections in CSV format.

The movement records stored in the *hareket.csv* files enabled chronological tracking of student activities throughout the day. Similarly, the *ogrenci.csv* and *mekan.csv* files provided detailed information regarding the duration of students' presence in specific locations.

Examples of these records are presented in Figures 4–6. The generated data demonstrate that the platform is capable of transforming raw surveillance information into structured and searchable datasets that can support administrative decision-making and incident investigations.

```
Öğrenci,Sınıfı,Numarası,Zaman
ogrenci1,sinif,numara,19:57:39
ogrenci2,sinif,numara,19:57:43
ogrenci1,sinif,numara,19:58:39
ogrenci2,sinif,numara,19:58:43
ogrenci1,sinif,numara,19:59:39
```

Figure 4. Sample Movement Records

```
Öğrenci,Sınıfı,Numarası,Süre (dk)
ogrenci1,sinif,numara,19
ogrenci2,sinif,numara,9
```

Figure 5. Sample Location Duration Records

```
Öğrenciler
  ogrenci1
    2025-01-04
    2025-01-05
    2025-01-06
      cevre.csv
      hareket.csv
      mekan.csv
  ogrenci2
    2025-01-04
    2025-01-05
    2025-01-06
      cevre.csv
      hareket.csv
      mekan.csv
```

```
Öğrenci,Sınıfı,Numarası,Süre (dk)
ogrenci1,sınıf,numara,19
ogrenci2,sınıf,numara,12
```

cevre.csv

```
Mekan,Sınıfı,Numarası,Zaman
Kantin,sınıf,numara,19:57:39
Kantin,sınıf,numara,19:58:39
Kantin,sınıf,numara,19:59:39
Bahçe,sınıf,numara,20:00:45
Bahçe,sınıf,numara,20:01:45
```

hareket.csv

```
Mekan,Sınıfı,Numarası,Süre (dk)
Kantin,sınıf,numara,19
Bahçe,sınıf,numara,13
```

mekan.csv

Figure 6. Student-Based Behavioral Monitoring Outputs

In addition, interaction records stored in the *cevre.csv* files enabled the identification of individuals who spent time together during the monitoring period. Such information may provide valuable support during investigations of disciplinary incidents or security-related events.

3.4. Detection of Potential Violation Elements

The platform successfully recorded detections of prohibited objects, including mobile phones, knives, and cigarettes. Each detection was stored together with the corresponding date, time, and monitored location.
Figure 7 presents a sample of the generated violation records. Unlike standard movement records, violation-related detections were recorded at a higher frequency in order to preserve detailed information regarding potentially critical incidents.

```
Tarih,Unsurlar,Mekan,Zaman
2025-01-06,Telefon,Kantin,19:45:55
2025-01-06,Telefon,Kantin,19:45:55
2025-01-06,Telefon,Kantin,19:45:56
2025-01-06,Telefon,Kantin,19:45:56
2025-01-06,Telefon,Kantin,19:45:57
2025-01-06,Telefon,Kantin,19:45:57
2025-01-06,Telefon,Kantin,19:45:58
2025-01-06,Bıçak,Bahçe,19:49:55
2025-01-06,Bıçak,Bahçe,19:49:55
2025-01-06,Bıçak,Bahçe,19:49:56
2025-01-06,Sigara,Bahçe,20:49:13
2025-01-06,Sigara,Bahçe,20:49:13
```

Figure 7. Sample Violation Detection Records

The collected records allow administrators to determine when and where a prohibited object was detected and to identify individuals present in the monitored area during the event. This capability supports evidence-based decision-making and enhances the practical usefulness of the platform.

3.5. Overall System Performance

The developed platform successfully integrated facial recognition, object detection, multi-camera monitoring, and structured behavioral data recording within a single framework. During testing, the system maintained stable operation while simultaneously processing multiple camera streams.

The integration of CUDA and cuDNN-based GPU acceleration significantly improved processing efficiency and reduced latency during image analysis operations. As a result, facial recognition and object detection tasks could be executed concurrently without substantial performance degradation.

Overall, the findings demonstrate that the proposed system can effectively support proactive school security management by combining real-time monitoring, intelligent object detection, and structured behavioral analysis within a unified AI-supported surveillance platform.

4. Discussion

The findings of this study demonstrate that artificial intelligence-based surveillance technologies can provide valuable support for improving school security through real-time monitoring, object detection, and behavioral analysis. The developed platform successfully integrated facial recognition, prohibited object detection, and structured data recording within a unified framework, thereby extending the functionality of conventional surveillance systems.

One of the most important findings is that the proposed system moves beyond the passive nature of traditional CCTV systems. Previous studies have reported that conventional surveillance cameras are generally effective for post-incident investigation but have limited capability in preventing incidents before they occur (Welsh & Farrington, 2009; Braga et al., 2014). In contrast, the developed platform continuously analyzes camera streams and generates real-time information regarding detected individuals and prohibited objects. This capability supports a more proactive security approach and may improve the ability of school administrators to respond to potentially risky situations.

The object detection results obtained in this study are consistent with previous research highlighting the effectiveness of YOLO-based architectures for real-time surveillance applications (Redmon et al., 2016; Bochkovskiy et al., 2020). The achieved mAP50 value of 84% demonstrates that the system can reliably identify prohibited objects under realistic operating conditions. Similar to previous deep learning studies, detection performance varied among object classes, with cigarette detection producing lower accuracy compared to mobile phones and knives. This limitation is likely associated with class imbalance within the training dataset and the relatively small number of cigarette images available during training.

The facial recognition component also produced promising results. The system was capable of identifying registered individuals across multiple camera streams and generating detailed movement histories. These findings support previous studies suggesting that AI-supported facial recognition technologies can significantly enhance situational awareness in surveillance environments (Chui et al., 2020; Zhang et al., 2021). The ability to automatically record location and interaction information may provide additional benefits for investigating incidents and understanding behavioral patterns within educational settings.

Another important contribution of the study is the structured recording of behavioral data. Unlike many surveillance systems that primarily store video recordings, the proposed platform converts surveillance outputs into searchable and analyzable datasets. This functionality enables administrators to review movement histories, interaction durations, and object detections without manually examining large volumes of video footage. Such capabilities may support evidence-based decision-making processes and improve operational efficiency.

Despite these promising results, several limitations should be acknowledged. First, the system was evaluated using a limited number of object categories, namely mobile phones, knives, and cigarettes. Future studies may extend the model by including additional categories such as firearms, dangerous substances, or other security-related objects. Second, environmental factors such as poor lighting conditions and facial occlusions negatively affected facial recognition performance. Improvements in camera quality, image preprocessing techniques, and training data diversity may help mitigate these limitations.

In addition to technical considerations, the use of AI-supported surveillance systems in educational environments raises important ethical and privacy-related concerns. Facial recognition and behavioral monitoring technologies involve the collection and processing of personal data, requiring compliance with relevant data protection regulations and ethical standards. Therefore, future implementations should balance security objectives with privacy protection and ensure transparent governance mechanisms regarding data collection, storage, and access.

Overall, the findings suggest that AI-supported surveillance systems have considerable potential to enhance school security by enabling proactive monitoring, rapid threat detection, and structured behavioral analysis. As artificial intelligence technologies continue to evolve, their integration into educational security infrastructures may provide increasingly effective and data-driven approaches for maintaining safer learning environments.

5. Conclusion

This study presented the design and implementation of an artificial intelligence-supported school security platform integrating facial recognition, prohibited object detection, and real-time movement monitoring within a multi-camera surveillance architecture. The proposed system was developed using Python-based computer vision and deep learning technologies, including OpenCV, YOLOv8, and face_recognition libraries.

The findings demonstrated that the developed platform successfully performed facial recognition, object detection, behavioral monitoring, and structured data recording in real-time monitoring environments. The facial recognition module enabled the identification and tracking of registered individuals, while the YOLOv8-based object detection model effectively detected prohibited objects such as mobile phones, knives, and cigarettes. In addition, the system automatically generated structured datasets containing temporal, spatial, and behavioral information that could support security management and incident investigations.

The results indicate that integrating artificial intelligence technologies into school surveillance systems can provide advantages beyond those offered by traditional CCTV infrastructures. Rather than functioning solely as passive recording tools, AI-supported systems can contribute to proactive monitoring, rapid threat identification, and data-driven decision-making processes. These capabilities may help educational institutions respond more effectively to potential security concerns while

improving overall situational awareness.
Despite the promising results, several limitations should be acknowledged. The current implementation focuses on a limited number of object categories and was evaluated under specific environmental conditions. Future studies may improve system performance by increasing dataset diversity, incorporating additional object classes, and integrating advanced anomaly detection and behavior analysis algorithms.
In conclusion, the proposed platform demonstrates the practical applicability of artificial intelligence technologies for school security applications. By combining facial recognition, object detection, and behavioral monitoring within a unified framework, the system provides a foundation for the development of safer, more responsive, and data-driven educational environments.

Acknowledgments

This study was presented as an oral presentation at the International Science and Art Student Congress organized by Muğla BİLSEM on June 11–13, 2025

Authors contributions

Belma Karanlık Tuna and Egemen Aslan contributed equally to this work.

Statements and Declarations

Ethical Considerations

Not applicable.

Consent to Participate

Informed consent was obtained from all participants before data collection.

Consent for Publication

All authors consent to the publication of this manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding Statement

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Data availability

Not applicable.

References

Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.

Braga, A. A., Papachristos, A. V., & Hureau, D. M. (2014). The effects of hot spots policing on crime: An updated systematic review and meta-analysis. *Justice Quarterly, 31*(4), 633–663. https://doi.org/10.1080/07418825.2012.673632

Chui, K. T., Lytras, M. D., & Visvizi, A. (2020). AI and smart cities: An empirical investigation of security surveillance applications. *Sustainable Cities and Society, 62*, 102389. https://doi.org/10.1016/j.scs.2020.102389

Farrington, D. P. (1992). Criminal career research in the United Kingdom. *British Journal of Criminology, 32*(4), 521–536.

Ferguson, A. G. (2017). The rise of big data policing: Surveillance, race, and the future of law enforcement. *NYU Press*.

Jocher, G., et al. (2023). YOLOv5 by Ultralytics. *GitHub repository*.

Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal, 79*(1), 1–14.

https://doi.org/10.1007/s10708-013-9516-8

Lum, K., & Isaac, W. (2016). To predict and serve? Significance, 13(5), 14–19. https://doi.org/10.1111/j.1740-9713.2016.00960.x

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100*(4), 674–701. https://doi.org/10.1037/0033-295X.100.4.674

Mohler, G., Short, M. B., Brantingham, P. J., Schoenberg, F. P., & Tita, G. E. (2015). Prediction and prevention of crime at hot spots. *Nature Human Behaviour, 1*(1), 1–7. https://doi.org/10.1038/s41562-015-0022-0

Perry, W. L., McInnis, B., Price, C. C., Smith, S. C., & Hollywood, J. S. (2013). *Predictive policing: The role of crime forecasting in law enforcement operations*. RAND Corporation.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788. https://doi.org/10.1109/CVPR.2016.91

Ristea, A., Leitner, M., & Langford, C. (2018). Predictive policing in urban areas: A review. *Computers, Environment and Urban Systems, 70*, 1–10. https://doi.org/10.1016/j.compenvurbsys.2018.02.003

Turkish Statistical Institute. (2024). *A snapshot on statistics on children in Türkiye 2024*. Ankara: TÜİK.

Welsh, B. C., & Farrington, D. P. (2009). Public area CCTV and crime prevention: An updated systematic review and meta-analysis. *Justice Quarterly, 26*(4), 716–745. https://doi.org/10.1080/07418820802506206

Zedner, L. (2009). *Security*. Routledge.

Zhang, Z., Qiu, S., & Li, X. (2021). Deep learning-based video surveillance for public safety: A review. *IEEE Access, 9*, 151730–151744. https://doi.org/10.1109/access.2021.3126789