

Investigating Enhanced YOLOv8 Model Applications for Large-Scale Security Risk Management and Drone-Based Low-Altitude Law Enforcement

Guancheng Chen^{1,*,#}, Wenzhuo Du^{2,#}, Tong Xu^{3,#}, Siyuan Wang^{3,#},
Xiang Qi^{1,#}, Yuxin Wang^{1,#}

¹School of Criminal Investigation, People's Public Security University of China, Beijing, China, 100038

²School of Information and Network Security, People's Public Security University of China, Beijing, China, 100038

³School of Public Security, People's Public Security University of China, Beijing, China, 100038

*Corresponding author: 13173211210@163.com

#These authors contributed equally.

Abstract. Urban expansion and complex environments are escalating security challenges. This research presents a drone-based risk assessment approach, significantly enhanced by the YOLOv8 algorithm. Through continuous iterations, the algorithm's performance has improved approximately fourfold, greatly boosting detection accuracy for low-altitude law enforcement and overcoming the constraints of existing surveillance technologies. Furthermore, the algorithm was applied to five major marathon events, conducting security risk assessments based on EWM-TOPSIS and FCM clustering. The final risk ranking of the events was Zhanjiang, Shanghai, Qingdao, Jiangsu, and Tianjin, with the Zhanjiang Marathon showing a people risk index of 72.12, indicating a need for strengthened security measures. This research applies artificial intelligence to public safety, delivering a scalable solution for a variety of security challenges.

Keywords: YOLOv8, Security control, Low-altitude law enforcement, Drone recognition.

1. Introduction

In recent years, China's society has been in a period of transition. With the continuous expansion of urban areas and the acceleration of urbanization, the urban three-dimensional spatial environment has become increasingly complex and variable^[1]. Various unstable factors and emergencies have put significant pressure on public security work. Moreover, the issue of insufficient grassroots police force is prominent, making traditional policing methods, such as helicopters and manpower, unable to meet the needs of modern police security^[2, 3]. As the demand for urban safety continues to grow, the application of drones in low-altitude law enforcement and security has become increasingly widespread^[4]. However, existing drone surveillance technology still has shortcomings in risk assessment^[5, 6]. Despite significant progress in deep learning technology in the field of target detection in recent years, research on safety risk assessment under drone low-altitude flight conditions, especially the identification of dynamic risk factors in complex environmental conditions, remains limited^[7, 8].

To address the above issues, this paper proposes a drone low-altitude law enforcement security risk assessment method based on an improved YOLOv8 algorithm. This method aims to significantly enhance the detection accuracy and speed of drones for small target risk factors in complex environments through deep customization and optimization of the YOLOv8 algorithm. By adding various attention mechanisms, updating the YOLOv8 network structure, and creating new network architectures and modules tailored to the characteristics of low-altitude law enforcement datasets, the method reduces computational weights. This not only improves the model's generalization ability in complex environments but also significantly reduces the consumption of computational resources. This paper also explores a real-time data-driven dynamic risk assessment framework that can

adaptively adjust risk assessment strategies based on environmental changes, enhancing the model's detection performance.

2. Construction of a large-scale security risk dataset

The experimental dataset was sourced from field reconnaissance labs and created using video data obtained via open-source intelligence tools. It comprises videos from five major security events covered by drones in 2024, including the Zhanjiang Marathon, Shanghai Marathon, Qingdao Marathon, Jiangsu Marathon, and Tianjin Marathon. Video clips from various critical security locations were selected to enrich the sample diversity for the experiments. The collection process involved analyzing low-altitude law enforcement videos returned by drones at different security sites, targeting the number of people and vehicles using the YOLOv8 algorithm. For each site, video images were selected to last about 10 seconds, with approximately 5 pictures chosen. The distribution of the dataset is outlined in Table 1.

Table 1: Data distribution of training set and testing set

Name	Training set		Testing set	
	Object	Image	Object	Image
Security	5162	2866	589	342

For the drone-captured video materials, this study employed advanced video processing techniques for decoding and frame extraction. Using the HMS Transcoder tool, videos were initially decoded, with a frame extracted every two frames to ensure a resolution of 1920x1080 pixels, capturing finer scene details. Redundant or unnecessary images were removed, maintaining 250 to 350 high-quality images per security scenario. The dataset, aiming for realism and broad coverage, includes images under partial occlusion, dynamic blur, and varying lighting conditions to mimic real surveillance challenges. The dataset comprises about 3100 carefully selected images and around 5800 precise labels for deep learning model training and validation. Image data was meticulously annotated with labeling software, outlining security-related targets without including irrelevant background elements, using the PASCAL VOC format and "security" as the uniform label name.

3. Construction and Optimization of the YOLOv8 Model

The experiments were conducted on a server featuring a 13th Gen Intel(R) Core(TM) i9-13900HX CPU at 2.2GHz, 32 GB of RAM, and an NVIDIA GeForce RTX 4060 GPU with 8 GB VRAM, running Windows 11 and leveraging CUDA 11.2 for GPU acceleration with Pytorch 1.8.0 as the deep learning framework. The training process spanned 200 epochs, initially freezing network layers for the first 80 epochs to stabilize foundational feature learning, with an initial learning rate of 0.001. For the subsequent 120 epochs, to refine adaptability and improve detection accuracy in complex scenarios, the learning rate was reduced to 0.0001, and all network layers were trained, using the Adam optimizer to ensure stability and speed in convergence.

3.1. Multi-Scale Dilated DilateFormer Architecture

In response to the demands of low-altitude police surveillance and large-scale security, we developed DilateFormer, a multi-scale dilated Transformer architecture optimized for drone-based visual recognition. Addressing the limitations of traditional Vision Transformers (ViT) in handling complex security scenes with high computational costs, DilateFormer incorporates a Multi-Scale Dilated Attention (MSDA) mechanism^[9]. This approach reduces computational redundancy and maintains essential spatial relationships by focusing on localized and sparse interactions between image blocks, improving the efficiency of drone surveillance systems. The result is a significant enhancement in real-time performance and accuracy for security tasks, leading to the upgraded YOLOv8 model, which is named YOLOv8_A.

3.2. Improvements to the SPD-Conv Module

Despite Convolutional Neural Networks (CNNs) excelling in tasks like image classification and object detection, they struggle with complex challenges due to the loss of detail from stride convolutions and pooling. We've developed a novel CNN module, SPD-Conv, to overcome this by using Spatial-to-Depth (SPD) layers and non-stride convolutions, preserving information while sampling feature maps. This enhancement is particularly beneficial for analyzing drone surveillance footage of small or low-resolution targets. Incorporating SPD-Conv into YOLOv8 and ResNet, we created the YOLOv8-SPD and ResNet-SPD models. Extensive testing on drone surveillance datasets revealed marked improvements in accuracy, especially for detecting small targets in security applications. The SPD-Conv module not only boosts existing CNN capabilities for drone surveillance but also offers a flexible solution for various deep-learning vision tasks, making real-time and precise security risk prevention achievable. The enhanced model is referred to as YOLOv8_B.

3.3. Switching to FasterNet as the Backbone Network

In drone policing and security, fast and accurate target identification is key. Traditional neural networks face real-time challenges, especially on resource-limited drones. We introduced PConv, an innovative local convolution operator that boosts spatial feature extraction by optimizing computation and reducing memory access. Our FasterNet neural network, built on PConv, speeds up image analysis on drones without sacrificing accuracy. For example, FasterNet-T0 runs up to 3.1 times faster than MobileViT-XXS on GPUs and CPUs and also improves accuracy. This emphasizes the need for FLOP efficiency over mere reduction, making FasterNet a rapid solution for drone vision tasks^[10]. The upgraded YOLOv8 model, YOLOv8_C, showcases this technology's potential to enhance drone surveillance efficiency.

3.4. Integrating the BiLevelRoutingAttention Dual-Level Routing Attention Mechanism

We developed a dynamic sparse attention mechanism for drone surveillance, balancing long-range dependency capture with lower computational costs. Named BiFormer, this visual transformer uses a two-level routing strategy to efficiently allocate computation and enhance content awareness, focusing on relevant tokens to reduce unnecessary processing. Ideal for dense prediction in drone video analysis, BiFormer outperforms existing methods in efficiency and performance. The improved YOLOv8 model, YOLOv8_D, demonstrates its effectiveness in various vision tasks.

3.5. Introducing the SEAttention Attention Mechanism

We explored CNN's feature extraction in drone surveillance, focusing on enhancing channel dependency representation. Introducing the "Squeeze-and-Excitation" (SE) block, our SENet architecture dynamically recalibrates channel features to improve network performance with minimal computational cost. Proven effective in drone policing and security tasks, SENet offers a new approach to modeling channel relationships, aiding in feature distinction for complex visual tasks. This innovation provides tools for network pruning and model compression. The improved YOLOv8 model is named YOLOv8_E.

3.6. Model Performance Comparison

Building on previous insights, we enhanced YOLOv8 by integrating various attention mechanisms, updating its network structure, and developing a new architecture tailored to the characteristics of low-altitude law enforcement drone datasets. This led to the creation of the YOLOv8_security model. Continuous iterations have resulted in performance evolution across different models, as outlined in Table 2.

Table 2: Comparing model efficacies

Model type	mAP/%	Params/M
YOLOv8	71.2	1.14
YOLOv8_A	80.2	3.67
YOLOv8_B	82.4	3.82
YOLOv8_C	81.8	4.13
YOLOv8_D	82.1	3.97
YOLOv8_E	81.3	3.16
YOLOv8_security	84.9	5.939

Through the integration of innovative mechanisms such as Multi-Scale Dilated Attention (MSDA), SPD-Conv modules, PConv operations, BiLevelRoutingAttention, and SEAttention, our YOLOv8 model variants have demonstrated significant improvements in mean Average Precision (mAP) scores, surpassing the original YOLOv8 model's performance. These enhancements have proven to be particularly effective in low-altitude police surveillance scenarios, enabling more accurate identification and classification of low-resolution or small-sized objects. The inclusion of mechanisms like MSDA and DilateFormer structures enhances the model's field of view and context-capturing capabilities, improving the identification of small targets and detailed features in complex monitoring scenes. Similarly, SPD-Conv's spatial to-depth conversion and the application of non-stride convolution layers have been crucial for preserving more detailed information, vital for target detection in complex environments. Furthermore, innovations such as PConv and FasterNet architecture, as well as BiLevelRoutingAttention and SEAttention mechanisms, have optimized spatial feature extraction, computation efficiency, and channel feature recalibration, strengthening inter-channel dependencies while reducing unnecessary computations and memory usage. These advancements not only validate the theoretical improvements but also showcase exceptional performance in real-world drone security applications, offering powerful tools for real-time monitoring and risk assessment, even on resource-constrained devices.

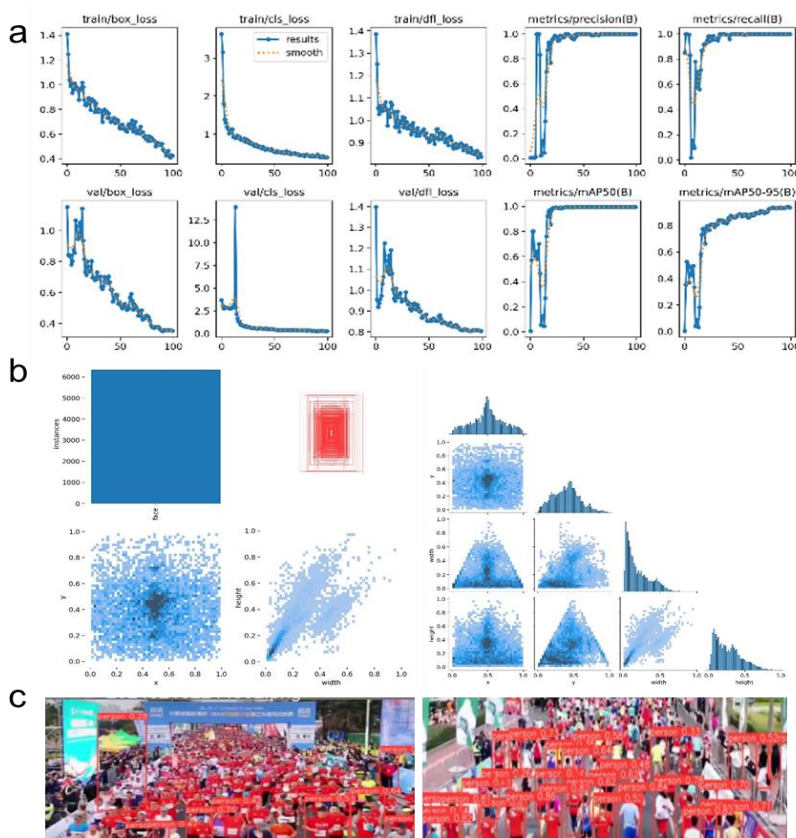


Figure 1: (a) Training and testing loss for YOLOv8_security (b) Training outcomes for YOLOv8_security (c) Model training iterations

As shown in Figure 1, compared to the original YOLOv8 model, the YOLOv8_security variant shows higher confidence and detection efficiency, particularly in complex security scenarios involving partial occlusions, dynamic blur, and targets moving in various directions. This improved model demonstrates superior adaptability and accuracy, accurately identifying and locating security risk elements in most situations, thus highlighting its powerful capabilities in drone surveillance applications. However, the detection process of YOLOv8_security is not flawless, as illustrated in Figure 1c, where errors stem from rapid target movement causing image blur, detection challenges due to other objects obstructing the dynamic scene, uneven lighting conditions, and difficulty in distinguishing pedestrians and vehicles from backgrounds with similar colors or textures. These issues point out the areas for future improvements in YOLOv8_security to enhance the accuracy and reliability of drone-based low-altitude law enforcement surveillance and large-scale security risk prevention tasks.

4. Drone-Based Security Risk Assessment Methods and Strategies

In this study, we initially utilized ArcGIS to visualize various risk indicators at different sites of five major marathon events in 2024, serving as typical security activities, as shown in Figure 2. Further, to assess the distinguishing effectiveness of the indicator system, we clustered the data using the FCM algorithm, identifying critical values through the coordinates of two of the five cluster centers, with threshold conditions presented in Table 3. We averaged the security risk levels of the EWM and TOPSIS models, designating the scoring outcomes as Mi1 for EWM, Mi2 for TOPSIS, and Mi3 for the combined EWM-TOPSIS model. The specific formulas for calculating the composite model are as follows:

$$Mi3 = \alpha Mi1 + \beta Mi2 \quad (1)$$

$$\lambda_1 = \frac{\mu_1 + \mu_2}{2} \quad (2)$$

$$\lambda_2 = \frac{\mu_2 + \mu_3}{2} \quad (3)$$

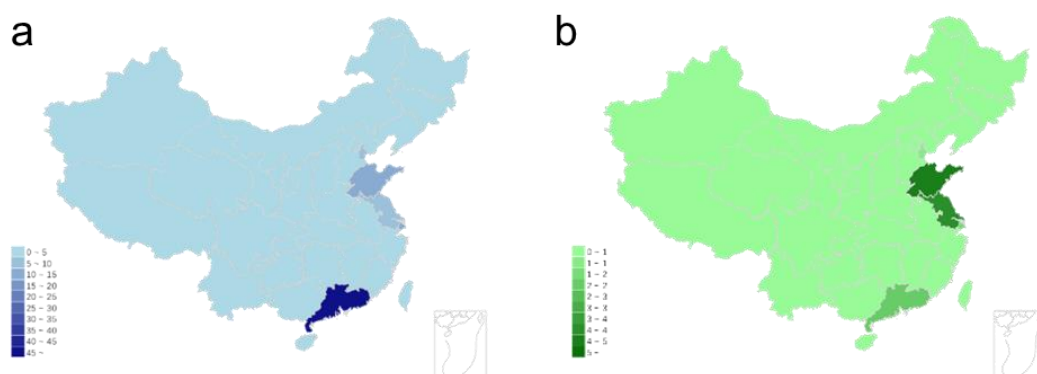


Figure 2: (a) Risk Indicators Based on Number of People (b) Risk Indicators Based on Number of Vehicles

Table 3: Security risk thresholds

Level	Start point	Key locations	Endpoint
High Risk	(15,26]	(22, 107]	(17,23]
Moderate Risk	(11,15]	(13, 22]	(12,17]
Medium Risk	(9,11]	(3,13]	(11,12]
Low Risk	(1,9]	(1, 3]	(4,11]
Minimal Risk	(0,1]	(0,1]	(0,4]

Table 4: Security risk ranking

Security Event	People risk index	Vehicle risk index	Risk ranking
Zhanjiang Marathon	72.12	4.21	1
Shanghai Marathon	14.19	3.52	2
Qingdao Marathon	11.92	1.92	3
Jiangsu Marathon	5.62	1.09	4
Tianjin Marathon	5.28	0.98	5

In the comprehensive assessment of safety risks for marathon events, as shown in Table 4, the data distinctly places the Zhanjiang Marathon at the highest risk, with a personnel risk index peaking at 72.12 and the highest vehicle risk index of 4.21, earning it the top spot in risk ranking. Although the Shanghai Marathon's personnel risk index is substantially lower than that of Zhanjiang, its vehicle risk index remains comparably high, thus securing the second rank. The Qingdao and Jiangsu Marathons exhibit more moderate risk indices, positioned at third and fourth, respectively. Despite the lack of a specific risk ranking for the Tianjin Marathon in the table, its lower risk indices suggest a relatively higher safety margin. This analysis underscores the imperative for event organizers to bolster safety measures for higher-risk events, including, but not limited to, augmenting security personnel, enhancing traffic management, and developing emergency response strategies to safeguard participants and spectators. Furthermore, the provided data is instrumental for participants when evaluating the personal risks of engagement in these sporting events.

To adequately prepare for any emergency scenarios, such as drone malfunctions or loss of control, establishing detailed contingency plans is essential. Ensuring the availability of vital rescue and incident response equipment, including drone capture nets and signal jammers, is a primary step. It is equally important to secure a sufficient supply of resources and equipment, like backup power, first-aid kits, and communication tools, to enable prompt response to any crisis. Additionally, forming an emergency response team and regularly conducting drills and simulations significantly bolsters the overall responsiveness and flexibility of the security risk control system in the face of significant incidents.

In the event of a substantial risk during security operations, immediate preventive actions are critical. Deploying police personnel for onsite management and using drones for broadcasting announcements are effective measures. Complementing these efforts, the integration of unmanned aerial vehicles with terrestrial-monitoring systems forms a comprehensive, dynamic detection network. This network is crucial for capturing real-time conditions and facilitating swift emergency responses. Assigning trained professionals to analyze live security footage from drones allows for the early identification of zones at risk for unforeseen events or threats. Dynamic risk assessments, continually refined based on real-time data and environmental shifts, enhance the precision of threat detection. Following security events, a thorough review of the effectiveness and adherence to safety protocols is conducted using preserved surveillance footage, drone recordings, and analytical reports. This evaluative process is key to gauging the performance of safety measures and identifying areas for improvement. Regular vulnerability assessments and penetration tests of IT systems are fundamental in uncovering and addressing security flaws, ensuring continuous improvement in security practices.

5. Conclusions

In summary, this study successfully constructed a comprehensive, large-scale security risk dataset through meticulous video data collection and processing from major marathon events, employing advanced algorithms and video processing techniques to enhance sample diversity and realism. The development and optimization of the YOLOv8 model variants, including the innovative YOLOv8_security model, demonstrated significant improvements in detection accuracy and efficiency, particularly in challenging low-altitude police surveillance scenarios. The integration of

cutting-edge mechanisms such as Multi-Scale Dilated Attention, SPD-Conv modules, and the attention mechanism, among others, has markedly advanced the capabilities of drone-based surveillance systems. This research not only provides valuable insights into the construction and optimization of high-performance deep learning models for security applications but also underscores the potential of drone technology in enhancing public safety and risk assessment methodologies. The findings offer a promising direction for future research in the field, aiming at refining detection processes and improving the reliability and effectiveness of drone-based surveillance for large-scale security risk prevention.

References

- [1] ZHANG K, WEI W, YIN L, et al. Spatial-Temporal Evolution Characteristics and Mechanism Analysis of Urban Space in China's Three-River-Source Region: A Land Classification Governance Framework Based on "Three Zone Space". Land [Z]. 2023: 12.
- [2] EVANS P. The police revolution[M]. Taylor & Francis, 2023.
- [3] CAWTHRAY T, BULL M. The Spatial Dimension of Police Legitimacy: An Exploration of Two Pacific Island States[J]. Asian Journal of Criminology, 2023,18(4): 459-478.
- [4] XU C, LIAO X, TAN J, et al. Recent research progress of unmanned aerial vehicle regulation policies and technologies in urban low altitude[J]. IEEE Access, 2020,8: 74175-74194.
- [5] YAACOUB J, NOURA H, SALMAN O, et al. Security analysis of drones systems: Attacks, limitations, and recommendations[J]. Internet of Things, 2020,11: 100218.
- [6] SONG Z, YE J. Selection of Development Mode of Police Unmanned Aerial Vehicle from the Perspective of Intelligent Policing[J]. Journal of Artificial Intelligence Practice, 2023.
- [7] WU X, SAHOO D, HOI S C. Recent advances in deep learning for object detection[J]. Neurocomputing, 2020,396: 39-64.
- [8] PAL S K, PRAMANIK A, MAITI J, et al. Deep learning in multi-object detection and tracking: state of the art[J]. Applied Intelligence, 2021,51: 6400-6429.
- [9] LIU X, XIE J, XU X, et al. Multi-source deep recommendation model of fusion dual attention mechanism[J]. Journal of Shaanxi Normal University (Natural Science Edition), 2023,51(05): 49-59.
- [10] WU K, XU Z, SHAN H. Improved Rapid Detection Method for Self-Exploding Defects in Glass Insulators Based on FasterNet and YOLOv5[J]. High Voltage Engineering, 2023: 1-11.