Advances in Computer, Signals and Systems (2024)
Clausius Scientific Press, Canada

DOI: 10.23977/acss.2024.080318
ISSN 2371-8838 Vol. 8 Num. 3

Research on Employee Abnormal Behavior Detection Algorithm Based on Improved SSD

Rong Zhou[1,a], Haibo Peng[2,b,*], Siyao Liu[3,c]

[1]*Yunnan KISC Electronic Information Technology Co., Ltd., Kunming, Yunnan, 650300, China*
[2]*Yunnan Open University, Kunming, Yunnan, 650300, China*
[3]*Kunming University of Science and Technology Oxbridge College, Kunming, Yunnan, 650300, China*
[a]*8573746@qq.com*, [b]*79517483@qq.com*, [c]*pp.colly@163.com*
[*]*Corresponding author*

Keywords: Improving SSD Algorithm, Employee Abnormal Behavior Detection, Real Time Monitoring, Energy Efficiency Ratio

Abstract: Detection of employee abnormal behavior is a hot topic in the field of video surveillance. In order to improve the accuracy and real-time performance of detection, this paper proposes an employee abnormal behavior detection algorithm based on SSD and improves the SSD algorithm. The improved SSD algorithm can effectively detect abnormal behaviors such as long-term immobility, sudden increase in activity, abnormal aggregation, abnormal body language and abnormal work performance. This paper introduces the technical route of enhanced SSD algorithm, including data preprocessing, network structure improvement, feature extraction, multi-scale prediction, target detection head design, loss function definition, post-processing technology, model training strategy and so on. The introduction of advanced feature fusion technology and the optimization of network structure make the improved SSD algorithm improve in three key performance indexes: detection time, detection accuracy and energy efficiency. Experimental results show that the maximum detection time of the improved algorithm is only 900ms, and the detection accuracy is 77.5%-95.4%. With the improvement of the energy efficiency ratio from 2.5 to 4.5, the changes of these indicators are very important for real-time monitoring system, which can greatly shorten the response time and reduce energy consumption.

1. Introduction

In recent years, deep learning technology has made a revolutionary breakthrough in the field of machine vision, so SSD, with both high accuracy and high efficiency, has attracted wide attention in the field of pedestrian detection. At present, with the wide application of intelligent monitoring systems in industrial safety and public management, automatic detection of employee abnormal behavior becomes a problem that needs to be solved in time. However, existing algorithms still have limitations in real-time performance, accuracy, and detection capabilities for small and occluded targets.

In this study, an improved SSD algorithm is proposed, which significantly improves the detection accuracy and real-time performance by introducing feature fusion techniques and optimizing the network structure and loss function. The contributions of this study are that an improved SSD algorithm combining multi-scale features and attention mechanism is proposed to enhance the detection capability of targets of different sizes; the energy-efficiency ratio of the algorithm is optimized by model compression and acceleration techniques to make it more suitable for real-time surveillance scenarios; and extensive experiments on several standard datasets are conducted to verify the effectiveness of the proposed algorithm.

Firstly, this paper reviews the related work and summarizes the current research progress of employee abnormal behavior detection algorithms. Then the framework and key technologies of the algorithm are introduced. The next part shows the experimental results and comparative analysis of performance. Finally, the conclusion summarizes the main findings of this paper.

2. Related Work

With the development of deep learning and computer vision technology, researchers have proposed a variety of algorithms to improve the detection effect of abnormal behavior in video surveillance. In order to improve the detection effect of abnormal behavior of the crowd in video surveillance, Tie Fuzhen extracted the corner points as feature points in the video frame of the crowd in video surveillance by using the improved single Gaussian model, calculated the motion speed and direction of the feature points based on the improved optical flow method, extracted the effective feature points, and obtained the target image of the crowd's motion, so as to put forward the algorithm for detecting abnormal behavior of the crowd in video surveillance based on the improved optical flow method [1]. Aiming at You Only Look Once version 5 (YOLOv5) algorithm which can detect abnormal behaviors, but the parameter scale is large and the Graphics Processing Unit has a large amount of computation, Zhu Xianyuan proposed the YOLOv5 algorithm that integrated Mobile Network Version 3. This algorithm improves the efficiency of the algorithm by improving the network structure, and improves the ability to recognize occluded multi target identification through small target anchor boxes [2]. Li Zhuoqing designed a deep learning-based embedded real-time abnormal behavior monitoring system in order to monitor the abnormal behaviors of people in a place and automatically send out alarm messages. The system collects image information by camera, then detects the key point coordinates of human body in the embedded device with posture detection network, and classifies the abnormal behaviors by using deep forest algorithm on the key point coordinates [3]. Zhao Lian proposed an abnormal behavior detection framework for safety production issues in current industrial scenarios, mainly targeting two special situations: workers sleeping and falling[4]. Zhu Qiang proposed a two-stage detection algorithm based on Full Convolutional One Stage(FCOS) to achieve abnormal behavior detection in smart construction sites. This algorithm mainly consists of two cascaded networks. Firstly, FCOS is used to identify and locate operators and abnormal behavior markers, and then multi-layer perceptrons are used to detect and classify abnormal behavior [5].

In addition, Liu D proposed a new algorithm combined with hierarchical structure to enhance features and improve the ability of classification and detection [6]. In order to solve the problem of garbage detection, Meng J proposed a mobile network -SSD model based on feature pyramid network. Compared with SSD model, this model has the advantages of reducing parameters, reducing internal space and improving the performance of small objects [7]. Zhao M proposed an enhanced feature pyramid scheme to improve the performance of SSD. In addition, he also proposed a cascade detection scheme to enhance the SSD positioning ability [8]. Chen W added an initial module to the extra layer of SSD, and proposed a detection algorithm based on SSD to improve the accuracy, especially for small vehicles [9]. In this paper, Wang Y combined six multi-scale feature maps extracted from the original SSD algorithm and proposed a new algorithm [10]. Although these researches have made remarkable progress in improving the accuracy and efficiency of abnormal behavior detection, the existing algorithms still have room for improvement in dealing with complex scenes, multi-target recognition and real-time performance. Based on the advantages of SSD, advanced feature fusion technology and optimized network structure, this paper proposes a new employee abnormal behavior detection algorithm.

3. Method

3.1 Algorithm Framework

In the study of employee abnormal behavior detection algorithm based on improved SSD, the design of the algorithmic framework includes key steps such as data preprocessing, improved network input layer, feature extraction, multi-scale prediction, target detection header, loss function, post-processing, model training, model evaluation and real-time performance optimization [11-12]. First, the video data is subjected to frame extraction, scaling and data enhancement to fit the network input size and improve model generalization. Next, a pre-trained deep learning model is used as a backbone network to extract image features, and the network structure is adapted to better capture scene features. The multi-scale prediction capability of SSD allows for target detection at different layers, with a particular focus on prediction modules for human-scale targets [13-14]. The target detection head predicts target classes and bounding box offsets on each feature map through a series of convolutional layers, and additional prediction heads may be added specifically to capture human gesture or behavioral patterns. Each scale l uses a convolutional layer C to predict the category c and bounding box b with the following expression:

$$(c^l, b^l) = C(X^l) \tag{1}$$

X is image features.

A multi-task loss function is used to simultaneously optimize the classification and localization tasks, and techniques such as focus loss may be introduced to solve the category imbalance problem:

$$L = L_c + L_b \tag{2}$$

L is the multitask loss function, L_c is the classification loss, and L_b is the bounding box regression loss.

Non-great value suppression is applied in the post-processing step to remove overlapping detection frames and low-confidence predictions, as well as possible additional behavioral analysis modules to identify coherent behavioral patterns. Among them, the prediction results are optimized through non maximum suppression, and the formula is:

$$R = NMSP(c_l, b_l) \tag{3}$$

$NMSP$ refers to the non maximum suppression operation used to optimize the prediction results, and R is the optimized result.

The model training uses annotated employee behavior video datasets and adopts an end-to-end training strategy. The training process can be expressed as an optimization problem, with the goal of finding the parameter set that minimizes the loss function:

$$\theta^* = \mathrm{argmin}_\theta \sum_{(x,y)\in Dtrain} L(f_\theta(x), y) \tag{4}$$

θ^* is the optimized model parameter, argmin_θ refers to finding the value of θ that minimizes the subsequent functions, $Dtrain$ is the training dataset, which includes multiple pairs of input and output samples (x,y), $L(f_\theta(x), y)$ is the loss function, which measures the difference between the model's predicted value $f_\theta(x)$ and the true value y.

Finally, in order to meet the needs of real-time behavior detection, the model is compressed and accelerated, and model pruning techniques are used to reduce the computational burden of the model.

3.2 Data Collection and Preprocessing

In the study of employee abnormal behavior detection algorithm based on improved SSD, it is first necessary to determine the data requirements, which include employee daily work videos and abnormal behavior videos, and then collect these video data through the camera in different work environments, and annotate them in order to identify the segments of normal and abnormal behaviors, and mark the regions of interest in the video [15-16]. The data related to data collection is shown in Table 1:

Table 1: Data Collection

Serial No.	**Data Type**	**Video Source**	**Behavior Category**	**Annotation Type**	**Frame Count**	**Resolution**	**Preprocessing Steps**	**Dataset Partition**	**Remarks**
1	Normal Work	Camera A	Normal	Bounding Box	3000	1920x1080	Scaling to 600x340	Training Set	None
2	Abnormal Act	Camera B	Abnormal	Bounding Box	1500	1280x720	Scaling, Enhancement	Training Set	Contains fighting behavior
3	Normal Work	Camera C	Normal	Bounding Box	2500	1920x1080	Scaling, Augmentation	Validation Set	None
4	Abnormal Act	Camera D	Abnormal	Bounding Box	1200	1280x720	Scaling, Augmentation, Denoising	Test Set	Contains falling behavior

Table 1 provides researchers with a clear data management framework to ensure consistency and completeness of data preprocessing, and provides detailed data records for model training and evaluation.

As the video is segmented, key frames are extracted and the image is resized to fit the model inputs, while the data diversity is enhanced by flipping, rotating, scaling, and color adjustments to improve the generalization ability of the model. The annotation process involves defining the annotation format and using annotation tools to allow experts to efficiently annotate the data and ensure that the annotations are converted to the format required for model training. In addition, the dataset needs to be divided into training, validation, and test sets, and the data loader needs to be designed to efficiently read and process the data during the training process.

In data normalization process, pixel values of image data need to be normalized for model training. Advanced data enhancement techniques such as simulating anomalous behavior and adding noise can further improve the recognition ability of the model. Finally, the preprocessed data is stored in a format suitable for reading by the deep learning framework and data validation is performed to ensure the data quality to construct a dataset suitable for the improved SSD algorithm for employee abnormal behavior detection.

3.3 Network Architecture Improvements

The improvement measures carried out on the SSD network structure involve several aspects [17-18]. Firstly, the backbone network suitable for video data is selected, and the backbone network is adjusted according to the characteristics of abnormal behavior detection, including increasing or decreasing the convolutional layers, and adjusting the depth and width of the network to better extract spatio-temporal features. Second, the feature fusion method is improved by using null convolution instead of the traditional convolution operation to extract multi-scale features and maintain the resolution, which helps to detect anomalous behaviors of different sizes. Meanwhile, the processing of temporal dimension is introduced in SSD to capture dynamic behavioral patterns in videos through temporal convolutional networks [19-20]. In addition, the introduction of an attention mechanism enhances the model's ability to identify the moment of abnormal behavior, and the optimization of the bounding box prediction module includes an improved bounding box regression algorithm or the addition of an extra prediction layer to improve localization accuracy. This study improves the category prediction module, using more complex classifiers or introducing additional feature representations to improve classification accuracy. In this study, we design or select a loss function that is more suitable for the task of abnormal behavior detection, so as to improve the ability of the model to recognize abnormal behavior. To meet the real-time requirements, the model is compressed and accelerated, including techniques such as weight pruning, quantization, and knowledge distillation to reduce the computational burden of the model. This study implements an end-to-end training strategy that goes directly from raw video data to the detection results of abnormal behavior without additional intermediate steps. In this study, specific data enhancement techniques such as simulating abnormal behavior and adding noise are used in the training process to improve the generalization ability and robustness of the model. As shown in Figure 1, it is an improved SSD network diagram:

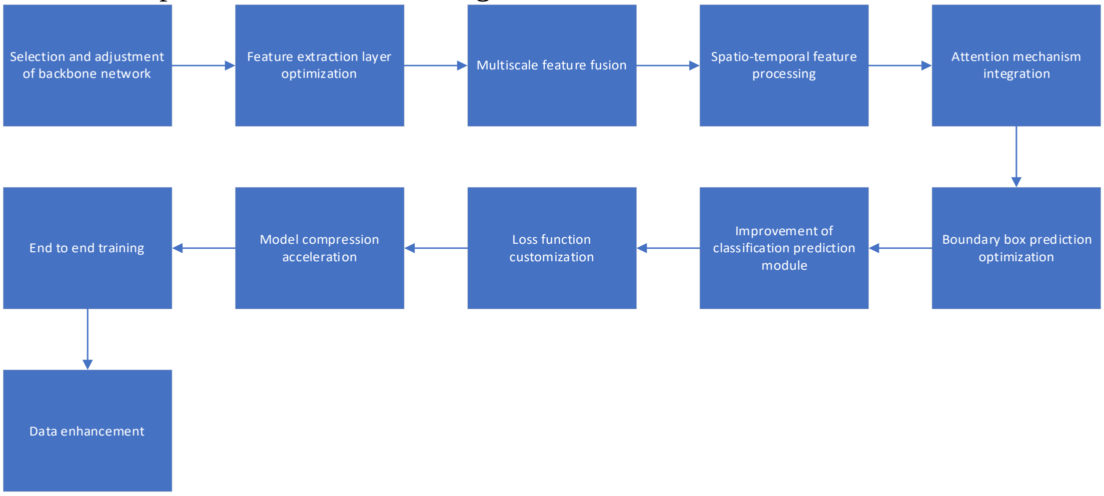

Figure 1: SSD network

4. Results and Discussion

This study aims to comprehensively evaluate the performance of the SSD algorithms before and after the improvement in the task of employee abnormal behavior detection through a series of comparative experiments. This study pays special attention to three key performance metrics: detection time, detection accuracy, and energy-efficiency ratio, with a view to revealing the specific impact of the algorithmic improvement on the efficiency and accuracy of the real-world application. The original SSD model and the improved SSD model were tested at the beginning of the experiment using the 22 sets of data collected above, the time required by each model to process a single video frame was measured, the accuracy was calculated, and the energy consumption data was recorded by running the two models under the same hardware conditions to calculate the energy-efficiency ratio, which was finally statistically analyzed.

4.1 Detection Time

Detection time is directly related to the real-time response capability of the monitoring system, and the results of its comparison test are shown in Figure 2:

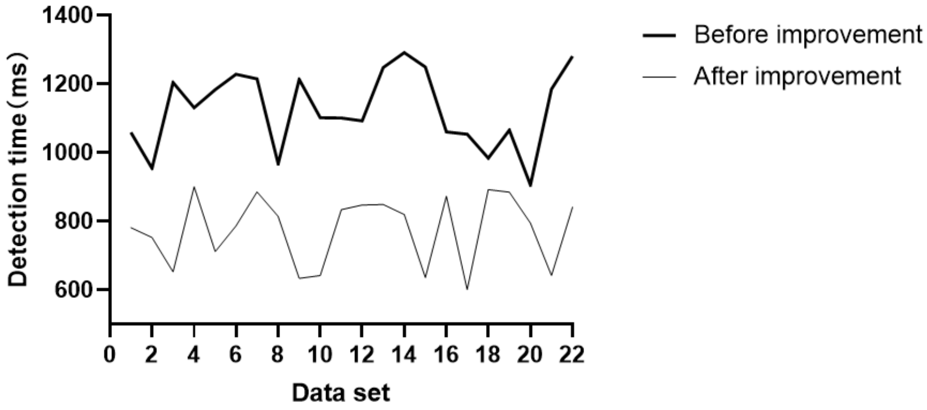

Figure 2: Comparison of detection time

As shown in Figure 2, the detection time of the improved SSD algorithm in employee abnormal behavior detection is much longer than that of the improved SSD algorithm. Among them, the highest detection time before the improvement is 1291ms and the lowest is 905ms, while the highest detection time after the improvement is only 900ms, which shows that the improved SSD algorithm has improved the detection speed. This improvement is very important for the real-time monitoring system that needs rapid response, because it reduces the delay time of the system's response to abnormal behavior and improves the real-time and reliability of the system.

4.2 Detection Accuracy

Detection accuracy not only reveals the specific contribution of algorithm improvement to improving detection accuracy, but also provides scientific basis for algorithm selection and optimization in monitoring systems, as shown in Figure 3:

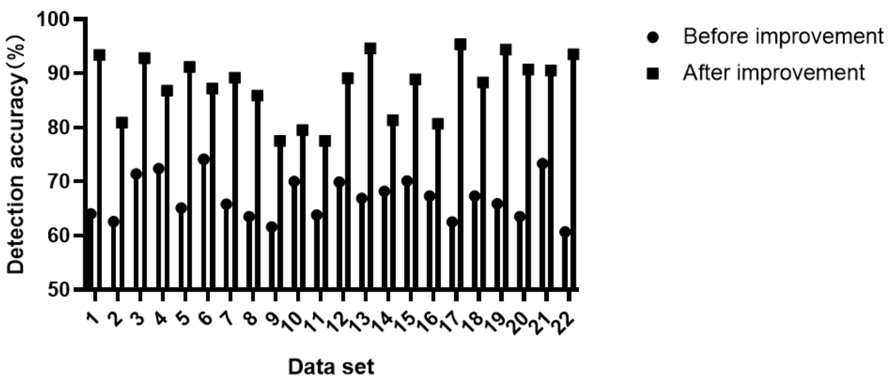

Figure 3: Comparison of detection accuracy

As can be seen from the data in Figure 3, the improved SSD algorithm has better detection accuracy than before. The detection accuracy before improvement is between 60.7% -74.1%, and the range after improvement is between 77.5%-95.4%. This significant improvement means that the algorithm is more accurate in distinguishing normal behavior from abnormal behavior, reducing false positives and false negatives, thereby improving the reliability and effectiveness of the monitoring system. The improvement in accuracy is due to the use of more advanced feature extraction techniques, optimized model structures, and enhanced data augmentation strategies in the improved algorithm.

4.3 Energy Efficiency Ratio

The significance of the energy efficiency ratio comparison experiment is that it can reveal the effectiveness of the algorithm improvement in terms of energy efficiency and provide decision support for realizing sustainable development and reducing operational costs. As shown in Figure 4, the results of energy efficiency ratio comparison of SSD algorithm before and after improvement are shown:

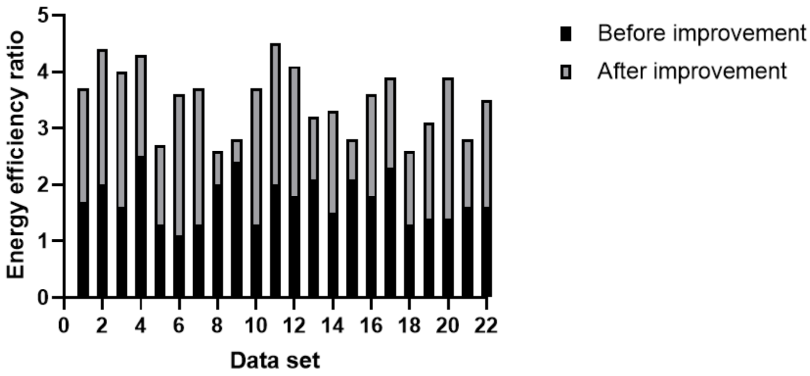

Figure 4: Energy efficiency comparison

As shown in Figure 4, the energy efficiency ratio of the SSD algorithm before and after the improvement has a large difference, the maximum energy efficiency ratio is 2.5 before the improvement, and after the improvement the energy efficiency ratio of the SSD algorithm reaches a maximum of 4.5, which is 2 different from the maximum energy efficiency ratio of the SSD algorithm before the improvement, which implies that, in accomplishing the same amount of the target detection task, the improved algorithm consumes less energy, or at the same energy consumption, the improved algorithm is able to process more data, thus providing higher performance.

5. Conclusion

This study successfully improved the SSD algorithm to enhance the performance of employee abnormal behavior detection. By introducing advanced feature fusion techniques and optimized network structure, the improved SSD algorithm in this study achieved significant improvement in both key metrics of accuracy and real-time performance. In addition, through model compression and acceleration techniques, this study also optimized the energy-efficiency ratio of the algorithm to make it more suitable for the application requirements of real-time monitoring scenarios.

The experimental results confirm the effectiveness of the improved SSD algorithm and its remarkable advantages in dealing with complex scenes, multi-target recognition and real-time processing. These improvements are of great significance to improve the response speed and detection effect of industrial safety monitoring system. The research in this paper can not only provide a new and effective solution for employee abnormal behavior detection, but also provide relevant reference and inspiration for future video analysis tasks based on deep learning. In the future work, the generalization ability and robustness of the algorithm to a wider range of scenarios and the actual effect in real monitoring environment will be further investigated. At the same time, the improved algorithm will be combined with other deep learning models, which will be a new breakthrough in current video analysis, which will make the motion detection technology more abundant and practical.

References

[1] Tie Fuzhen. Algorithm for detecting abnormal behavior of crowds in video surveillance based on improved optical flow method [J]. Modern Electronic Technology, 2024,47 (7): 45-48.

[2] Zhu Xianyuan, Wang Songlin, Zhou Yefan, Han Haifeng. An Improved Classroom Abnormal Behavior Perception Detection Algorithm for YOLOv5 [J]. Journal of Qiqihar University (Natural Science Edition), 2023,39 (1): 46-52.

[3] Li Zhuoqing, Jia Zhentang. Abnormality behavior monitoring system and algorithm design based on deep learning [J]. Microcomputer Application, 2024,40 (3): 7-10.

[4] Zhao Lian, Zhou Lei, Guo Yuheng, Chen Huagui. Abnormal Behavior Detection in Factory Scenarios [J]. Software Guide, 2024,23 (1): 57-62.

[5] Zhu Qiang, Sun Chen, Xu Pan Yuchi, Yan Yunfeng. A two-stage detection algorithm for abnormal behavior in smart construction sites based on FCOS [J]. Zhejiang Electric Power, 2023,42 (4): 65-71.

[6] Liu D, Gao S, Chi W, et al. Pedestrian detection algorithm based on improved SSD[J]. International Journal of Computer Applications in Technology, 2021, 65(1): 25-35.

[7] Meng J, Jiang P, Wang J, et al. A mobilenet-SSD model with FPN for waste detection[J]. Journal of Electrical Engineering & Technology, 2022, 17(2): 1425-1431.

[8] Zhao M, Zhong Y, Sun D, et al. Accurate and efficient vehicle detection framework based on SSD algorithm[J]. IET Image Processing, 2021, 15(13): 3094-3104.

[9] Chen W, Qiao Y, Li Y. Inception-SSD: An improved single shot detector for vehicle detection[J]. Journal of Ambient Intelligence and Humanized Computing, 2022, 13(11): 5047-5053.

[10] Wang Y, Liu X, Guo R. An object detection algorithm based on the feature pyramid network and single shot multibox detector[J]. Cluster Computing, 2022, 25(5): 3313-3324.

[11] Yan C, Zhang H, Li X, et al. R-SSD: Refined single shot multibox detector for pedestrian detection[J]. Applied Intelligence, 2022, 52(9): 10430-10447.

[12] Feng T. Mask RCNN-based single shot multibox detector for gesture recognition in physical education[J]. Journal of Applied Science and Engineering, 2022, 26(3): 377-385.

[13] Farokhah L. Perbandingan Metode Deteksi Wajah Menggunakan OpenCV Haar Cascade, OpenCV Single Shot Multibox Detector (SSD) dan DLib CNN[J]. Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), 2021, 5(3): 609-614.

[14] Rogelio J, Dadios E, Bandala A, et al. Alignment control using visual servoing and mobilenet single-shot multi-box detection (SSD): A review[J]. International Journal of Advances in Intelligent Informatics, 2022, 8(1): 97-114.

[15] Zhang H, Hong X, Zhu L. Detecting small objects in thermal images using single-shot detector[J]. Automatic Control and Computer Sciences, 2021, 55(2): 202-211.

[16] Liu S, Huang L, Zhao Y, et al. Lightweight Single Shot Multi-Box Detector: A fabric defect detection algorithm incorporating parallel dilated convolution and dual channel attention[J]. Textile Research Journal, 2024, 94(1-2): 209-224.

[17] Zhang H, Niu M, Chen X, et al. (Retracted) Target recognition and localization based on lightweight single-shot multibox detector network for robotics[J]. Journal of Electronic Imaging, 2022, 31(6): 618-619.

[18] Kabolizadeh M, Abbasi M. Automatic Detection and Extraction of Qanat from Satellite Images with High Spatial Resolution of Google Earth based on Conventional Neural Networks Single-Shot Multi-box Detection (SSD)[J]. Engineering Journal of Geospatial Information Technology, 2023, 11(3): 85-102.

[19] Wen G, Cao P, Wang H, et al. MS-SSD: Multi-scale single shot detector for ship detection in remote sensing images [J]. Applied Intelligence, 2023, 53(2): 1586-1604.

[20] Juneja A, Juneja S, Soneja A, et al. Real time object detection using CNN based single shot detector model[J]. Journal of Information Technology Management, 2021, 13(1): 62-80.