Journal of Information Systems Engineering and Management
2023, 8(4), 22898
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Research on Optimization of Boundary Detection and Dangerous Area Warning Algorithms Based on Deep Learning in Campus Security System

Baitong Zhong[1,2*], Johan Bin Mohamad Sharif[1], Sah Salam[1], Chengke Ran[2], Yizhou Liang[3], Zijun Cheng[4]

[1] Lecturer, College of Information Engineering, Universiti Teknologi Malaysia, Johor, Malaysia
[2] Lecturer, College of Information Engineering, Hunan Mechanical Electrical Polytechnic, Changsha, China
[3] Lecturer, School of Computer Science and Engineering, Central South University, Changsha, China
[4] Lecturer, Changde City Economic Construction Investment Group Co., LTD, Changde, China
* **Corresponding Author:** bvbv1234weih@163.com

Citation: Zhong, B., Sharif, J. B. M., Ran, C., Liang, Y., & Cheng, Z. (2023). Research on Optimization of Boundary Detection and Dangerous Area Warning Algorithms Based on Deep Learning in Campus Security System. *Journal of Information Systems Engineering and Management*, 8(4), 22898. https://doi.org/10.55267/iadt.07.13844

ARTICLE INFO

Received: 03 Aug 2023
Accepted: 02 Oct 2023

ABSTRACT

This study designs and implements a boundary detection and dangerous area warning algorithm based on deep learning from the perspective of typified campus security situation resources such as data, information, and knowledge. Based on integrating multiple campus security factors, real-time perception and further prediction of campus security situation can be achieved. Through coordinated operation among various algorithm modules, object intrusion in specific areas can be accurately identified and early warning can be given. The research results show that when an object invades a specific area, the difference coefficient will increase, and the larger the change value in the intrusion area, the larger the corresponding difference coefficient. By using this feature, the threshold of the difference coefficient can be determined. When a region is invaded, the contour length of the foreground will sharply increase. Based on the statistical information of the contour length of the foreground, the threshold can be set to determine whether someone has invaded the region. The deep learning algorithm in this study accurately extracts the contour of moving targets and can identify foreground targets. The real-time performance of the algorithm is also guaranteed, and it has high practical value in intelligent video monitoring. This algorithm greatly improves the efficiency of intrusion detection by utilizing the joint constraints of two types of time-domain and scene-space transformations in monitoring images. This method is not affected by the brightness of the regional environment, nor will it cause misjudgment due to significant differences in brightness of the regional environment. The detection and inference time of deep learning-based detection methods is controlled within 2-3ms, and the FPS value of the detection method is always at a high level, which can quickly increase to over 350frames/s after transmission begins. The detection method based on deep learning has higher detection efficiency.

Keywords: Deep Learning, Campus Security, Intrusion Detection, Regional Environment.

INTRODUCTION

In recent years, the security issues on university campuses have become increasingly severe. To ensure the personal safety of students, video surveillance systems have been widely applied to campuses. To ensure a good teaching and research atmosphere in universities, especially to ensure a good public security environment, it is necessary to establish a comprehensive security prevention system on university campuses (Ahmed, Parvin, Haque, & Uddin, 2020; Liu, 2021). As an important technical means of campus security management, the security monitoring system plays an important role in preventing property damage on campus, reducing and avoiding various unsafe factors such as safety accidents and illegal criminal cases (Accattoli, Sernani, Falcionelli, Mekuria, & Dragoni, 2020; Liu, Shi, Zhang, Ou, & Wang, 2020). The widespread application of video surveillance technology has provided new solutions for the safety prevention methods of college students. Fast and accurate judgment depends on the accurate identification of risk warning modes for college students' safety prevention, including the construction and integration of functional

modules such as data collection systems, information analysis and prediction systems, and decision alarm systems (Zhang, Liu, Zhang, Zhang, & Zhu, 2019; Rahmatov, Paul, Saeed, & Seo, 2021). Universities can establish an indicator system and weight that covers factors such as students' daily habits, online speech, physiological reasons, interests, campus life, and interpersonal relationships. By utilizing modern online monitoring technology and data mining technology, they can provide real-time feedback on students' behaviors, improve the accuracy of analyzing and predicting group safety risk levels, and achieve accurate and diverse dynamic monitoring of students on campus (Lei & Wu, 2020; Qin, Cao, & Ji, 2021). The current safety management system for college students in Chinese universities is far from fully aware of the changes brought about by this technology, and scattered and multifaceted safety risks still exist (Khan, Wang, Riaz, Elfatyany, & Karim, 2021). In this regard, big data and intelligent video detection technology can be regarded as effective means for colleges and universities to achieve the refinement of student safety management (Zhang, Xiao, Yang, Xiang, & Zhong, 2019).

Yang et al. (2023) stated that efforts to solve these technology problems have persisted over many years. As an important technical measure to ensure campus safety, the layout of the security monitoring system determines the coverage range of the security status parameters that can be collected (Guo, Che, Shahidehpour, & Wan, 2021). However, in actual scenarios, it is often limited by the low resolution of the monitoring equipment and the distance between the detection target and the camera, resulting in low accuracy of the video monitoring system established to reflect the risk status of different areas of the campus (Huang, Zhao, Yan, Liu, & Zhou, 2020). Therefore, upgrading the existing campus security monitoring system to a more accurate, efficient, and intelligent security detection system has become one of the effective ways to improve campus security (Chen & Shi, 2021). This study applies the continuity of target motion and the spatiotemporal continuity of image sequences in intrusion detection and proposes an improved three-frame difference algorithm based on spatiotemporal constraints, which improves detection performance by combining the direct continuity of image sequences with the spatial structure of the scene. In intelligent monitoring systems, boundary detection and dangerous area warning have successfully ensured personal safety (Rafiq, Shi, Zhang, Li, Ochani, & Shah, 2021; Brandon & Price, 2020). Boundary detection and dangerous area warnings are important intelligent recognition and automatic alarm technologies. The so-called boundary detection and dangerous area warning refer to users adding virtual warning lines with position directions on video images based on their actual needs (Zhu, Zhang, Cui, Wang, & Tang, 2022; Alahmadi, Hussain, Aboalsamh, & Zuair, 2020). In the monitoring field of view, once a moving object enters the set area, the system will automatically trigger an alarm signal to remind security personnel to pay attention and promptly handle the intrusion event. How to extract information quickly and effectively will become an important research topic (Kim, 2022). From the definition of boundary detection and dangerous area warning, implementing their functions includes moving object detection and intrusion condition determination (Xiao, Xu, Xing, Luo, Dai, & Zhan, 2021; Liu, Zhao, Yan, Huang, Mueed, & Meng, 2020). The moving object detection method utilizes background modeling to detect the motion of the target. In practical situations, commonly used background modeling methods are susceptible to changes in lighting and physical environment (Teng, Wang, Zhang, & He, 2020). On the other hand, the displacement of the detection area caused by the projection of three-dimensional scenes onto a two-dimensional plane in real scenes also poses technical difficulties for boundary detection and dangerous area warnings (Liu, 2022). Previous research has mainly focused on campus security data collection and single-level utilization, with limited involvement in multi-level data collection and fusion, making it difficult to provide timely feedback on campus security. Particularly demanding will be the real-time and granularity of performance needed for telesurgery and holographic communication (Hermawan, Putri, & Kartanto, 2022). Therefore, there is an urgent need to research efficient algorithms and data operation models suitable for campus security, in order to build a campus security monitoring system.

The issue of campus safety in social security governance is crucial and has become a hot topic of concern for all sectors of society. This paper focuses on issues such as incomplete utilization of campus security data, low intelligence of campus security monitoring, and untimely response to campus security monitoring. As the foundation for ensuring the construction of a smart campus security environment, situational awareness is an important solution to campus security issues. This article aims to design and implement a boundary detection and danger zone warning algorithm based on deep learning. From the perspective of typified campus security situation resources such as data, information, and knowledge, this study integrates multiple campus security factors to perceive and further predict the campus security situation in real time. When campus security is threatened, automatic alarms are triggered to alert administrators. Through coordinated operations between various algorithm modules, object intrusion in specific areas can be accurately identified and early warning can be given.

BOUNDARY DETECTION METHOD

Hazardous Area Drawing

This article mainly studies the video region detection algorithm based on OpenCV and proposes to improve detection performance by utilizing the constraints of temporal and spatial topological structures between image sequences (Pan, Li, & Zhao, 2021). The algorithm structure is shown in **Figure 1**. This function mainly utilizes the cvCvtColor function, cvSmooth function, cvSobel function, and cvThreshold function in the OpenCV1.0 library to smooth images. The cvCvtColor function is used to convert the RGB three-channel image obtained from a digital camera into a grayscale single-channel image (Shao, Zhu, Tan, Hao, & Ma, 2020); Smooth the image frames using the cvSmooth

function; Using the cvSobel function to extract edges from images; Finally, the edge image is binarized through the cvThreshold function to obtain the Binary image required by the detection algorithm. Compared with previous methods, the method in this study has significant computational advantages in data acquisition, bad point removal, data operation, and model data optimization. The algorithm in this study belongs to end-to-end learning, and the results can be obtained by inputting data, which is convenient and fast. Without manually designing rules, deep learning can optimize the loss function as much as possible to learn rules. It is possible to explore the potential features of the data as much as possible.

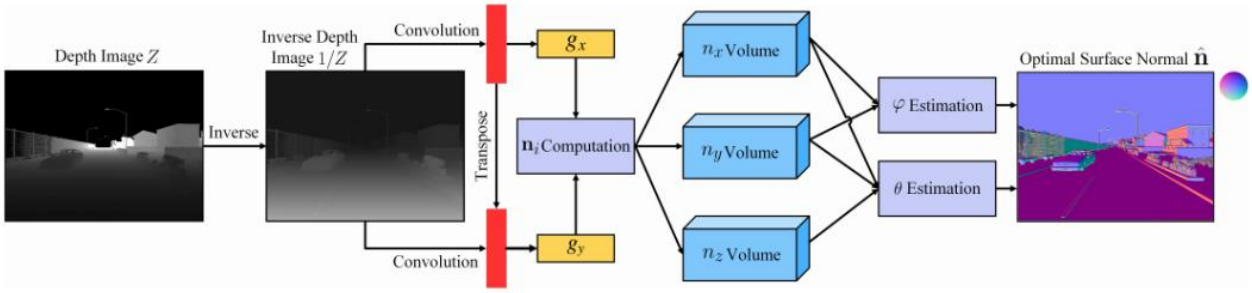

Figure 1. OpenCV Video Region Detection Algorithm

The method of drawing detection areas is to allow users to directly use the mouse to draw "dangerous areas" on the video. The specific method used is to set the trip wire number in the program, as shown in **Figure 2**: the judgment conditions for fully entering the detection area and partially entering the detection area are not the same. Target intrusion zone detection methods can be divided into two types of trip line problems: one is that only a portion of the moving target enters the depicted area; Another approach is for the moving target to fully enter the depicted area (Shao et al., 2020). Represent the moving target as several rectangles, and when any boundary of the rectangle intersects with the boundary of the set danger area, the system will immediately alarm. When the detection target enters the set area, it is necessary to prove that any line segment of the target matrix is on both sides of the dangerous area boundary, that is, to prove that the cross product of line segments p1p3 and p1p4 is different from line segments p2p3 and p2p4. When the moving target invades the "dangerous area", point (x1, y1) is p1, point (x2, y2) is p2, point (x3, y3) is p3, and point (x4, y4) is p4.

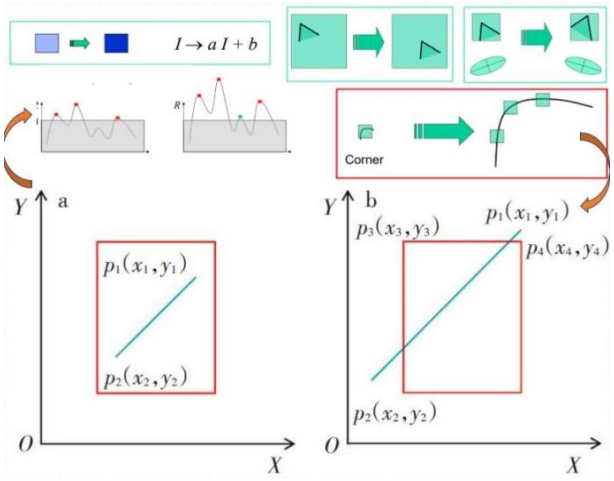

Figure 2. Schematic Diagram of Moving Target Entering the Setting Area

Constructing an intelligent campus security detection model can achieve campus security detection functions. This model utilizes video surveillance equipment to capture real-time images of the campus, including teaching buildings, dormitories, and laboratories (Arbeiter, Maier, & Spöck, 2021). It uses JetsonAGXXavier edge devices to call security detection algorithms and malignant load detection algorithms to monitor the safety of the campus in real time (Liu et al., 2020). Taking fire and electricity safety on the campus as examples, this model uses video surveillance equipment to capture real-time scenes on the campus, Intelligent sockets are used to measure the air conditioning sockets in each dormitory building, and JetsonAGXXavier's edge devices, hazard detection algorithms, and malicious load detection algorithms are used to monitor the fire situation in the school in real-time. If there is a malignant load situation, the system can issue a power outage command to the intelligent sockets in the dormitory building, teaching building, or experimental building, and notify the relevant management personnel to issue a warning to the relevant building staff (Zahra et al., 2021). When the camera device transmits Real Time Streaming Protocol (RTSP) video streams to the edge, the video processing module is first called to extract video frame image data from the video stream and input it into the hazard detection algorithm module. Secondly, the program calls the hazard detection algorithm model deployed in the JetsonAGXXavier edge device to detect smoke and flame targets (Zhang et al., 2020). Finally, the recognition results are transmitted to the alarm module, and when a power outage alarm is received, the alarm information is transmitted to cloud computing to inform relevant management personnel. After discovering dangerous targets, based on the size of the target, the level of the dangerous event is determined, and it is sent to the relevant security personnel through the network, enabling them to quickly go to the scene for investigation and achieve intelligent detection and management of campus security (Brandon & Price, 2020).

Three-frame Difference Method Based on Spatiotemporal Constraints

Due to the ability of the inter-frame difference method to quickly detect the range of motion caused by moving targets in adjacent images, the extracted moving targets are often larger than the actual targets, and there is usually a phenomenon of "shadow" (Tan, Sia, & Tang, 2022; Huang et al., 2020). The detected objects are the parts that change relative to the front and back frames, and the overlapping parts cannot be detected, leading to the phenomenon of a "hole" in the detected targets. So based on this, this study proposes a three-frame difference method, which uses adjacent three-frame images as a set of data for further differentiation, which can effectively detect the shape contour of moving targets in the middle frame (Zhang et al., 2020; Benouareth, 2021). At the same time, this article adds time and regional constraints to the three-frame difference method, further strengthening its application in practical scenarios. Based on multi-frame constraints in the time domain, utilizing the geometric attributes of the warning area in the image space to draw joint constraints in the spatiotemporal domain can improve the accuracy and practicality of intrusion detection algorithms.

Algorithm process: Select three consecutive frames of a video image sequence as $P_{i-1}^{t-\Delta t}(x,y)$, $P_i^t(x,y)$ and $P_{i+1}^{t+\Delta t}(x,y)$, t is the current detection image time point, Δ T is any time interval set, and the size can be set manually. When there is a high traffic volume during school or after school Δ T can be set to 0.5 seconds and can be set to 2 seconds when the traffic is low, which shares the computational pressure with the real-time monitoring system, making the system more efficient and intelligent. Calculate the interpolation of adjacent two frames of images using formula (1):

$$\begin{aligned} |d_{(i-1)}^{tl}(x,y) &= |P_i^t(x,y) - P_{i-1}^{t-\Delta t}(x,y)| \\ d_{(i,i+\eta)}^{t2}(x,y) &= |P_i^t(x,y) - P_{i+1}^{t+\Delta t}(x,y)| \end{aligned} \quad (1)$$

By selecting an appropriate threshold for the obtained difference image τ Perform binarization:

$$\begin{aligned} m_{(i,i-l)}(x,y) &= \begin{cases} 1 d_{(f,i-)}^{tl}(x,y)\tau \\ 0 d_{(f,i-1)}^{tl}(x,y) < \tau \end{cases} \\ m_{(i,i+l)}(x,y) &= \begin{cases} 1 d_{(f,i+1)}^{t2}(x,y)\tau \\ 0 d_{(f,i+1)}^{(2} (x,y) < \tau \end{cases} \end{aligned} \quad (2)$$

At each pixel (x, y), logically "AND" the obtained Binary image to obtain the Binary image of the middle frame of three images:

$$M_i(x,y) = \begin{cases} 1 m_{(i,i-1)}(x,y) \cap m_{(i,i+l)}(x,y) = 1 \\ 0 m_{(i,i-l)}(x,y) \cap m_{(i,i+1)}(x,y) \neq 1 \end{cases} \quad (3)$$

Compared with mixed Gaussian background modeling and the ViBe method, this method has faster iteration speed and lower program design complexity. Its disadvantage is the threshold τ The selection is difficult (Alahmadi, Hussain, Aboalsamh, & Zuair, 2020; Liu, Xiang, Shi, Zhang, & Wu, 2020). If the value is too low, the background noise of the selected frame image is too large, which is prone to incorrect detection of moving targets; if τ If the value is too high, it is easy to lose some moving targets when multiple moving targets appear in the "dangerous area", resulting in false detection, τ The appropriateness of the value directly affects the detection efficiency. Two typical application scenarios were studied for threshold selection in algorithms: day and night monitoring scenarios (Liu et al., 2020; Shao et al., 2020). During the day, there is a high traffic volume, Δ T and τ The selection of values should be appropriately small, increasing the interval between time-domain sampling and improving accuracy; On the other hand, there is less traffic at night, which is suitable for larger parameters from the perspective of resource conservation. Therefore, this article presents experimental results under two parameter configurations. When the three-frame difference algorithm based on spatiotemporal constraints is applied to practical scenarios, intrusion detection is divided into two categories in the time domain: one is daytime detection. If a target moves to the designated danger area, the system will not immediately alarm but will issue a prompt to alert the observer (Liu, 2021; Zahra et al., 2021). When the target stops in the area and stays for too long, the system will immediately alarm; Another type is night detection, where no matter how the target moves, as long as it invades the "dangerous area", the system will immediately alarm (**Figure 3**). Based on alarm accuracy, the concepts of false alarm rate and missed alarm rate were introduced to better evaluate the efficiency value of this algorithm (Liu et al., 2020). Assuming there are a total of T positive cases of alarms and F negative cases without alarms, and after confirmation of the alarm, there are Tt system alarms that are evaluated as positive cases and Ft system alarms that are evaluated as negative cases, then Tt+Ft=T; If there are Ff systems that are not alarmed and are judged as negative examples by humans, and Tf systems that are not alarmed and are judged as positive examples by humans, then Ff+Tf=F.

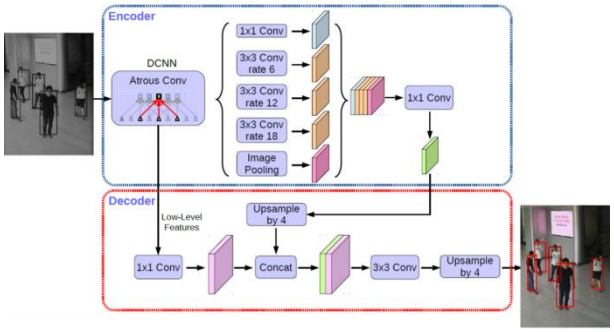

Figure 3. Architecture of Intrusion Alarm Situational Awareness Recognition Model

The introduction of spatiotemporal joint constraints effectively reduces the false alarm rate of the system. Compared with other algorithms, the three-frame difference method detects background images with no accumulation, fast update speed, and low computational complexity, overcoming the defects of large and blurry detection areas for moving objects. Intelligently perceive and intelligently analyze newly constructed video resource images, combine feature aggregation methods, use the intermediate feature

layer generated by the backbone network, and improve the detection effect of detection methods on security targets based on increasing a small amount of complexity. To further improve the generalization ability of the detection method, data augmentation technology is introduced to enhance the existing limited samples. Four randomly cut images are concatenated into one image to obtain a new training map. This algorithm can provide more background information for different training images. In addition, image stitching can increase batch size, and when performing batch normalization operations, four images can be simultaneously processed, improving training efficiency.

METHODOLOGY

Improved FAM-YOLOv4 Algorithm Model

The Scaled YOLOv4 algorithm proposes a network scaling method that allows the network to freely scale up and down. The scaled network is suitable for different scenarios, balancing speed and accuracy. The reduced YOLOv4 network model has a certain improvement in speed, but it has a significant loss in accuracy and cannot detect dangerous targets effectively. To further apply to embedded edge deployment, an improved method is proposed based on the original algorithm. This article selects the YOLOv4 small model as the baseline, which further improves the speed and accuracy of the algorithm in dangerous target detection tasks. Given the problem that the original algorithm has too many parameters and calculations to achieve the reasoning speed of real-time detection on an Edge device, the YOLOv4 small model is selected, and the original backbone network is replaced by MobileNetv2 after synchronized reduction to improve the reasoning speed of the algorithm on Edge device. A feature aggregation method is proposed to address the issue of poor detection performance of dangerous targets. At the same time, the CBAM attention mechanism is introduced in the network output section to improve the algorithm model's ability to capture global information. Mosaic and Mixu data augmentation methods are used in the training phase to improve the model's generalization ability. The improved algorithm model is called FAM-YOLOv4. The specific structure of the improved FAM-YOLOv4 network model is shown in **Figure 4**, where the area of the FAM box is the feature aggregation method proposed in this paper, the InvertedResdual module is the backbone submodule in the synchronously reduced MobileNetV2, SPPCSP, and BottleneckCSP2 are the next submodule in YOLOv4, Conv module is the ordinary convolution module, and CBAM is the attention mechanism module.

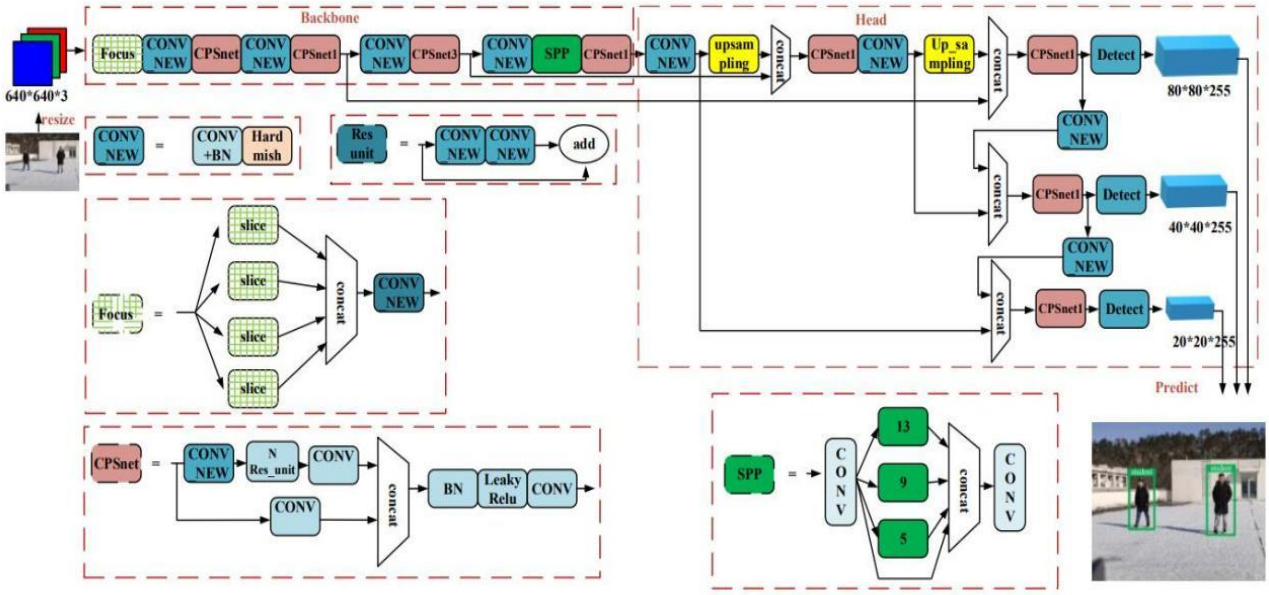

Figure 4. FAM-YOLOv4 Network Model Structure

Lightweight Backbone Network

To improve the reasoning speed of the algorithm on the Edge device and reduce the complexity of the network model, the synchronized reduced MobileNetV2 network is introduced to replace the backbone network of the original network, reduce the redundancy of the feature extraction network, and realize the lightweight of the backbone network. MobileNetV proposes a deep separable convolution suitable for the construction of lightweight neural networks, which mainly consists of deep-wise and point-wise convolutions. **Figure 5** shows a comparison of the structures of standard convolutions and deep separable convolutions, with the standard convolution structure on the left and the deep separable convolution structure on the right.

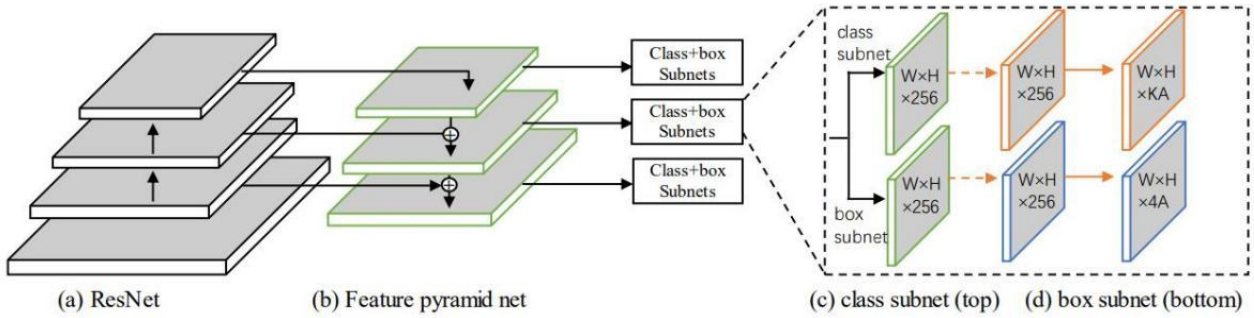

Figure 5. Standard Convolution and Depth Separable Convolution Structures

Compared with standard convolutions, deep separable convolutions have smaller complexity and fewer parameter quantities, thus adapting to fewer embedded devices, constructing lightweight network structures, and effectively improving the inference speed of algorithms. The complexity of standard convolution and depth separable convolution is described in the following formula.

$$\frac{\text{DW Conv}}{\text{Standard Conv}} = \frac{1}{K^2} + \frac{1}{C_{out}} \sim \frac{1}{K^2} \quad (4)$$

Where is the number of output channels, and is the size of the convolutional kernel. According to the formula, the complexity of deep separable convolutions is much lower than that of standard convolution structures. When obtaining feature maps of the same size, the computational complexity of deep separable convolutions is much lower than that of standard convolutions. MobileNetV2 further introduces residual thinking, improving accuracy based on lightweight. In the process of replacing the original network backbone network, follow YOOv4 small's original scaling principle (depth=0.33, width=0.5), and scale the MobileNetV2 network at the same level, that is, reduce the depth of each stage to 1/3 of the original depth, and reduce the number of output channels of the convolution core in the module to 1/2 of the original network. The scaled MoibileNetV2 only contains 0.3M parameter quantity, which is 12 times lower than YOLOv4 small's backbone parameter quantity, further improving the inference speed of the model without significant loss of model accuracy.

Although the use of the MobileNet backbone network has improved the speed to a certain extent, it cannot solve the problem of the poor detection effect of the algorithm in dangerous targets. In this paper, a feature aggregation method is proposed, which makes reasonable use of the intermediate feature layer generated by the backbone network and further improves the detection effect of the model in smoke and flame targets with a small amount of complexity. **Figure 6** shows the on-site application of FAM in hazardous area identification.

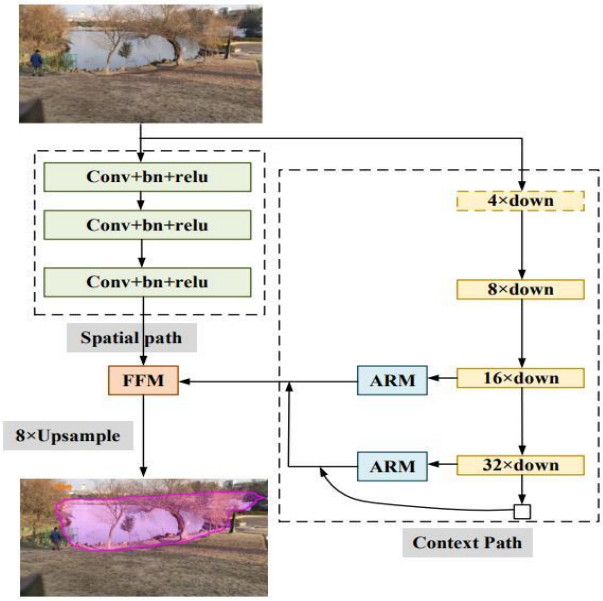

Figure 6. The On-site Application of FAM in Hazardous Area Identification

Data Augmentation Methods

Big data technology has powerful functions such as anytime, anywhere fidelity recording, permanent storage, and restorative imagery. The data in this study are all campus security monitoring data, sourced from digital monitoring equipment installed on the campus. The obtained data complies with ethical and ethical standards, does not involve sensitive or private data, and has been approved and supervised by the school in terms of data utilization and acquisition. The obtained data will be destroyed after the security assessment is completed. Data augmentation is a data augmentation technique that refers to the use of existing limited data to create as much new data as possible. Good data augmentation methods can improve algorithm performance and enhance the generalization ability of network models. In addition to conventional data augmentation methods such as random flipping, random rotation, random cropping, and deformation scaling, this article also adopts a mixture of Mosaic and Mixup data augmentation methods to enhance the dangerous dataset and improve the model's generalization ability. The main idea of the Mosaic data augmentation method is to randomly crop four images first, and then randomly concatenate the cropped images into one image as a new training image,

which is an effective data augmentation method. Its advantage lies in enriching the background of different training images while increasing the number of small targets and increasing the complexity of training data. Mixup is a mixed-category data augmentation method that effectively solves the problem of single data categories. By controlling the proportion of different categories through hyperparameters, data fusion is carried out to expand the training dataset. It is a data-independent data augmentation method that constructs virtual data for data augmentation during the training phase.

Dynamic Identification Model for Hazardous Areas

Dynamic Background Shadow Elimination: In the detection and tracking of moving targets, the segmentation of moving targets is a crucial step. Moving targets always appear with their shadows, causing changes in their physical shape. Therefore, shadow elimination is a key point in identifying moving targets. Therefore, the article analyzes the characteristics of shadow HSV color space, utilizes the different characteristics of shadows and moving targets in the H, S, and V components, calculates their corresponding thresholds in specific situations, and uses this threshold to segment and eliminate shadows. Imax and Imin are important parameters for color space conversion. Imax is the maximum of I (R, G, B) three-channel values, and Imin is the minimum of I (R, G, B) three-channel values. The conversion formula for converting RGB color spaces to HSV color space is as follows. Perform difference calculation, i.e. calculate the following values separately:

$$\begin{aligned} D_H &= F_H - B_H \\ D_V &= F_V - B_V \\ D_S &= F_S - B_S \end{aligned} \quad (5)$$

Obtain the difference between each component of the foreground and background images in the new color space model, and perform threshold segmentation after calculating the difference. By performing binarization based on the difference, the shadow and actual foreground area can be separated after segmentation according to the above threshold.

$$f_H(x,y) = \begin{cases} 255 & D_H(x,y) \leqslant \alpha \\ 0 & D_H(x,y) > \alpha \end{cases}$$

$$f_S(x,y) = \begin{cases} 255 & D_S(x,y) \leqslant \delta \\ 0 & D_S(x,y) > \delta \end{cases} \quad (6)$$

$$f_V(x,y) = \begin{cases} 255 & D_V(x,y) \leqslant \varphi \\ 0 & D_V(x,y) > \varphi \end{cases}$$

Global Grayscale and Contour Background Extraction

The global grayscale means statistical information is used to handle false foreground targets caused by sudden changes in light or excessive accumulation of light over a long period. The change in light can lead to a decrease in the matching degree of background pixels, which is considered a foreground target, resulting in the appearance of false foreground targets. If the global grayscale mean is Avggray, calculate the global grayscale mean as follows:

$$Avg_{gray} = \sum_{i=0,j=0}^{i=W,j=H} I_{ij}\,(\,gray\,) \quad (7)$$

The Learning rate of the Gaussian model can be adjusted according to the gray mean α， This enables the model to adapt to the brightness changes of the scene. To reflect the changes in grayscale, the mean change rate of grayscale is introduced θ:

$$\theta = \left| \frac{Avg_{gray}\ (n+1) - Avg_{gray}\ (n)}{Avg_{gray}\ (n)} \right| \quad (8)$$

After the above steps of processing, a relatively accurate foreground image can be extracted. Extracting contour information from the foreground image can determine whether there are any intrusions in a specific area and the number of intruders. The Canny operator is used in the article to extract foreground contour information. The contour information is obtained by connecting the edge points. Determine regional intrusion by counting the length of contours. The key findings show that when a region is invaded, the contour length of the foreground will sharply increase. Based on the statistical information of the contour length of the foreground, a threshold can be set to determine whether someone has invaded the region. Through the results of video testing, it can be seen that the background modeling method in the article has achieved good results, which also makes the segmentation effect of foreground targets very good. The contour extraction of moving targets is very accurate, and the moving foreground targets can be marked. The real-time performance of the algorithm is also guaranteed, and it has high practical value in intelligent video monitoring.

DISCUSSION

Experimental Results and Analysis of Static Background Model

The following is the rendering after each step of the algorithm. The background contour is shown in **Figure 7** (a), the real-time frame contour is shown in **Figure 7** (b), and the differential image is shown in **Figure 7** (c). This method can effectively extract foreground targets and determine whether a specific area has been invaded by people through differential coefficients.

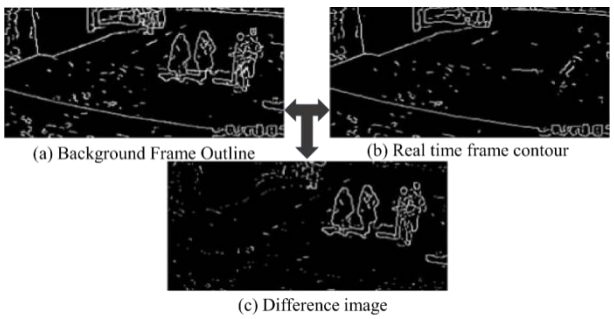

Figure 7. Effect Diagram of Each Step of the Algorithm

From the figure, it can be seen that when an object invades a specific area, the difference coefficient will increase, and the larger the change value in the invaded area, the larger the corresponding difference coefficient. By using this feature, the threshold of the difference coefficient can be

determined. The success rate of this model is still high, but it has a serious drawback that the background is fixed and unchanging, which can cause false positives for dynamic backgrounds. Analyze in depth the general characteristics of the changes in the contour, including length or other measures that reflect the overall size of the contour. The contour length of foreground image frames was examined and tested on the video set. From the above figure, it can be seen that when an area is invaded, the contour length of the foreground will sharply increase. Based on the statistical information of the contour length of the foreground, a threshold can be set to determine whether someone is invading the area. Through the results of video testing, it can be seen that the background modeling method in the article has achieved good results, which also makes the segmentation effect of foreground targets very good. The contour extraction of moving targets is very accurate, and the moving foreground targets can be marked. The real-time performance of the algorithm is also guaranteed, and it has high practical value in intelligent video monitoring. **Figure 8** shows the effect of character intrusion detection.

Figure 8. Comparison of Character Intrusion Detection Effects

The number of people counting algorithms in the article also has certain flaws. When characters overlap, as shown in the situation of two people on the right in the second row of **Figure 8**, it can cause the foreground to overlap, and the contour information of the foreground target will be incorrect. This way, the number of contours cannot represent the number of objects in the intrusion area, but it can still accurately determine that the area has been invaded. This is also the core issue that needs to be addressed in the next step. When the target object is obstructed, other features can be used to discover and label moving objects.

Comparative Analysis of the Accuracy of Different Algorithms

The experiment was divided into two periods for detection, and the parameters set for the three-frame difference method during the daytime period are as follows: τ Value is 10; Night period parameters τ The value is 12. Select different areas at the entrance of the Computer Academy building for testing, and use equipment such as Kinect2.0 and SOMITA tripod for video shooting. You can draw any geometric shape of a "dangerous area" on video images to simulate an "invasion" event when pedestrians cross the designated area on campus. When the target passes through an area, if it intersects with the area, it is considered an intrusion, as shown in **Figure 9**.

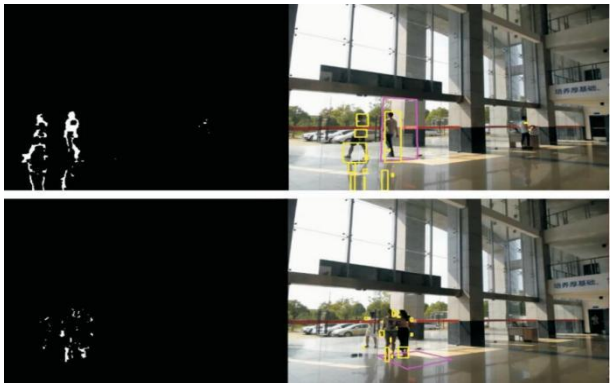

Figure 9. Detection of Different Areas During the Day

The shape of the detection area drawn in this experiment is a triangle, and the video recording period is from 2 p.m. during the day to 10 p.m. During the day, the school pond is defined as a dangerous area. When pedestrians "invade" the pond for a long time, to prevent accidents, it is defaulted to the target being in danger and promptly reporting to the police. Based on daytime experimental testing, experiments were added to detect the same area at different periods at night based on spatiotemporal constraints. At night, when pedestrians "invade" the pond area, the system will alarm immediately due to the small change in monitoring background frames to avoid timeout and false alarms. After evening classes, the entrance gate of the college is also defined as a "dangerous area". When students leave the college after class and the system sets the direction of the trip line to be consistent with the rules, the system will not alarm. If the detected target's movement direction is inconsistent with the rules, the system will immediately alarm to prevent unauthorized personnel from entering. The campus environment used in this system is not complex, and the main detection target is pedestrians. When the identification matrix of the target movement intersects with the boundary of the set area, it is determined whether the target violates the direction rules and whether to alarm or not. The results of the daytime experiment are shown in **Table 1**.

Table 1. Comparison of Experimental Data from Different Methods

Algorithm	Accuracy	False alarm rate	Missed alarm rate
Mixed Gaussian background	79.8%	37.5%	32.1%
ViBe	82.3%	27.4%	25.4%
Three Image Difference	83.7%	38.9%	43.4%
An Algorithm in this article	89.9%	17.2%	19.8%

Experimental data shows that when τ When the value is 10, the system has the highest accuracy. The three-frame difference method based on spatiotemporal constraints has improved the accuracy by 10.1%, 7.6%, and 6.2% compared to the mixed Gaussian background method, ViBe method, and three-frame difference method, respectively; The false alarm rate decreased by 20.3%, 10.2%, and 21.7% respectively; The missed alarm rate decreased by 12.3%, 5.6%, and 23.6% respectively. The data indicate that this algorithm greatly improves the efficiency of intrusion detection by utilizing the joint constraints of two types of time-domain and scene-space transformations in monitoring images. During the experimental process, it was found that this method is not affected by the brightness of the regional environment, nor will it cause misjudgment due to excessive differences in brightness of the regional environment. The main reason for the missed detection of experimental error types is that the overlapping parts of the targets between frames cannot be segmented, and can be misjudged as a single target, especially when the camera is placed far away and the distance between moving targets is small.

To test parameters τ impact of values on the performance of recognition algorithms has been experimentally investigated by adding accuracy experiments under different value conditions, as shown in **Figure 10**. It can be seen that τ When the value is 10, the algorithm has the best recognition performance. In future research, further research will be conducted on the adaptive value problem.

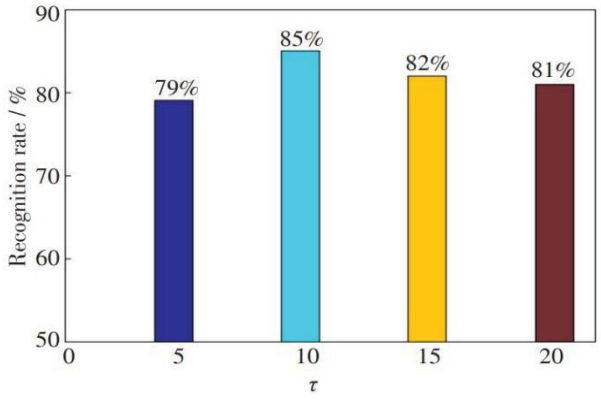

Figure 10. Different τ Accuracy of Value

Application and Effect Analysis of Detection Business of Importing Edge Device

To verify the application feasibility of the proposed detection method based on deep learning, it is introduced into Edge device to carry out detection business applications and analyze and study the application effect. The JetsonAGXXavier edge device uses the Linux operating system of Ubuntu 18.4, which sets the address in the device to the same network segment as the address of the host computer. Enable remote connection services for boundary terminal devices using SSH Server, with the main computer using shnvidia@192.168.3.249 Command to connect Jason AGXXavier boundary terminal device. To visually demonstrate the application advantages of the detection method, the results obtained by this detection method are compared with the detection results of the CNN-GRU-based detection method and adaptive detection method, respectively. Infer its efficiency and record the comparison results obtained, as shown in **Figure 11**.

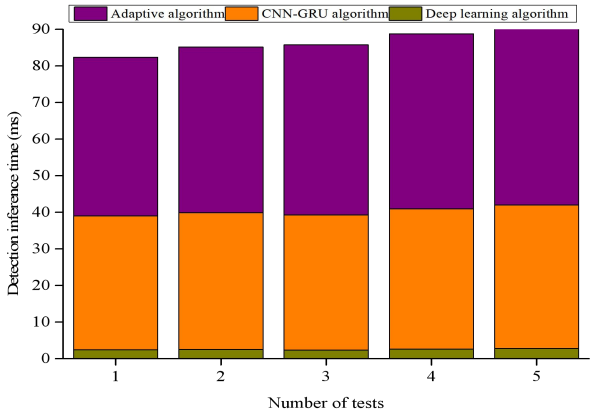

Figure 11. Comparison of Inference Time for Three Detection Methods

As shown in **Figure 11**, under the same detection conditions, the detection and inference time of the deep learning-based detection method is controlled within 2-3ms, while the detection and inference time of the other two detection methods exceeds 30ms. Therefore, the deep learning-based detection method has higher detection efficiency. Compare the Frames Per Second (FPS) of three detection methods to determine their video resource transmission capabilities. The higher the FPS, the stronger the video resource transmission capacity, and vice versa. Record the FPS values of the three detection methods and the results obtained are shown in **Figure 12**.

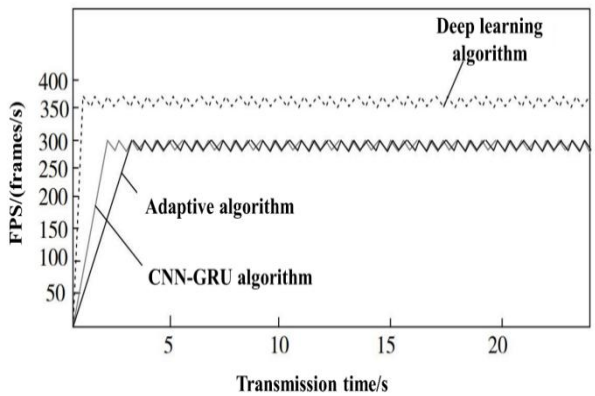

Figure 12. Comparison of FPS Values Among Three Detection Methods

From **Figure 12**, it can be seen that among the three detection methods, the deep learning-based detection method always has a high FPS value, which can quickly increase to over 350 frames/s after starting transmission. However, the other two detection methods have lower FPS values and require a longer time to stably transmit video resources after starting transmission. From this, it can be

seen that deep learning-based detection methods have higher detection efficiency.

CONCLUSION

This study designs and implements a boundary detection and dangerous area warning algorithm based on deep learning from the perspective of typified campus security situation resources such as data, information, and knowledge. Based on integrating multiple campus security factors, real-time perception and further prediction of the campus security situation can trigger automatic alarms to alert administrators when campus security is threatened. Through coordinated operations between various algorithm modules, object intrusion in specific areas can be accurately identified and early warning can be provided. The main research conclusions are as follows:

(1) In this paper, a Mixture model for hazard area identification is designed to model the background to extract moving targets, which achieves the goal of suppressing noise and improving target integrity. When an object invades a specific area, the difference coefficient will increase, and the larger the change value in the invaded area, the larger the corresponding difference coefficient. By using this feature, the threshold of the difference coefficient can be determined.

(2) When a region is invaded, the contour length of the foreground will sharply increase. Based on the statistical information of the contour length of the foreground, a threshold can be set to determine whether someone has invaded the region. Through the results of video testing, it can be seen that the background modeling method in the article has achieved good results, which also makes the segmentation effect of foreground targets very good. The contour extraction of moving targets is very accurate, and the moving foreground targets can be marked. The real-time performance of the algorithm is also guaranteed, and it has high practical value in intelligent video monitoring.

(3) This algorithm greatly improves the efficiency of intrusion detection by utilizing the joint constraints of two types of time-domain and scene-space transformations in monitoring images. During the experimental process, it was found that this method is not affected by the brightness of the regional environment, nor will it cause misjudgment due to excessive differences in brightness of the regional environment. The detection and inference time of deep learning-based detection methods is controlled within 2-3ms, while the detection and inference time of the other two detection methods exceeds 30ms. The FPS value of deep learning-based detection methods is always at a high level, and can quickly increase to over 350frames/s after starting transmission. Deep learning-based detection methods have higher detection efficiency.

(4) Artificial intelligence algorithms perform deep processing on sample data. The algorithms in this study provide timely security responses and feedback while processing campus security data. The databases in the samples have a significant impact on the accuracy of this study. In the future, this study will further train and provide feedback on models based on larger databases to achieve more effective data evaluation results.

REFERENCES

Accattoli, S., Sernani, P., Falcionelli, N., Mekuria, D. N., & Dragoni, A. F. (2020). Violence detection in videos by combining 3D convolutional neural networks and support vector machines. Applied Artificial Intelligence, 34(4), 329-344. https://doi.org/10.1080/08839514.2020.1723876

Ahmed, T., Parvin, M. S., Haque, M. R., & Uddin, M. S. (2020). Lung cancer detection using CT image based on 3D convolutional neural network. Journal of Computer and Communications, 8(03), 35. https://doi.org/10.4236/jcc.2020.83004

Alahmadi, A., Hussain, M., Aboalsamh, H. A., & Zuair, M. (2020). PCAPooL: unsupervised feature learning for face recognition using PCA, LBP, and pyramid pooling. Pattern Analysis and Applications, 23, 673-682. https://doi.org/10.1007/s10044-019-00818-y

Arbeiter, M., Maier, T., & Spöck, G. (2021). A cyber-physical environment for detecting exceptional and dangerous human behavior in the home by sensors and its verification by computer simulation. Adaptive Behavior, 29(6), 579-600. https://doi.org/10.1177/1059712320930420

Benouareth, A. (2021). An efficient face recognition approach combining likelihood-based sufficient dimension reduction and LDA. Multimedia Tools and Applications, 80(1), 1457-1486. https://doi.org/10.1007/s11042-020-09527-9

Brandon, N., & Price, P. S. (2020). Calibrating an agent-based model of longitudinal human activity patterns using the Consolidated Human Activity Database. Journal of exposure science & environmental epidemiology, 30(1), 194-204. https://doi.org/10.1038/s41370-019-0156-z

Chen, W., & Shi, K. (2021). Multi-scale attention convolutional neural network for time series classification. Neural Networks, 136, 126-140. https://doi.org/10.1016/j.neunet.2021.01.001

Guo, W., Che, L., Shahidehpour, M., & Wan, X. (2021). Machine-Learning based methods in short-term load forecasting. The Electricity Journal, 34(1), 106884. https://doi.org/10.1016/j.tej.2020.106884

Hermawan, D., Putri, N. M. D. K., & Kartanto, L. (2022). Cyber Physical System Based Smart Healthcare System with Federated Deep Learning Architectures with Data Analytics. International Journal of Communication Networks and Information Security, 14(2), 222-233. https://doi.org/10.17762/ijcnis.v14i2.5513

Huang, C., Zhao, Y., Yan, W., Liu, Q., & Zhou, J. (2020). A new method for predicting crosstalk of random cable bundle based on BAS-BP neural network algorithm.

IEEE Access, 8, 20224-20232. https://doi.org/10.1109/ACCESS.2020.2969221

Khan, M., Wang, H., Riaz, A., Elfatyany, A., & Karim, S. (2021). Bidirectional LSTM-RNN-based hybrid deep learning frameworks for univariate time series classification. The Journal of Supercomputing, 77, 7021-7045. https://doi.org/10.1007/s11227-020-03560-z

Kim, D. (2022). Research On Text Classification Based On Deep Neural Network. International Journal of Communication Networks and Information Security, 14(1s), 100-113. https://doi.org/10.17762/ijcnis.v14i1s.5618

Lei, Y., & Wu, Z. (2020). Time series classification based on statistical features. EURASIP Journal on Wireless Communications and Networking, 2020, 1-13. https://doi.org/10.1186/s13638-020-1661-4

Liu, Q. (2022). Aerobics posture recognition based on neural network and sensors. Neural Computing and Applications, 1-12. https://doi.org/10.1007/s00521-020-05632-w

Liu, Q., Zhao, Y., Yan, W., Huang, C., Mueed, A., & Meng, Z. (2020). A novel crosstalk estimation method for twist non-uniformity in twisted-wire pairs. IEEE Access, 8, 38318-38326. https://doi.org/10.1109/ACCESS.2020.2976136

Liu, Y., Liu, W., Shen, Y., Zhao, X., & Gao, S. (2021). Toward smart energy user: Real time non-intrusive load monitoring with simultaneous switching operations. Applied Energy, 287, 116616. https://doi.org/10.1016/j.apenergy.2021.116616

Liu, Z., Liu, G., Zhang, L., & Pu, J. (2020). Linear regression classification steered discriminative projection for dimension reduction. Multimedia Tools and Applications, 79, 11993-12005. https://doi.org/10.1007/s11042-019-08434-y

Liu, Z., Shi, K., Zhang, K., Ou, W., & Wang, L. (2020). Discriminative sparse embedding based on adaptive graph for dimension reduction. Engineering Applications of Artificial Intelligence, 94, 103758. https://doi.org/10.1016/j.engappai.2020.103758

Liu, Z., Xiang, L., Shi, K., Zhang, K., & Wu, Q. (2020). Robust manifold embedding for face recognition. IEEE Access, 8, 101224-101234. https://doi.org/10.1109/ACCESS.2020.2997953

Pan, H., Li, Y., & Zhao, D. (2021). Recognizing human behaviors from surveillance videos using the SSD algorithm. The Journal of Supercomputing, 77, 6852-6870. https://doi.org/10.1007/s11227-020-03578-3

Qin, Y. Y., Cao, J. T., & Ji, X. F. (2021). Fire detection method based on depthwise separable convolution and yolov3. International Journal of Automation and Computing, 18, 300-310. https://doi.org/10.1007/s11633-020-1269-5

Rafiq, H., Shi, X., Zhang, H., Li, H., Ochani, M. K., & Shah, A. A. (2021). Generalizability improvement of deep learning-based non-intrusive load monitoring system using data augmentation. IEEE Transactions on Smart Grid, 12(4), 3265-3277. https://doi.org/10.1109/TSG.2021.3082622

Rahmatov, N., Paul, A., Saeed, F., & Seo, H. (2021). Realtime fire detection using CNN and search space navigation. Journal of Real-Time Image Processing, 18, 1331-1340. https://doi.org/10.1007/s11554-021-01153-4

Shao, Z., Zhu, H., Tan, X., Hao, Y., & Ma, L. (2020). Deep multi-center learning for face alignment. Neurocomputing, 396, 477-486. https://doi.org/10.1016/j.neucom.2018.11.108

Tan, K. L., Sia, J. K. M., & Tang, K. H. D. (2022). Examining students' behavior towards campus security preparedness exercise: The role of perceived risk within the theory of planned behavior. Current Psychology, 41(7), 4358-4367. https://doi.org/10.1007/s12144-020-00951-6

Teng, Q., Wang, K., Zhang, L., & He, J. (2020). The layer-wise training convolutional neural networks using local loss for sensor-based human activity recognition. IEEE Sensors Journal, 20(13), 7265-7274. https://doi.org/10.1109/JSEN.2020.2978772

Xiao, Z., Xu, X., Xing, H., Luo, S., Dai, P., & Zhan, D. (2021). RTFN: A robust temporal feature network for time series classification. Information sciences, 571, 65-86. https://doi.org/10.1016/j.ins.2021.04.053

Yang, Y., Rosni, N. A., Zainol, R. B., Yang , X., Hu, H., & Wang, T. (2023). Daily Activities of Elder Adults Using Optimized Deep Learning Model in China. International Journal of Communication Networks and Information Security (IJCNIS), 15(3), 273-292. https://doi.org/10.17762/ijcnis.v15i3.6270

Zahra, S. B., Khan, M. A., Abbas, S., Khan, K. M., Al-Ghamdi, M. A., & Almotiri, S. H. (2021). Marker-based and marker-less motion capturing video data: Person and activity identification comparison based on machine learning approaches. Computers, Materials & Continua, 66(2), 1269-1282. https://doi.org/10.32604/cmc.2020.012778

Zhang, L., Liu, J., Zhang, B., Zhang, D., & Zhu, C. (2019). Deep cascade model-based face recognition: When deep-layered learning meets small data. IEEE Transactions on Image Processing, 29, 1016-1029. https://doi.org/10.1109/TIP.2019.2938307

Zhang, Y., Xiao, X., Yang, L. X., Xiang, Y., & Zhong, S. (2019). Secure and efficient outsourcing of PCA-based face recognition. IEEE Transactions on Information Forensics and Security, 15, 1683-1695. https://doi.org/10.1109/TIFS.2019.2947872

Zhu, H., Zhang, J., Cui, H., Wang, K., & Tang, Q. (2022). TCRAN: Multivariate time series classification using residual channel attention networks with time correction. Applied Soft Computing, 114, 108117. https://doi.org/10.1016/j.asoc.2021.108117