International Journal of Digital Crime and Forensics
Volume 16 • Issue 1 • January-December 2024

Laboratory Abnormal Behavior Detection Based on Multimodal Information Fusion

Dawei Zhang
https://orcid.org/0009-0006-2519-0244
Liaodong University, China

ABSTRACT

The traditional laboratory anomaly detection methods mainly focus on the hidden dangers caused by chemical leaks and other items, ignoring the impact of abnormal behaviors such as incorrect operations and improper behavior on safety in the laboratory. This paper proposes a laboratory abnormal behavior detection method based on multimodal information fusion. The method generates a dense optical flow field of RGB image sequences based on optical flow theory and global smoothing constraints, and mines motion mode information. Meanwhile, the contour modal information of behavior is captured through convolution and adjacency matrix operations. Using decision level and proximity functions to integrate student behavior motion mode information and contour mode information, and using the maximum value as the behavior detection result. The experimental results show that the method can effectively detect abnormal behavior in the laboratory environment, with small detection errors and a specificity close to 1.00, effectively ensuring the safety of the laboratory environment.

KEYWORDS

Multimodal Information Fusion, Optical Flow Theory, Abnormal Behavior Detection, Motion Mode Information, Contour Modal Information

INTRODUCTION

Behavioral management in laboratory environments is crucial for the safety of students and the normal operation of the laboratory. Although most students can abide by laboratory rules, unsafe behavior arises due to their weak safety awareness, lack of self-protection awareness, weak psychological resilience, and tendency to make unwise decisions in the event of laboratory accidents. At the same time, group factors also have an indirect impact on laboratory safety. The group environment in which students live is significant, and their goals and pressures have a direct impact on individual behavior. If the group safety atmosphere is poor, individuals are easily affected in both learning and living environments, and many students are driven by conformity psychology to unconsciously engage in unsafe behavior (Zhang et al., 2022; Li et al., 2022). Therefore, the detection of abnormal behavior in university laboratories has become an issue of high concern.

At present, research on unsafe human behavior pattern recognition, both domestically and internationally, mainly focuses on abnormal behavior detection using artificial intelligence. For example, Zhang et al. (2003) established an abnormal behavior detection network model using the fifth version of the “you only look once” (YOLO) family of object detection models network and a masked convolutional attention model. They learned and extracted behavior features through each

DOI: 10.4018/IJDCF.350265

layer of the network and classified normal and abnormal behavior features using cross entropy loss function to achieve abnormal behavior detection. Xiao et al. (2002) used a spatiotemporal encoder to mine the spatiotemporal features of each frame in behavioral video images. They used attention mechanisms to weigh these features and used sigmoid functions and cross entropy loss functions to obtain the error between predicted and actual frames. Based on this result, they judged whether the current behavior was abnormal. Li et al. (2023) used the third version of the YOLO family of object detection models and regression methods to extract behavioral features, classified them through multitask learning, and completed abnormal behavior detection. Qian et al. (2020) established an abnormal behavior monitoring system using residual networks, obtained behavioral characteristics through convolutional kernel operations, and judged whether the behavior was abnormal based on the loss function value of each residual block. Although the existing technology for detecting abnormal behavior in school laboratories has made some progress, it still faces many challenges. Abnormal behavior detection requires a large amount of accurate and complete data to train and validate the model. In practical applications, data may have issues such as noise, missing components, or errors, which can affect the accuracy and reliability of the model. Liu et al. (2024) proposed a dangerous driving behavior prediction method based on the convolutional neural network-long short-term memory network and self-attention mechanism, which was used for historical driving data of trucks in a certain province. Through feature screening, spatial feature extraction, temporal information capture, and self-attention mechanism prediction, high prediction accuracy was achieved. However, the combination of the convolutional neural network-long short-term memory network and self-attention mechanism may lead to high model complexity, resulting in good performance on training data but poor generalization ability on unknown data. Shen et al. (2024) designed a composite network based on the rotating object detection model rRetinaNet and the convolutional recurrent neural network text recognition algorithm combined with attention mechanism to solve the problem of number plate recognition in athlete identity recognition with abnormal behavior, especially optimizing the number plate tilt distortion and small changes in aspect ratio. Although rRetinaNet can handle rotating targets, recognition performance may decrease for number plates with extreme angles or irregular shapes. Ren et al. (2024) designed an intelligent vision based auxiliary monitoring system for abnormal behavior of personnel in substations. By connecting hardware modules with a CC2530 wireless chip and combining artificial intelligence and image processing technology, they achieved preprocessing, background modeling, and information description of shadows in monitoring videos. They also used the histogram of oriented gradients feature recognition technology to detect abnormal behavior, improving the accuracy of monitoring recognition. In complex monitoring scenarios, there may be various interference factors, such as personnel occlusion, motion blur, etc., which may affect the accuracy of abnormal behavior detection.

Multimodal information fusion is a technology that processes multiple sources of information, aiming to integrate these different sources to obtain more comprehensive, accurate, and reliable information. In laboratory abnormal behavior detection, multimodal information fusion technology can integrate information from different data sources, such as video, audio, sensor data, etc., to achieve high-precision and efficient capturing and descriptions of complex behaviors and events in the laboratory environment, thereby more accurately detecting abnormal student behavior. To this end, a laboratory abnormal behavior detection based on multimodal information fusion is proposed.

RESEARCH METHOD

Laboratory Behavior Modal Information Mining

Acquisition of Student Behavioral Movement Modal Information

The optical flow field can capture the movement patterns of pixels or feature points in the laboratory student behavior video sequence, thereby reflecting the students' dynamic behavior (Bohan et al., 2023; Abhirami et al., 2022). Through the analysis of the optical flow field, students' behavior types, such as running, operating experimental instruments, data recording, instrument arrangement, jumping, etc., can be accurately identified (Wang et al., 2022; Zhou et al., 2022). To this end, the optical flow field is used to mine the motion modal information of students' behavioral acceleration and angular velocity.

First, the experimental student behavior red, green, blue (RGB) image sequence is obtained through the infrared sensor. Let $A(x, y, t)$ describe the gray value of point (x, y) on the student behavior RGB image sequence at time point t. $u(x, y)$ and $u(x, y)$ respectively describe the horizontal and vertical movement components of optical flow $b = (u, v)$ on (x, y). The solution process for the displacement of point (x, y) in the horizontal u and vertical v directions is as shown in Equation 1.

$$u = \frac{dx}{dt}, v = \frac{dy}{dt} \tag{1}$$

In Equation 1, d is the distance.

After a period Δt, $A(x, y, t) \rightarrow A(x + dx, y + dy, t + \Delta t)$. A remains unchanged when $\Delta t \rightarrow 0$, resulting in $A(x, y, t) = A(x + dx, y + dy, t + \Delta t)$. Then, it is rewritten according to Taylor's equation, ignoring the second order and infinitesimal terms, to obtain the mathematical expression for optical flow constraint, which is shown in Equation 2.

$$-\frac{\partial A}{\partial t} = \frac{\partial A}{\partial x} u + \frac{\partial A}{\partial y} v \rightarrow 0 = A_x u + A_y v + A_t \tag{2}$$

In Equation 2, ∂ is the derivative, and A_x, A_y, and A_t represent the gradients of A in the x, y, and t directions, respectively.

Based on Equation 2, a global smoothing constraint is added to avoid abrupt and discontinuous optical flow changes in the generated RGB optical flow map of laboratory student behavior conditions (Qi et al., 2023; Sun et al., 2022), thereby improving the accuracy and stability of optical flow calculation. The essence of this constraint is to calculate the sum of squares of the optical flow gradient mode, that is,$|\Delta u|^2 + |\Delta v|^2$. The smaller this value, the smaller the optical flow intensity, and the better the smoothing effect of the optical flow field in the sequence. The process of solving the global smoothing constraint term is as shown in Equation 3.

$$C_s = \iint (|\Delta u|^2 + |\Delta v|^2) dxdy = \iint \left(u_x^2 + u_y^2 + v_x^2 + v_y^2 \right) dxdy \tag{3}$$

According to Equation 2, solve the optical flow constraint term C_c, which is as shown in Equation 4.

$$C_c = \iint \left(A_x u + A_y v + A_t \right)^2 dxdy \tag{4}$$

According to Equations 3 and 4, the optical flow field should be minimized in accordance with C_s and C_c, and denoted by C_{sc}, then we have Equation 5.

$$\underset{min}{C_{sc}} = \min \iint \left[\beta \left(u_x^2 + u_y^2 + v_x^2 + v_y^2 \right) + \left(A_x u + A_y v + A_t \right)^2 \right] dxdy \tag{5}$$

In Equation 5, βdescribes the weight factor. If external noise has a significant impact on the behavior of RGB image sequences, the larger the value of β, the smoother the optical flow field of the sequence, indicating that the larger the value of β, the better the effect of suppressing the drastic changes in optical flow in the image. Conversely, the smoothness of the optical flow field is determined by C_c. At this point, it is required that β take a relatively small value for C_c to obtain a greater weight.

$C_{sc_{min}}$ is equivalent to C_{sc} solving for u and v separately, with a derivative of 0. The detailed operation process is as shown in Equation 6.

$$\begin{cases} A_x^2 u + A_x A_y v = -\beta \Delta u - A_x A_t \\ A_y^2 u + A_x A_y u = -\beta \Delta v - A_y A_t \end{cases} \tag{6}$$

Assume $\bar{u}$and $\bar{v}$represent the neighborhood means of u and v, then $\Delta u = u - \bar{u}$, $\Delta v = v - \bar{v}$, and rewrite Equation 6, then we have Equation 7.

$$\begin{cases} \left(A_x^2 + \beta^2\right) u + A_x A_y v = \beta^2 \bar{u} - A_x A_t \\ \left(A_y^2 + \beta^2\right) v + A_x A_y u = \beta^2 \bar{v} - A_y A_t \end{cases} \tag{7}$$

Using first-order difference to derive Equation 7, obtain the v and u of optical flow b, which are shown in Equation 8.

$$\begin{cases} u = \bar{u} - \dfrac{A_x\left(A_x \bar{u} + A_y \bar{v} + A_t\right)}{\beta^2 + A_x^2 + A_y^2} \\ v = \bar{v} - \dfrac{A_x\left(A_x \bar{v} + A_y \bar{v} + A_t\right)}{\beta^2 + A_x^2 + A_y^2} \end{cases} \tag{8}$$

According to Equation 8, obtain the dense optical flow field of the RGB image sequence of laboratory student behavior. Based on the motion information carried by the optical flow field, obtain the motion mode information of student behavior, namely acceleration information and angular velocity information.

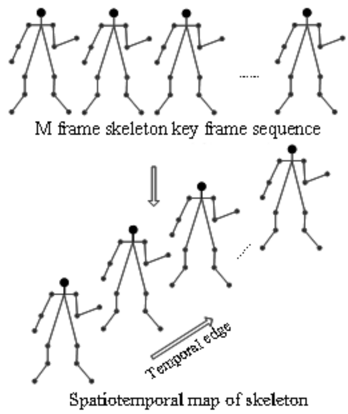
Multibehavior Contour Modal Information Mining

In laboratory behavioral abnormality detection, although the acceleration information, angular velocity information, etc. of student behavioral motion modal information are obtained, because the key points of the skeleton contain a large amount of human motion information, such as the skeleton body structure characteristics and the spatiotemporal relationship, which make up the student behavior contour mode, there is too much information. This can lead to misjudgments or missed judgments in abnormal behavior detection (Fei et al., 2023; Zhou et al., 2023). To this end, spatiotemporal graph convolution is used to further obtain student behavior profile modal information.

Firstly, a spatiotemporal graph of the human skeleton structure must be constructed. $G = (P, V)$, where $P = \left\{ p_{Tj} \middle| T = 1,2,...,M, j = 1,2,...,12 \right\}$ represents the set of key nodes in the skeleton, T represents the time dimension, j represents the number of key skeleton nodes in students, and V represents the set of images. It is composed of a single frame of the internal connection set of the human skeleton $V_z = \left\{ p_{Tj} p_{Tk} \middle|, (j,k) \in R \right\}$ (R is the set of inter joint connections) and a set of identical key point connections between adjacent frames $V_o = \left\{ p_{Tj} p_{(T+1)j} \right\}$. The spatiotemporal map of the human skeleton is shown in Figure 1.

After establishing the spatiotemporal map of the student skeleton, considering that human actions continuously change over time, using spatiotemporal convolutional networks to perform convolution operations on V_z and V_o in each frame of student behavior can effectively model the dynamic posture of students, identify the temporal and spatial relationships between different actions, and improve the accuracy and stability of action recognition. Mining the spatial and temporal motion characteristics of

Figure 1. Spatiotemporal map of human skeleton



skeleton sequences, namely behavioral contour information, involves the calculation process shown in Equation 9.

$$f_2 = e\left(MD^{-2}\left(I+B\right)D^{-2}f_1 W\right) \tag{9}$$

In Equation 9, I describes the adjacency matrix, which refers to the connections within the skeleton nodes of a single frame of student behavior. B represents the connection between laboratory video frames. D e refers to the degree matrix and activation function of the training network. f_1 describes the key points of student skeleton information. W is the learning factor. f_2 is the output result of spatiotemporal graph convolution, which is the modal information of the behavior profile of laboratory students. Each frame of the image is convolved according to Equation 9 to extract the modal information of student behavior contours.

Identification of Laboratory Abnormal Behavior Based on Multimodal Information Fusion

There are many methods for multimodal information fusion, such as data level, feature level, and decision level. Since the motion mode information and contour mode information of laboratory students belong to two different spatial dimensions, if the data level and feature level are directly collected to fuse and reduce the dimensionality of each modal feature, the differences in feature information and intermodal interference will be ignored. This will to some extent affect the accuracy of abnormal behavior detection in experiments. Decision level fusion has strong anti-interference ability and good fault tolerance (Ning et al., 2023; Ma et al., 2022). Therefore, decision level is selected to achieve the fusion of motion mode information and contour mode information of laboratory behavior, thereby improving detection accuracy. The fusion idea is to first solve the conditional probability of all modal information belonging to a certain category through a classifier (Qin et al., 2022), combine certain rules to fuse the metric information of each modality, and then select the category corresponding to the maximum value from the fusion results as the result of laboratory behavior detection.

The closeness function considers the connections and correlations between samples when measuring the similarity between categories of laboratory behavior, which can better capture the relationships between different modal information in laboratory behavior. By measuring the distance between abnormal behavior and normal behavior (Araya et al., 2022; Bayrami et al., 2022), it is possible to more clearly distinguish between the two types of behavior, provide more comprehensive feature expression, and enhance the discriminative ability of laboratory behavior categories. The calculation process is shown in Equation 10.

$$\gamma(l,r) = \exp[-H(l,r)] \tag{10}$$

In Equation 10, $H(l,r)$ describes the elements in the matching cost matrix. In the spatial domain, $H(l,r)$ describes the degree of matching between the experimental behavior sequence l frame and the template column r frame. In the time domain, $H(l,r)$ describes the degree of matching of shape changes between adjacent frames in the experimental behavior sequence and adjacent frames in the template sequence. Overall, $H(l,r)$ focuses on shape matching between single frames in the spatial domain; while in the temporal domain, it focuses on shape change matching between consecutive frames. The probability calculation process for experimental behavior detection is presented in Equation 11.

$$P^* = w_1\gamma_1(l,r) + w_2\gamma_2(l,r) \tag{11}$$

In Equation 11, $\gamma_1(l,r)$ and $\gamma_2(l,r)$ respectively describe the matching distance between the experimental behavior motion mode and the contour mode. w_1 and w_2 are the weighting factors of motion mode and contour mode, which are the proportion of the final representation of behavior. The solution process is as shown in Equation 12.

$$w_1 = \frac{p_1^*}{\sum_{j-1}^{2} p_j^*}, w_2 = \frac{p_2^*}{\sum_{j-1}^{2} p_j^*} \tag{12}$$

In Equation 12, p_1^* and p_2^* respectively represent the estimated probabilities of the motion mode and contour mode of the current behavior occurring in a certain category, which is the conditional probability that each modal information belongs to a certain category.

Different proportions of allocation result in different behavior recognition results, and the proportion allocation corresponding to each behavior category is also different. Therefore, a weighted average algorithm is used to solve for its proportion allocation value, which is shown in Equation 13.

$$w_3 = \frac{p_j^*}{\sum_k^2 p_k^*} \tag{13}$$

After extensive experimental verification, it has been found that the optimal detection effect for abnormal behavior is achieved when the matching threshold is 0.9 for multimodal fusion of experimental behavior sequences. If both the motion mode information and the contour mode information belong to a certain behavior (normal behavior category or abnormal behavior category), the maximum conditional probability in the fusion result is selected. If it is greater than 0.9, it is considered that this behavior belongs to the normal behavior category/abnormal behavior category. Conversely, if it is different, the experimental abnormal behavior detection task is ultimately completed.

In summary, to achieve abnormal behavior detection in the laboratory, this article first uses an optical flow field to mine the motion mode information of student behavior acceleration and angular velocity. Firstly, obtain the RGB image sequence of experimental student behavior through infrared sensors, and based on optical flow constraints, ensure the continuity and smoothness of the entire optical flow field in space. Add global smoothing constraints and calculate the dense optical flow field of the RGB image sequence of laboratory student behavior, that is, obtain the student behavior motion mode information, namely acceleration information and angular velocity information, based on the motion information carried by the optical flow field. At the same time, a spatiotemporal graph of the human skeleton structure is constructed, and a spatiotemporal convolutional network is used to perform convolution operations on the internal connections of a single frame of the human skeleton within each frame of student behavior and the same key point connections between adjacent frames, mining the spatial and temporal motion characteristics of the skeleton sequence. This is the behavioral contour information. Based on obtaining the movement and contour information of laboratory students, a decision level is selected for information fusion. In the fusion process, the conditional probability of each modality belonging to a certain category is solved by a classifier, and the metric information of each modality is fused according to certain rules. Then, the maximum value corresponding to the category is selected from the fusion results as the result of laboratory behavior detection. As a result, it is possible to capture the characteristics of student laboratory behavior more comprehensively and accurately and obtain reliable laboratory behavior detection results, providing more effective support for laboratory safety management and monitoring.

In summary, the laboratory abnormal behavior detection process based on multimodal information fusion is shown in Figure 2.

Experimental Analysis

Experimental Setup

The experimental site is a laboratory equipped with professional measuring equipment and monitoring systems. The laboratory space is spacious enough to accommodate 40 students participating in the experiment simultaneously, and ensures that each student's movements and postures can be monitored and recorded in real-time by the camera. At the same time, the temperature and humidity in the laboratory are controlled within an appropriate range (20°C~24°C, 40%~60%), ensuring that students conduct experiments in a comfortable environment and avoiding interference from environmental factors on the experimental results. The laboratory has sufficient and uniform lighting, ensuring that the camera can clearly capture the movements and postures of students, while avoiding measurement errors caused by insufficient lighting. In addition, the laboratory adopts sound insulation design and strictly controls the internal noise level to ensure that it is not disturbed by external noise during the experimental process.

The laboratory has installed cameras to monitor and record the movements and postures of students in real-time. The camera system layout is reasonable, ensuring comprehensive coverage of the experimental area and capturing the details of each student's movements. Accelerometers and gyroscope sensors are installed around students to collect student behavior data. The relevant instrument parameters are shown in Table 1.

The equipment referenced in Table 1 are common instruments in real life and can be deployed in laboratories of different scales, while ensuring the stability and reliability of the system.

After completing the above hardware settings, install multimodal information fusion software on the central server and configure the proposed algorithm. Set up a data storage and backup system to ensure the security and accessibility of data. Configure network devices, including routers, switches, etc., to ensure real-time transmission of video streams and sensor data. Integrate the multimodal information fusion system with existing laboratory security systems to ensure information sharing and collaborative work, thereby completing algorithm operation and achieving abnormal behavior detection in the laboratory.

Experimental Dataset

Considering that cross validation can more accurately evaluate the performance of the model, ensuring the reliability and credibility of experimental results, this approach not only helps to discover the performance differences of the model on different datasets, thereby evaluating its generalization ability, but also improves the rigor and standardization of the research, ensuring its scientific and credibility. Therefore, divide the dataset into multiple groups.

A total of 40 students participated in this experiment, with 20 male and 20 female students each. To ensure the accuracy of the experimental results, the height and weight of these students were measured, and the range of their height and weight was found to be 155 cm -189 cm and 45.35 kg - 88.12 kg, respectively. All students are required to perform relevant operations in the assigned

Figure 2. Laboratory abnormal behavior detection process

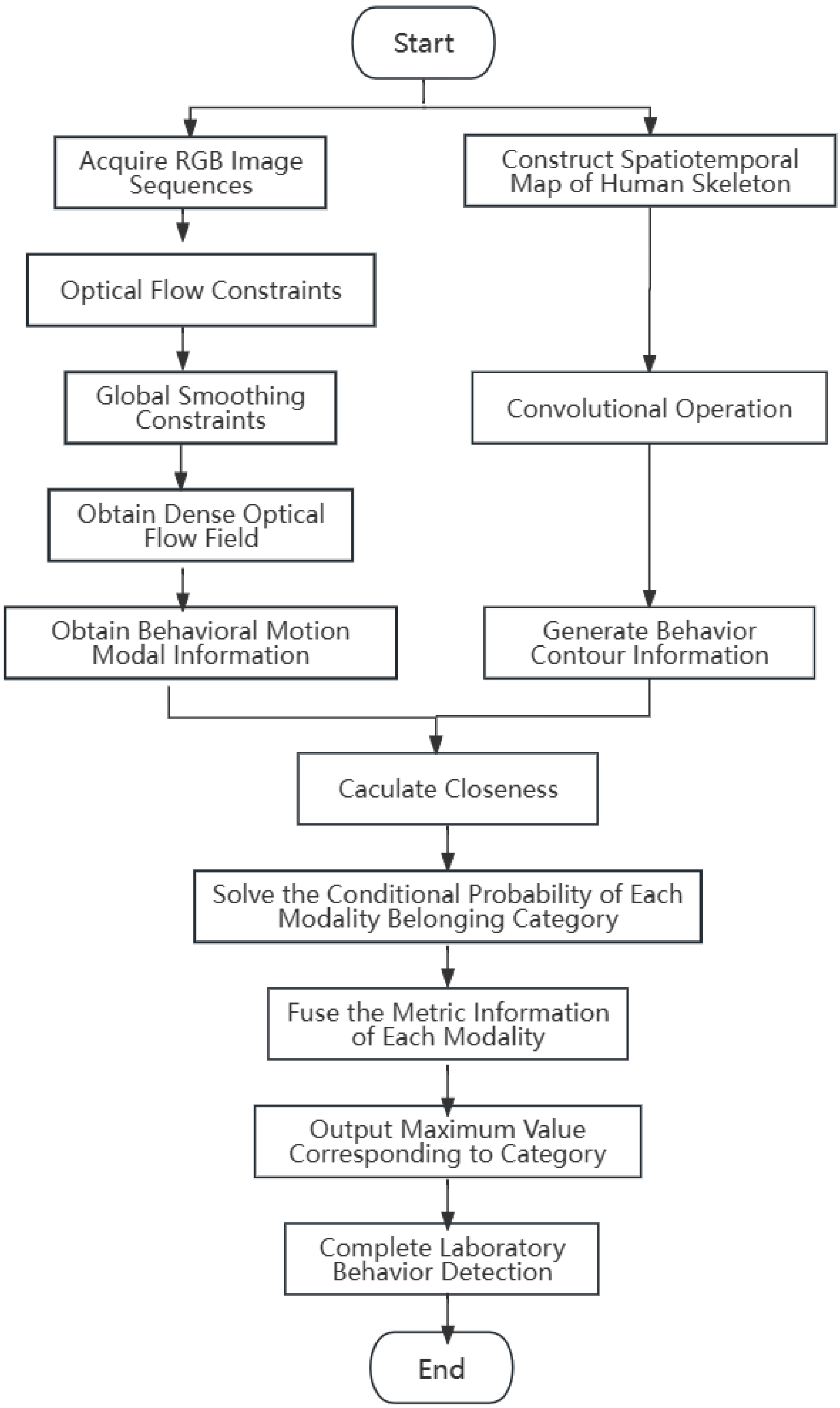


Table 1. Behavioral information collection instrument related parameters

Instrument Name	Related Parameters	Parameter Details
CS-3LAS-03three-axis Accelerometer	Supply voltage	12 ± 0.5 VDC
	Supply current	≤ 12 mA
	Resolution ratio	≤0.01 g
	Bandwidth	≤10 kHz
MPU9250gyroscope	Magnetic induction Intensity measurement	±4800μT
	Communication rate	0-400kHz/s
	Working voltage	2.4V-3.6V(±5%)
	Operation temperature	-40°C~+85°C

Figure 3. Abnormal behavior detection results

(a) Before Information Fusion

(b) After Information Fusion

laboratory according to the course requirements, and a video sequence of laboratory behavior monitoring is collected based on the environmental arrangement in the experimental setup section.

Analysis of the Effectiveness of Laboratory Abnormal Behavior Detection

According to the video sequence of laboratory behavior monitoring, there are 10 behavioral postures, namely standing, operating experimental instruments, measuring, data recording, instrument organization, instrument cleaning, rapid movement, jumping, bending and walking, and playing, in this experiment. Among these behavioral postures, rapid movement, jumping, bending and walking, and playing are abnormal behaviors in the laboratory. Select 1000 data for each behavioral posture, totaling 10000 row posture data. Mix the 10000 behavioral posture data and randomly generate 10 datasets as experimental research objects. To verify the laboratory abnormal behavior detection results obtained by the proposed technology, the detection results before and after information fusion were compared, as shown in Figure 3.

Table 2. Results of ablation experiment

Data Set	Multimodal Information Fusion (%)	No Modal Information Mining (%)	No Information Fusion (%)
1	99.6	79.4	61.9
2	99.5	75.3	61.6
3	99.2	71.2	61.3
4	99.0	69.8	61.0
5	98.7	67.8	59.8
6	98.6	62.5	59.5
7	98.5	61.9	59.2
8	98.4	58.9	58.9
9	98.2	58.3	58.6
10	98.0	57.7	58.0

From Figure 3, before information fusion, both normal and abnormal laboratory behaviors were detected. However, after the application of the proposed method, two types of abnormal behaviors, namely standing abnormality and operating the experimental instrument abnormality, can be accurately detected from left to right. This detection indicates that by integrating the information of student behavior movement modes and behavior contours, it is possible to more accurately detect abnormal behaviors of students in the laboratory.

The ablation experiment refers to gradually reducing a portion of abnormal behavior detection in a laboratory, observing changes in system performance, and thus understanding the contribution of that portion to the entire system. The experimental results are shown in Table 2.

Analyzing Table 2, neither laboratory behavior mode information mining nor information fusion can to some extent reduce the detection accuracy and affect the final detection results.

Comparative Experimental Analysis

To further verify the effectiveness of multimodal information fusion in detecting abnormal behavior, the experiment measured the detection error from the perspective of abnormal behavior. The experimental dataset remained unchanged, and the YOLO network in reference Zhang et al. (2023) and the attention mechanism in reference Xiao et al. (2022) were used as comparative methods to compare and analyze the detection errors of abnormal behavior in 10 datasets of the three methods. The results are shown in Figure 4.

In Figure 4, the multimodal information fusion method has 10 errors in dataset 2 (measurement), which is the largest error data in all datasets. Therefore, detecting complex behaviors may introduce more errors. The error count in dataset 6 (instrument cleaning), dataset 8 (jumping), and dataset 9 (bending and walking) is 1, which is the least error data among all datasets. Different experimental behaviors have different complexities and dynamic characteristics. For example, measurement behavior may involve fine hand movements and prolonged static retention, which may not be as obvious or easily recognizable in optical flow fields and contour extraction as other behaviors, such as jumping, bending, and walking. Therefore, the detection of complex behaviors may introduce more errors. However, overall, the number of abnormal behavior detection errors in each dataset of multimodal information fusion methods is always less than that in YOLO networks and attention mechanism methods, because they use optical flow method and spatiotemporal graph convolution to mine the motion mode information and contour mode information of various behavioral postures of students. By fusing this complementary information, more comprehensive and accurate detection results can

Figure 4. Comparison of detection errors

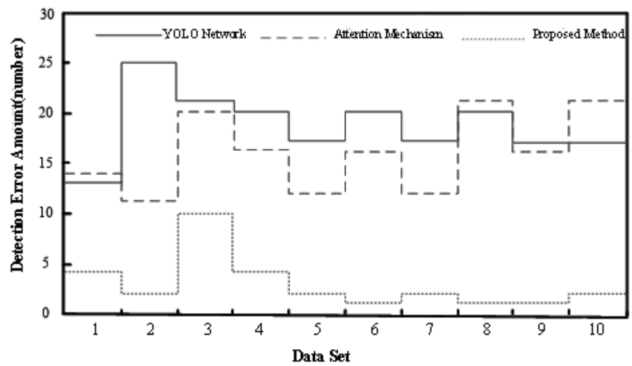
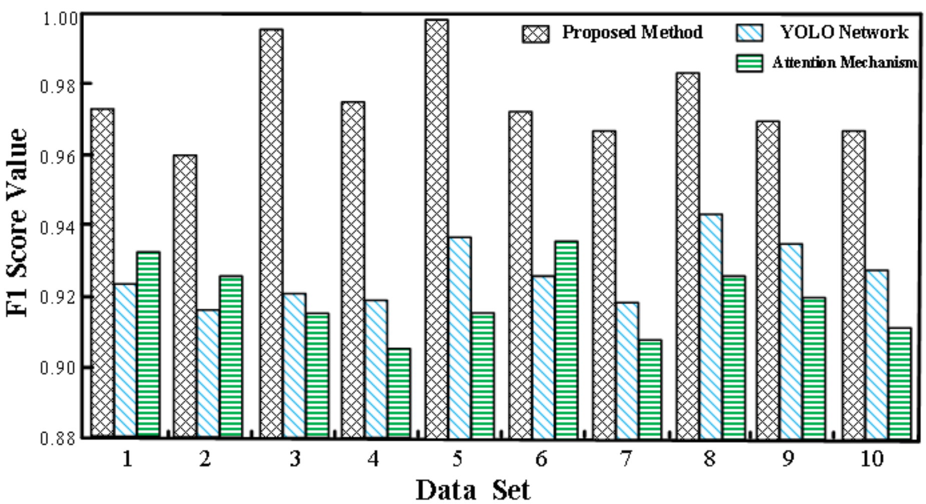


Figure 5. F1 score results



be obtained. However, comparing the two methods, they can only capture motion mode information, obtain fewer features, and have a higher number of errors in detecting abnormal behavior.

F1 score is one of the key indicators for detecting abnormal behavior among laboratory students. To quantify the detection performance of the method, the proposed method, the YOLO network in reference Zhang et al. (2023) and the attention mechanism in reference Xiao et al. (2022) were used to perform anomaly detection on these behavioral datasets, and the F1 score values of each data set were analyzed, as detailed in Figure 5.

A shown in Figure 5, the F1 scores for each dataset classification of the multimodal information fusion method are all greater than 0.96. After applying YOLO network and attention mechanism methods to detect abnormal behavior of laboratory students, the F1 scores obtained fluctuate between 0.90 and 0.94, which is relatively high but lower compared to multimodal information fusion methods. In summary, the multimodal information fusion has a good effect on student abnormal behavior detection. This is because the proposed method utilizes decision level methods to fuse behavioral motion modal information with contour modal information, improving the ability to represent

Table 3. Accuracy test results

Data Volume (number)	Proposed Method (%)	YOLO Network (%)	Attention Mechanism (%)
100	98.7	94.2	92.1
200	98.5	93.4	91.6
300	98.1	92.6	91.2
400	97.6	91.9	90.6
500	97.3	90.7	89.2
600	95.1	90.1	88.4
700	94.5	89.6	87.9
800	93.5	88.4	86.7
900	92.2	87.5	86.4
1000	91.0	87.1	85.3

behavioral posture features, obtaining richer and more comprehensive feature representations, improving behavior classification performance and accuracy, and ultimately obtaining a higher F1 score.

To further verify the detection performance of the proposed method, images from the University of California, San Diego (UCSD) anomaly detection dataset were used as experimental objects. The method proposed in this article, the YOLO network Zhang et al. (2023) and the attention mechanism (Xiao et al. 2022) were tested, using accuracy as the testing indicator. The obtained experimental results are shown in Table 3.

According to Table 3, the highest accuracy values of the YOLO network and attention mechanism methods are 94.2% and 92.1%, respectively. Although they have achieved high accuracy in detecting abnormal behavior, there is still room for improvement. The accuracy of multimodal information fusion in detecting abnormal behavior of images in the UCSD anomaly detection dataset is above 91.0%, with a maximum value of 98.7%, indicating that the proposed method has good practicality in abnormal behavior detection.

Specificity refers to the percentage of samples without abnormal behavior that are correctly identified as non-abnormal behavior by the model, that is, the calculation process is as shown in Equation 14. Specificity can measure the ability of the detection model to correctly identify non-abnormal behavior. For this reason, specificity is selected to evaluate the abnormal behavior detection performance of the laboratory.

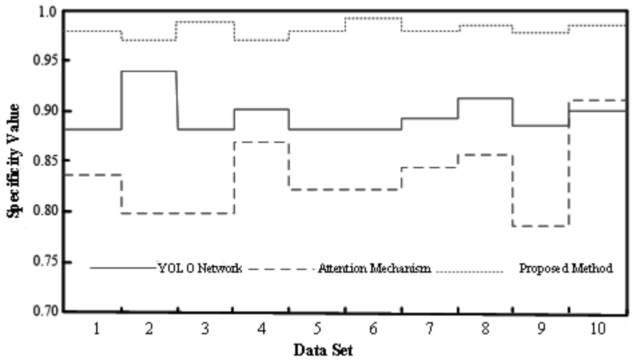
$$S = \frac{TN}{TN + FP} \times \% \tag{14}$$

In Equation 14, TN and FP are true negative and false positive.

The higher the specificity, the better the performance of the detection model, which can effectively eliminate non-abnormal behavior and reduce false positives. The experimental abnormal behavior dataset remains unchanged, and the specificity of multimodal information fusion, YOLO network, and attention mechanism abnormal behavior detection is compared and analyzed. The results are shown in Figure 6.

In Figure 6, the average specificity of abnormal behavior detection in each dataset using the YOLO network and attention mechanism methods is around 0.92 and 0.85. The specificity of abnormal behavior detection in various datasets of multimodal information fusion is extremely close to 1.00. Due to its use of probability functions and weighted average algorithms to adjust the matching weights

Figure 6. Comparison of specificity



of information fusion appropriately and make the optimal decision, the abnormal behavior detection performance is superior to the comparison of the two methods.

Based on the above comparative experiments, it can be concluded that the multimodal information fusion method performs well in detecting abnormal behavior among laboratory students. This method outperforms the YOLO network and attention mechanism methods in key performance indicators such as F1 score, accuracy, and specificity. Specifically, the F1 score of the multimodal information fusion method is higher than 0.96, with an accuracy of 98.7% on the UCSD anomaly detection dataset, and a specificity close to 1.00, demonstrating its significant advantages in classification accuracy, accuracy in anomaly behavior detection, and correct recognition of normal behavior. Although some errors may be introduced when detecting complex behaviors, overall, the number of errors in multimodal information fusion methods is less than that of comparative methods, demonstrating their robustness and efficiency in handling laboratory anomaly detection tasks.

DISCUSSION

Based on the above experimental verification results, the method mentioned in this article integrates student behavior motion mode information and behavior contour information in the detection process, which can more effectively detect abnormal behavior of laboratory students. By integrating multimodal information methods, mining behavior and posture information can obtain a richer amount of abnormal behavior features compared to traditional methods, thereby achieving higher detection accuracy. The mining of modal information related to behavior detection and the fusion of multimodal information directly constrain the results of abnormal behavior detection. Multimodal information is processed using decision level methods to obtain richer and more comprehensive feature representations, optimize the accuracy and recall of detection models, effectively eliminate non-abnormal behaviors, reduce false alarm rates, and better improve model detection performance.

The proposed laboratory abnormal behavior detection method needs to comprehensively consider performance indicators such as accuracy, speed, and computational complexity when facing the practical application needs of daily laboratory behavior management. In the subsequent algorithm research, the algorithm performance will be further improved. While balancing various indicator parameters, the design cost will be further considered, and the lightweight model algorithm will make it easier to deploy on embedded mobile terminals. Due to limitations in equipment conditions, in addition to using public datasets to validate algorithms, self-built datasets contain fewer types of abnormal behaviors. In real laboratory operating environments, the on-site situation is complex, and abnormal behaviors are diverse. Therefore, future research will collect more abnormal behaviors in daily laboratory operating environments, enrich self-built datasets, conduct in-depth detection experiments, and prove the effectiveness of the proposed method.

CONCLUSION

To help teachers and administrators detect student misconduct in a timely manner, intervene and educate in a timely manner, and avoid accidents, a laboratory abnormal behavior detection based on multimodal information fusion is proposed. The research methods involved using the optical flow method and global smoothing constraint conditions to obtain student behavior motion characteristics, using the spatiotemporal graph convolution to capture the contour features of behavior, and using the decision level fusion method to achieve the fusion of various modal features. The experimental results show that the multimodal information fusion method can effectively improve the accuracy and reliability of abnormal behavior detection and reduce detection errors. Multimodal information fusion requires processing a large amount of data, which may affect the real-time performance of detection methods. How to improve real-time performance while ensuring accuracy is a challenge that this method needs to face in practical applications. In the future, this method can be further optimized by introducing more advanced fusion strategies and algorithms, such as deep learning models, graph neural networks, etc., to more effectively fuse multimodal information and improve real-time performance. Further improvements can be made to feature extraction and classification algorithms, optimizing algorithm parameters, and enhancing the performance and efficiency of abnormal behavior detection. At the same time, in-depth research is being conducted in the applicable fields of multimodal information fusion methods, such as intelligent monitoring, robot vision, etc.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this article.

FUNDING STATEMENT

This work is supported by Liaoning Provincial Department of Education's Basic Research Project for Universities [Grant Number: JYTMS20230711], Liaoning Province Science and Technology Plan Joint Program (Fund) Project [Grant Number: 2023JH2/101700009] and Dandong City Guiding Science and Technology Plan (Liaodong University Joint Technology) Project [Grant Number: DDZDKJ202204].

PROCESS DATES

July 15, 2024
Received: March 25, 2024, Revision: June 26, 2024, Accepted: June 27, 2024

CORRESPONDING AUTHOR

Correspondence should be addressed to Dawei Zhang (China, zdw@elnu.edu.cn)

REFERENCES

Abhirami, K., & Devi, M. K. (2022). Student behavior modeling for an e-learning system offering personalized learning experiences. *Computer Systems Science and Engineering*, *2022*(40), 1127–1144. 10.32604/csse.2022.020013

Araya, I., Beas, V., Stamulis, K., & Allede-Cid, H. (2022). Predicting student performance in computing courses based on programming behavior. *Computer Applications in Engineering Education*, *2022*(30), 1264–1276. 10.1002/cae.22519

Bayrami, R., Ebrahimi, S., Rasouli, J., & Feizipour, H. (2022). Knowledge, attitude, and behavior in avoiding environmental tobacco smoke exposure at home among pregnant women. *Current Women's Health Reviews*, *2022*(18), 80–86.

Bohan, C., & Smyth, S. (2023). The effect of schedule thinning on student behavior during the caught being good game. *Behavior Modification*, *2023*(47), 644–669. 10.1177/01454455221129993363734 36

Fei, S. M., Zhao, H. T., Yang, Y., & Li, C. F. (2023). Temporal topology unshared graph convolution and multi-scale temporal convolution for skeleton-based action recognition. *Information and Control*, *2023*(52), 758–772.

Li, J., Xie, H. Z., & Wang, G. S. (2022). Human abnormal behavior recognition algorithm based on pose estimation. *Journal of Beijing University of Technology*, *2022*(48), 710–720.

Li, Z. H., Zhang, J., We, Y. J., Dong, Y., & Xu, W. T. (2023). A multiscale fusion YOLOV3-based model for human abnormal behavior detection in special scenarios. *Journal of Transportation Engineering, Part A: Systems*, 2023(149), 4022150.1-4022150.12.

Liu, P. F. (2024). LULu, J. G., Xu, L., Tang, X. H., & Liu F. J. (2024). Research on intelligent prediction technology of dangerous driving behavior in highway freight. *Automobile Technology*, (3), 56–62.

Ma, Y. T., Wang, S., & Liu, Y. F. (2022). research on human action recognition method by fusing multimodal data. *Computer Engineering*, *2022*(48), 180–188.

Ning, D. H., & Zheng, S. (2023). An object detection algorithm based on decision-level fusion of visible and infrared images. *Infrared Technology*, *2023*(45), 282–291.

Qi, M., Xu, H., Li, S., Zhang, Y., & Sun, H. (2023). An action recognition method based on two-stream network. [Science Edition]. *Journal of Jilin University*, *2023*(61), 347–352.

Qian, H. F., Zhou, X., & Zheng, M. M. (2020). Abnormal behavior detection and recognition method based on improved resnet model. *Computers, Materials & Continua*, *2020*(65), 2153–2167. 10.32604/cmc.2020.011843

Qin, B. X., Lu, H., Qiu, C., & Wan, W. M. (2022). A pedestrian abnormal behavior detection based on motion analysis. *Telecommunication Engineering*, *2022*(62), 457–465.

Ren, Y. L., Liu, Y. M., Jie, L. B., & Jing, Y. Z. (2024). Auxiliary monitoring system for abnormal behavior of substation personnel based on intelligent video. *Techniques of Automation and Applications*, *66-70*(43), 107.

Shen, J. B., Song, S. Y., Huo, S., & Li, J. H. (2024). Research on marathon directory recognition method based on deep learning. *Journal of Suihua University*, *2024*(44), 144–148.

Sun, R. X., Zhu, G. L., Xie, S. Y., Guo, X. L., & Cai, Z. L. (2022). High-speed computing of pyramid LK optical flow based on embedded GPU. *Jisuanji Yingyong Yanjiu*, *2022*(39), 1966–1972.

Wang, Z. J., Shen, C. M., Zhao, C., Liu, X. M., & Chen, J. (2022). Recognition of classroom learning behaviors based on the fusion of human pose estimation and object detection. [Natural Science]. *Journal of East China Normal University*, *2022*(2), 55–66.

Xiao, J. S., Shen, M. Y., Jiang, M. G., Lei, J. F., & Bao, Z. Y. (2022). Abnormal behavior detection algorithm with video-bag attention mechanism in surveillance video. *Acta Automatica Sinica*, *2022*(48), 2951–2959.

Zhang, H. M., Zhuang, X., Zheng, J. T., & Fang, X. B. (2023). Optimizing human abnormal behavior detection method of YOLO network. *Computer Engineering and Applications*, *2023*(59), 242–249.

Zhang, X. P., Ji, J. H., & Wang, L. (2022). Overview of video based human abnormal behavior recognition and detection methods. *Control and Decision*, *2022*(37), 14–27.

Zhou, H., Liu, Y. X., Gong, Y., Kou, F. W., & Xu, G. L. (2022). Action recognition algorithm based on dense trajectories and optical flow binarization image. *Computer Engineering and Applications*, *2022*(58), 174–180.

Zhou, X. H., & Wu, W. B. (2023). driving behavior recognition based on improved densenet. *Jisuanji Fangzhen*, *2023*(40), 197–202.

Professor DaWei Zhang was born in Dandong, Liaoning, P.R. China, in 1978. He received his master's degree from Shenyang Aerospace University, P.R. China. Now, he works in the School of Information Engineering at Liaodong University. His research interests include electronic information engineering, embedded system, and computational intelligence.