Sensors and Materials, Vol. 35, No. 12 (2023) 4251–4267
MYU Tokyo

S & M 3479

Research and Implementation of Intelligent Learning Desk Based on Visio Sensor in AI IoT Environments for Smart Education

Yanjun Zhu,* Shi Wang, Yong Peng, Yuqiang Chen, Rongli Chen, Gang Ke, and Jianxin Li

School of Electronic Information, Dongguan Polytechnic, Dongguan, 523808, China

(Received February 1, 2023; accepted November 21, 2023)

Keywords: visio sensor, AI, IoT, smart education, sitting posture detection, intelligent learning desk

Nowadays, the number of applications of IoT and AI has increased rapidly to provide personalized learning environments that provide control to learners. To promote good sitting posture and support the health of primary and middle school students, intelligent learning desks with visio sensors, which can be used to evaluate the health-related effects of sitting posture based on three dimensions, namely, human posture, critical threshold, and abnormal posture duration, are proposed. In accordance with the joint point model obtained from the OpenPose algorithm, we identified five abnormal sitting postures, namely, head tilt, body tilt, head lowering, reading at a close distance, and sitting for a long duration. Using this information, we designed a detection process for these postures to improve the accuracy of posture evaluation in our intelligent learning desks. We optimized the OpenPose model for mobile terminals by utilizing deep separable convolution to replace some convolution cores in the two-branch multistage network. This approach effectively reduced the amount of network structure parameters and significantly decreased the computational load required for the model. As a result of this optimization, we were able to more than double the video recognition speed compared with the original model. This improvement enables our intelligent learning desks to operate with greater efficiency on mobile devices without sacrificing accuracy or performance. According to our experiments and practical tests, our system can effectively monitor and warn students of common abnormal sitting postures. The recognition rate of abnormal sitting postures, such as prolonged learning, head tilt, body tilt, and head bow, has been optimized to over 92%. This high level of accuracy enables our intelligent learning desks to provide timely feedback and alerts to students when they exhibit poor posture habits, which can help prevent long-term health issues associated with prolonged sitting or incorrect posture.

*Corresponding author: e-mail: 26307320@qq.com
https://doi.org/10.18494/SAM4339

ISSN 0914-4935

1. Introduction

Maintaining a good learning sitting posture is crucial for the growth and development of teenagers. According to statistical data, students spend around 7 h of their day studying, and 75% of them have incorrect sitting posture during this time. This has led to over 90% of students developing physical illnesses associated with unhealthy sitting posture, such as hunchbacks, myopia, scoliosis, and oblique shoulders. Although teachers and parents attempt to pay attention to students' sitting posture, it is challenging to achieve continuous supervision and reminders. To address this problem, we propose utilizing intelligent learning desks based on visio sensors to achieve real-time detection of students' postures. This approach allows for the continuous monitoring and feedback of students' posture habits, while also allowing for personalized suggestions and adjustments to be made. Our system uses a visio sensor installed on the intelligent desk to collect data on students' learning sitting posture and then employs human posture estimation to monitor their postures in real time. When the system detects an abnormal sitting posture, it provides timely reminders through various modalities including voice, text, screen flicker, and vibration signals. By doing so, it can effectively prevent harm associated with abnormal sitting posture and promote healthy growth and development of teenagers.

2. Related Work

2.1 Artificial Intelligence Internet of Things

Artificial Intelligence Internet of Things (AIoT) is a combination of AI and IoT technologies. It allows for the collection and analysis of vast amounts of data from a wide range of sources across various dimensions. This data is then stored in both the cloud and edge, utilizing the capabilities of IoT,[1–4] and subsequently realizes the digitalization and intelligent connection of everything through big data analysis and higher forms of AI. IoT technology and AI integration aims to create a cohesive intelligent ecosystem that amalgamates various intelligent terminal devices, system platforms, and application scenarios. Achieving this requires constant technological innovation, development of AIoT technical and testing standards, efficient implementation of relevant technologies, and the widespread promotion of typical use cases. These critical issues need to be tackled effectively within the IoT and AI sectors.

As IoT and AI technologies continue to mature, an increasing number of enterprises have prioritized AIoT as their primary development direction. The term "AIoT" has been attracting significant attention within the IoT industry since 2017. It refers to the integration of AI and IoT technologies in practical applications, enabling the creation of intelligent systems that can provide deeper insights into data analysis and decision-making processes. This integration has opened up new possibilities for advanced automation, predictive maintenance, and real-time monitoring, facilitating greater efficiency and productivity across a wide range of industries[5–8]. An increasing number of industrial applications are incorporating AI into IoT technology. Several manufacturers, including Xiaomi, Skyworth, and Hisense, have introduced their own AIoT televisions, highlighting the growing demand for intelligent devices that can connect and communicate with each other seamlessly. These smart TVs leverage AI algorithms to enable device control, voice recognition, facial recognition, and personalized recommendations, providing users with a more interactive and intuitive viewing experience. The integration of AI and IoT is transforming traditional devices into "smart" systems, facilitating the development of new business models and services across various industries.

2.2 Smart education

Education has undergone a significant transformation in recent years, extending the learning process beyond traditional classrooms. Nowadays, learners have access to vast digital libraries, online courses, and the ability to submit assignments electronically from any mobile device. Educational institutions are increasingly offering mobile services that enable students to attend classes remotely via videoconferencing and live streaming, providing an immersive and engaging learning experience that can be tailored to fit their lifestyles and schedules. This shift towards mobile education is empowering individuals to take control of their learning journeys and pursue their educational goals more flexibly than ever before.

Smart education is a tech-driven approach to improve different aspects of education, including management, teaching methodologies, and research. It aims to provide learners with personalized, interactive learning experiences and help educators optimize their teaching strategies and resource management using technologies such as AI, big data, and cloud computing.(9–10) It is not only about making education digital but also involves applying emerging technologies to school management, teaching methodologies, and other aspects to improve education quality and equity on a larger scale. Through this approach, smart education can create a new ecological model for education that aims to promote better results and outcomes for both learners and educators alike.

2.3 Human posture estimation based on visio sensor

Human posture estimation typically involves the use of sensors to collect and monitor data on a person's posture. This information is then analyzed and processed to determine their sitting posture status. In some cases, vision sensors are used to obtain images of the person, and image processing technologies are applied to identify the key characteristics of the human body in the images to accurately determine their sitting position. Lan *et al.* proposed a method for analyzing a user's sitting posture by comparing the proportion of their face in video footage and the spatial relationship between specific points on their body.(11) Tariq *et al.* developed an approach for classifying human sitting posture using image processing techniques, which they enhanced by combining image information with motion data collected from smartwatches in an IoT environment.(12) Kumara *et al.* improved on the state-of-the-art gait analysis by developing novel deep-learning-based algorithms designed to identify occluded frames in a gait sequence and use spatiotemporal information to reconstruct them. This approach involved leveraging the power of deep learning techniques to detect and fill in missing frames caused by occlusion, resulting in more accurate and complete gait analyses.(13) Chen *et al.* proposed an image authentication method based on the residual histogram shifting technique.[14] Hsia *et al.* proposed a method that uses the time-of-flight (ToF) for an assisting device.[15] Pai *et al.* designed a control interface for dual-input video/audio recognition consisting of two input interface systems, hand posture, and speech recognition, with the use of specific hand postures or voice commands for control without the need for wearable devices.[16] Qu *et al.* proposed a human fall detection algorithm that combines human posture, support vector machine (SVM), and quadratic threshold decision.[17]

3. Feature Extraction of Abnormal Sitting Posture

Human posture estimation refers to the method of estimating the relationship between bone joint points in the human body and reconstructing the limbs and trunk by detecting the position information of these joint points. It is widely used in advanced applications such as human behavior recognition, posture tracking, character image generation, and human–computer interaction. Human posture estimation methods can be divided into traditional methods and deep-learning-based methods. Traditional methods rely on manual feature annotation, treat posture estimation as a regression problem, and directly return the coordinates of relevant nodes, which may result in lower accuracy. In contrast, deep-learning-based methods, especially convolutional neural networks (CNNs) such as the hourglass model and its variants, have shown robust performance in human posture estimation tasks. DeepPose was one of the first methods to apply deep learning to human posture estimation, while convolutional posture machines (CPMs) have greater accuracy for predicting joint point positions. CPM studies the relationship between human joint points, learns the expression of spatial information through convolution networks, and uses different receptive field sizes to deal with variations in key parts of the human body in images.[18] To address the issue of variation in joint point size, Yang *et al.* proposed the Pyramid Residual Module (PRM), a novel module for obtaining multiscale feature information from images. The PRM performs feature extraction using multiple branches with different scale sizes, each of which obtains a feature map with different sizes through downsampling. This enables the module to capture the features of different scales and improve the accuracy of joint point prediction in human posture estimation.[19] Tang and Wu proposed the Deeply Learned Composite Model (DLCM) to address low-level ambiguity in target image recognition. The DLCM utilizes a joint point connection mode representation method, which contains rich feature information and demonstrates efficient performance in joint point estimation. The use of this approach may improve the design of human joint point data in future studies, leading to enhanced accuracy and performance in human posture estimation tasks.[20] Feng *et al.* proposed a solution to the issue that most current research on human posture estimation is focused on enhancing network model generalization and overlooks efficiency concerns, by introducing fast posture interpretation (FPD). This approach prioritizes training lightweight human posture estimation network models, optimizing for both model accuracy and computational efficiency.[21] Taking into account the characteristics of hyperactive students and the need to deploy applications on intelligent desks with low configuration tablet PCs, researchers use OpenPose to analyze human joint point data, focusing specifically on sitting postures. They analyze the characteristics of different sitting postures and use duration analysis to determine whether a given posture represents an abnormal learning posture. This system provides real-time reminders to help students maintain proper sitting posture while studying. These optimizations represent an effective solution for promoting healthy learning habits, particularly in the context of hyperactive students. OpenPose is an open-source computer vision library that can be used to acquire human joint point data. By analyzing the characteristics of a person's sitting posture and judging the duration of any abnormal sitting posture, OpenPose can identify abnormal learning postures and provide real-time reminders. The library uses 18 joints to map the acquired human joint map and obtain more accurate analysis results for the sitting posture: V0 (nose), V1 (neck), V2 (right shoulder), V3 (right elbow), V4 (right wrist), V5 (left shoulder), V6 (left elbow), V7 (left wrist), V8 (right hip), V9 (right knee), V10 (right ankle), V11 (left hip), V12 (left knee), V13 (left ankle), V14 (right eye), V15 (left eye), V16 (right ear), and V17 (left ear). The schematic diagram of specific human joint points is shown in Fig. 1.

When using a visual sensor on a tablet to monitor sitting posture behavior, it is necessary to extract specific joint points that align with the guiding principle of "head straight, shoulder flat, and body straight." In this context, V0 (nose), V1 (neck), V2 (right shoulder), V5 (left shoulder), V14 (right eye), V15 (left eye), V16 (right ear), and V17 (left ear) are optimal choices as they can accurately identify and track the relevant body parts needed for monitoring proper sitting posture. By utilizing these joint points and adhering to the guiding principle, it is possible to detect and monitor abnormal sitting posture behaviors in real time.

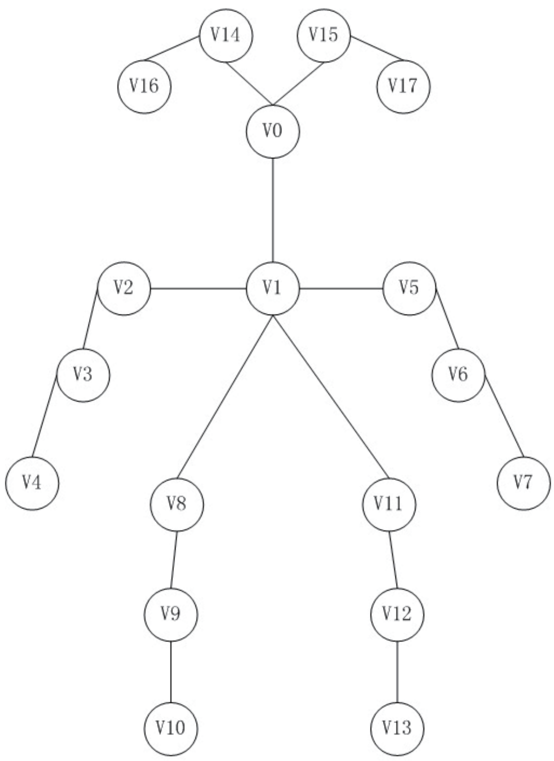


Fig. 1. Schematic diagram of human joint point data obtained by OpenPose.

To better monitor and classify abnormal sitting postures, we can divide them into five distinct states: head tilt, body tilt, head lowering, reading too close, and extended learning duration. However, it is important to consider the unique characteristics of each student's hyperactivity behavior when identifying abnormal posture states. To address this concern and reduce false positives, we recommend incorporating the duration of each state in calculating its corresponding abnormal index score. Below are the specific characteristics and index calculation formulas for each abnormal state.

(1) The abnormal sitting posture of body tilt is identified by monitoring whether the body tilt angle exceeds preset or recommended thresholds and distinguishing whether it is left-leaning or right-leaning from to the specific body tilt angle. The body tilt angle is calculated as the arctangent function value of the difference between the left and right eye coordinates by the difference between the abscissae. The specific calculation formula for the head tilt angle is

$$Head\,Tilt\,Angle = \arctan\left(\frac{Y_{v14} - Y_{v15}}{X_{v14} - X_{v15}}\right) * \frac{180}{\pi}. \tag{1}$$

(2) The abnormal sitting posture of body tilt is identified by monitoring whether the body tilt angle exceeds preset or recommended thresholds and distinguishing whether it is left-leaning or right-leaning from the specific body tilt angle. The body tilt angle is calculated as the arctangent function value of the difference between the left and right shoulder coordinates divided by the difference between the left and right shoulder abscissae. The specific calculation formula for the head tilt angle is

$$Body\,Tilt\,Angle = \arctan\left(\frac{Y_{v2} - Y_{v5}}{X_{v2} - X_{v5}}\right) * \frac{180}{\pi}. \tag{2}$$

(3) Head lowering refers to an abnormal sitting posture where students read or write with their head lowered due to mismatched table and chair heights. This posture can lead to posture-related health problems such as hunchback. To detect head lowering, we measure the distance between the nose and neck using the y-axis and divide it by a threshold value to obtain the head lowering amplitude. If the head lowering amplitude exceeds the preset or recommended threshold value, the student is considered to be in a head lowering state. The specific calculation formula for the head tilt angle is

$$Head\,Lowering\,Amplitude = \left|Y_{v0} - Y_{v1}\right|. \tag{3}$$

(4) Reading too close refers to a situation where students' eyes are positioned too close to books or tablet PCs when reading, which can increase the risk of myopia. To measure the reading distance degree, we track changes along the x-axis of a student's left and right eyes. If the reading distance degree exceeds a preset or recommended threshold, it is considered that reading is happening too close to the eyes. The specific calculation formula for the head tilt angle is

$$Reading\ Distance\ Amplitude = \left| X_{v14} - X_{v15} \right|. \quad (4)$$

(5) A very long learning duration refers to situations where students exceed the recommended duration of continuous learning. To prevent eye strain and improve learning efficiency, primary and secondary school students are generally advised to take a 10–15 min break every 45 min to one hour of study. We can detect learning duration by tracking coordinates of the nose and left and right shoulders. When the duration of continuous learning exceeds a preset or recommended threshold, it is considered that the student has been studying for too long and needs a break to rest their brain and eyes. This information helps promote healthy learning habits and prevent potential long-term problems associated with excessive study time.

4. Design of Abnormal Sitting Posture Detection Process

4.1 System architecture design

The intelligent desk continuously monitors children's abnormal sitting posture on the basis of abnormal sitting posture monitoring algorithms and returns relevant data to the management server. The data storage and management server automatically records the number of instances of abnormal sitting posture, the duration of each instance, and the video footage of the abnormal sitting posture (applicable to the home version) detected by the intelligent desk. The guardian app and the teacher app can display detailed information about the child's abnormal sitting posture over a recent period of time. The intelligent desk has two types of user: home users and school users. In schools, administrators or homeroom teachers can set abnormal sitting posture thresholds by using system-recommended values or personal experience values for each class. If there are special circumstances, the student system also supports the ability to adjust settings for individual students. Homeroom teachers can view an analysis report for their managed classes on the management server that lists the number and duration of abnormal sitting postures for each student. With this information, teachers can focus on students who exhibit prolonged periods of abnormal learning posture and help them correct it. In home settings, parents can adjust the threshold for abnormal sitting posture on the terminal continuously on the basis of their child's actual situation. Parents can log in to the server to view analysis reports of their child's abnormal sitting posture during specific time periods and watch videos to help correct their child's sitting posture. The entire system architecture uses the cloud service model, with data storage and management servers, intelligent desk apps, guardian apps, and teacher apps. The intelligent desk continually monitors children's abnormal sitting posture using monitoring algorithms and returns relevant data to the management server. The server automatically stores a record of the number of abnormal sitting postures observed by the intelligent desk, including video footage of duration and occurrence (this only applies to home users). Both the guardian app and teacher app can display detailed information on a child's abnormal sitting posture over time. The system architecture for intelligent learning desk is shown in Fig. 2.

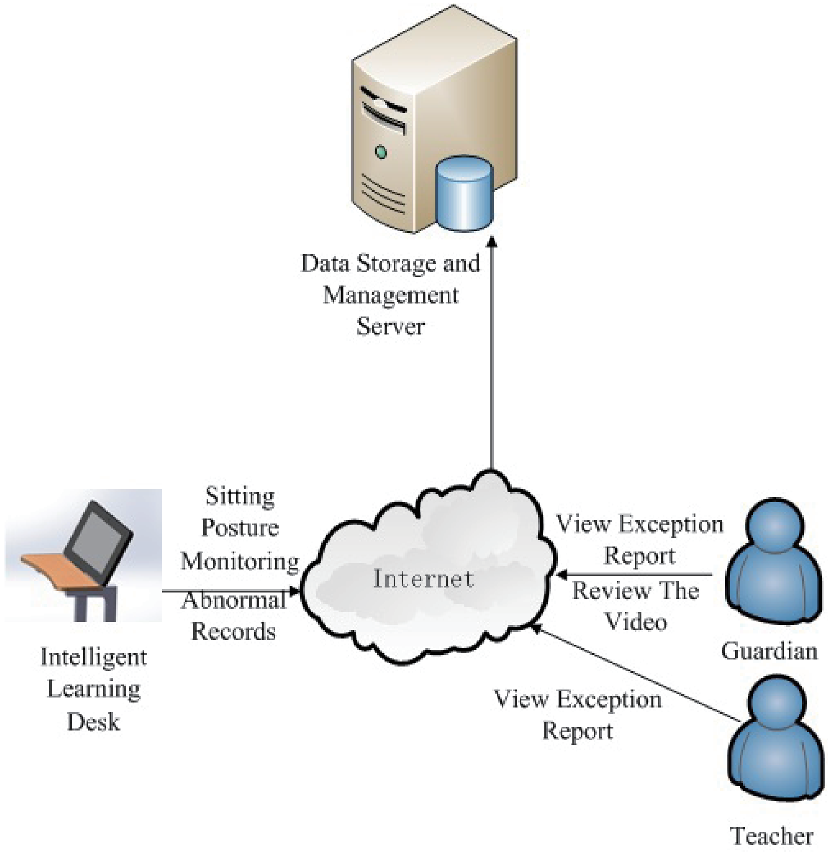


Fig. 2. (Color online) System architecture for intelligent learning desk.

4.2 Detection process design

The intelligent desk is tailored for primary school senior students and junior high school students between the ages of 9 and 18, who experience significant growth in height and weight during this period. As such, the monitoring system cannot rely on a fixed threshold to judge abnormal sitting posture across all students. Instead, guardians or teachers are required to set individualized thresholds based on the unique physical characteristics of each student through the interface provided by the system. Additionally, the system automatically shares reference values from other terminals to assist in setting appropriate thresholds relative to the student's actual situation. To determine the presence of abnormal sitting posture, the system utilizes Eqs. (1) to (4) in combination with the set thresholds to monitor students' sitting positions. The system records the duration of any instances of abnormal posture and notifies students if this duration exceeds the set threshold. The system also records the type of abnormal sitting posture present to provide additional insight into the student's behavior. The specific algorithm business flow chart is shown in Fig. 3.

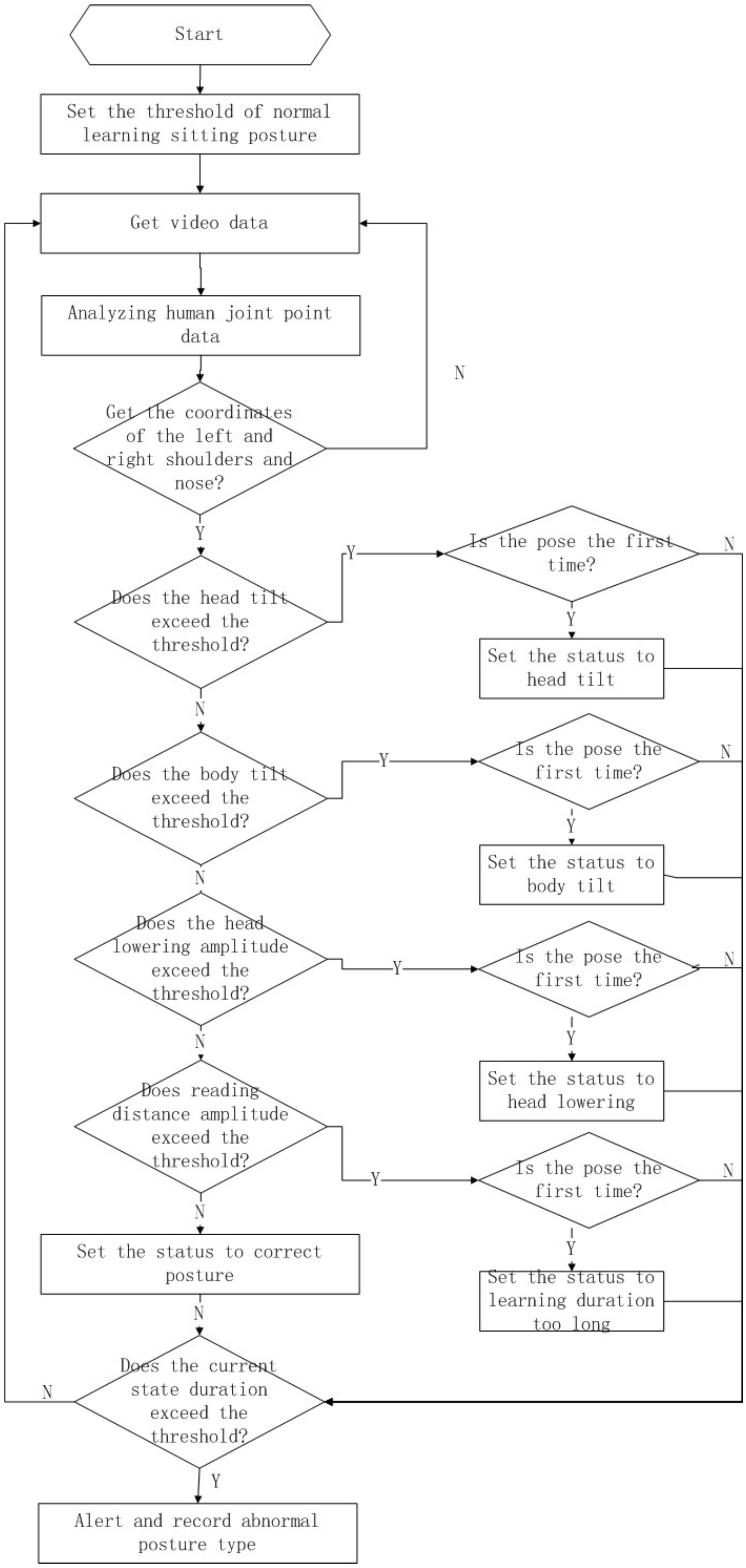


Fig. 3. Abnormal learning sitting posture detection process.

5. Lightweight improvement of OpenPose

To enable the deployment of OpenPose on the intelligent desk's embedded device, we addressed the issue of the large parameter size of the OpenPose model by replacing part of the convolution kernel in the prediction network with deep separable convolution. This modification reduces the number of parameters within the network, resulting in faster and more efficient calculations while ensuring the recognition accuracy. Deep separable convolution is derived from Google's Mobilenet neural network, which was designed specifically for mobile or embedded devices. This technique involves feeding an image with N input features (both length and width D) to a group of convolution kernels (convolution size H) to output depthwise feature maps of N feature images. The depthwise feature map contains spatial features related to each channel present in the original image. By convolving these depthwise feature maps with K 1×1 convolution kernels, a pointwise feature map is produced, which serves as the final result of the process. This approach reduces computational complexity throughout the network, making it possible to use OpenPose on the intelligent desk's embedded device without sacrificing recognition accuracy. The specific process is shown in Fig. 4.

Therefore, the amount of computation required for deep separable convolution is as follows:

$$Conv_{dp} = D \times D \times N \times H \times H + D \times D \times N \times K. \tag{5}$$

If traditional convolution is used, the calculation amount is as follows:

$$Conv_{comm} = D \times D \times N \times H \times H \times K. \tag{6}$$

The ratio of the calculation amount between depthwise separable convolution and traditional convolution is as follows:

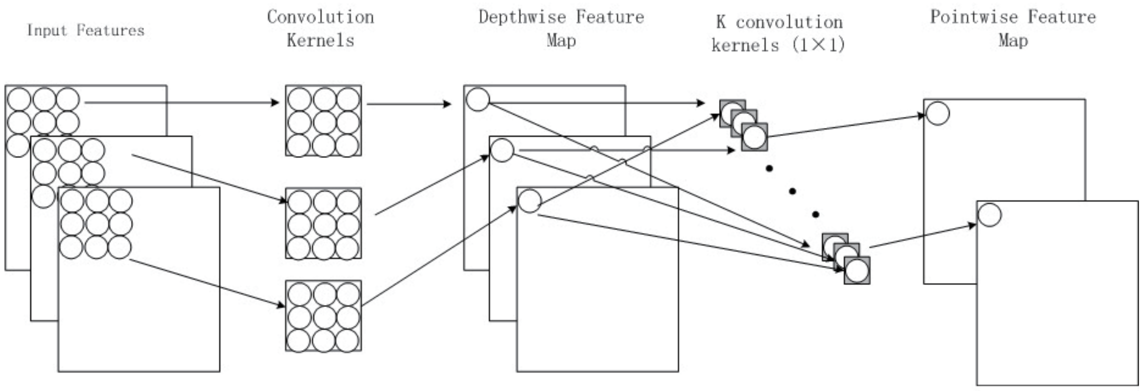


Fig. 4. Deep separable convolution process.

$$\frac{Conv_{dp}}{Conv_{comm}} = \frac{D\times D\times N\times H\times H + D\times D\times N\times K}{D\times D\times N\times H\times H\times K} = \frac{1}{K} + \frac{1}{H\times H}. \tag{7}$$

OpenPose, based on CNN, detects key points of the human skeleton in the supervised learning mode and adopts a large convolutional kernel to obtain large receptive fields. The network extracts the features of the input picture through the traditional CNN vgg19 to obtain the feature map and then inputs the feature map into the two-branch multistage network, which includes upper branch prediction affinity and lower branch prediction confidence. Both branches adopt the fall prediction method. The specific network structure is shown in Fig. 5.

Figure 6 shows the internal structure of the OpenPose network. Except for the first stage in the dual-branch multistage network, the other stages adopted the large convolution kernels (7 × 7). Although large convolution kernels can obtain a larger receptive field, they also cause a large amount of computation.

To optimize the network structure and reduce computational complexity, a technique used in OpenPose is depthwise separable convolution, which replaces large convolution kernels with a series of smaller ones. Specifically, a 7 × 7 kernel is replaced by three 3 × 3 kernels in series to achieve the same receptive field while markedly reducing the number of parameters. To prevent the vanishing gradient problem as the network deepens, skip connections are added between every three consecutive convolutions, with a 1 × 1 convolution used for dimensionality reduction. This architecture modification reduces the computational cost and improves the overall performance of the system. The specific structure is shown in Fig. 7.

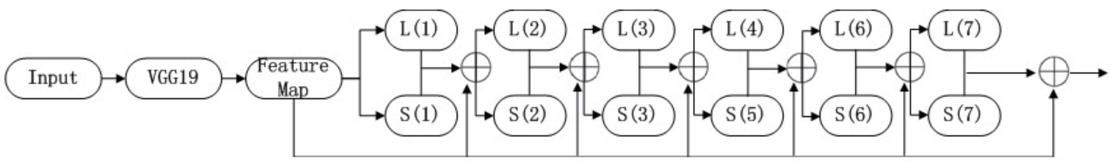


Fig. 5. Main structure of OpenPose.

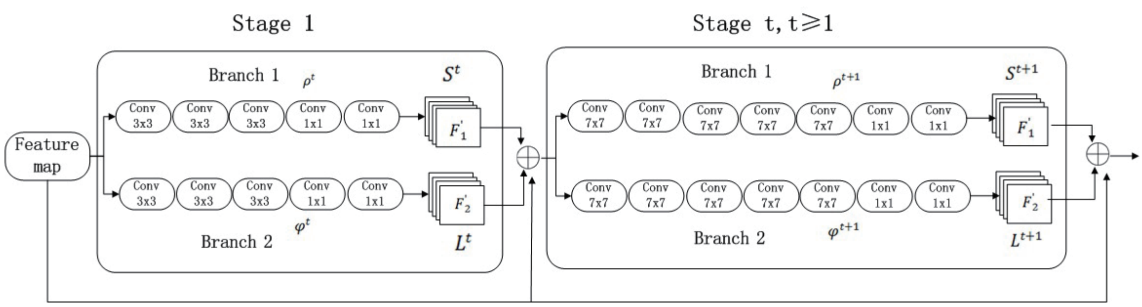


Fig. 6. Internal structure of OpenPose.

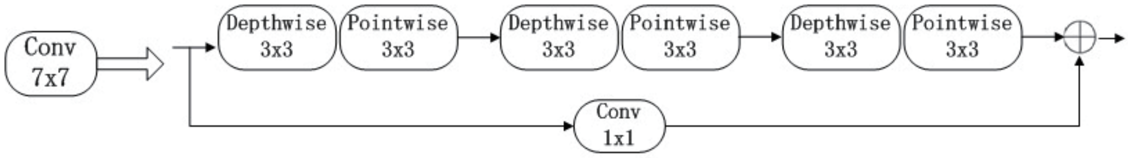


Fig. 7. Convolution kernel structure after replacement.

Deep separable convolution is also used to improve each convolution kernel (3×3) in the first stage of the two-branch multistage network. The improved double-branch multistage network can achieve up to 1/9th of the computational cost compared with the original network through this optimization method, as shown in Eq. (7). This technique helps make the network more efficient and faster without compromising performance.

6. Experimental Results and Analysis

6.1 Experimental results and analysis of OpenPose lightweight improvement

The aim of optimizing OpenPose is to make it suitable for use on lightweight mobile devices such as intelligent desk and mobile terminals. The hardware environment was set up using NVIDIA's Jetson TX2 development board and NVIDIA 1080ti GPU. The software environment was Ubuntu 20.04 and Python 3.9. To evaluate the performance of the optimized model, the COCO2017 dataset was used for comparison. The initial learning rate was set to 4×10^{-5}, and the back propagator was the Adam optimizer. The number of iterations for training was set to 250000, and the training batch size was set to 8. The mean average precision (mAP) was selected as the evaluation metric, with the ap50 and ap75 used to measure the key point predictors with thresholds of 0.5 and 0.75, respectively. By optimizing the network architecture and reducing computational complexity, the optimized OpenPose model achieved good results on the COCO2017 dataset while remaining lightweight enough to be used on mobile devices. Table 1 shows that the prediction score of the lightweight OpenPose model is slightly lower than the original, but the processing speed per second for video has been significantly improved. Prior to optimization, the average video recognition speed was approximately 1.6 s per frame, while after optimization, it increased to around 0.7 s. Thus, the computational load of the entire network can be considerably reduced without compromising accuracy.

6.2 Experimental results and analysis of abnormal sitting posture monitoring

Currently, there is no standardized dataset available for evaluating sitting posture health that meets the specific requirements of this study. Therefore, we utilized self-collected sitting posture testing data. In accordance with the GB/T 26158-2010 "Chinese Minor Body Size" standard, 24 primary school students (12 males and 12 females) ranging from 9 to 18 years old were selected as test subjects. Two students (1 male and 1 female) were selected from each age group, ensuring

Table 1
Evaluation results of OpenPose lightweight improvement.

Model	AP (%)	AP50 (%)	AP75 (%)
OpenPose	63.5	84.3	68.2
This study	61.7	81.5	66.8

Table 2
Relevant data of testers.

Project	Age (years)	Height (cm)	Weight (kg)
Maximum value	9	142.5	41.2
Minimum value	18	181.3	73.5

Table 3
Threshold data of abnormal sitting posture of tester.

Age (years)	Head tilt (°)	Body tilt (°)	Head lowering (px)	Reading distance (px)	Learning duration (min)
9–10	±70	±30	0.13	0.013	45
11–12	±60	±20	0.14	0.013	45
13–15	±55	±16	0.16	0.014	45
16–18	±50	±13	0.18	0.015	45

a diverse sample. All the subjects had no history of back pain, had normal or corrected vision, were right-handed, and could maintain accurate sitting posture while holding a pen. Additionally, all the subjects were confirmed to be healthy for the study. The specific data of the subjects are shown in Table 2.

Owing to the variation in body size and shape among students of different ages, it is not feasible for the monitoring system to use a constant value to judge whether a student's sitting posture is abnormal. Therefore, before recognition, a threshold value for abnormal sitting posture must be set in accordance with common sense, which can be adjusted using the system's visual setting functions. To determine the sitting posture, the system uses a portrait alignment box, which allows for the accurate detection and analysis of the student's posture. The threshold values of abnormal sitting posture set for different ages are shown in Table 3.

To reduce false positives and increase accuracy, the monitoring system sets the duration of abnormal posture to a uniform 15 s before triggering a warning. The health status of the corresponding sitting posture is then determined by analyzing the threshold data for abnormal sitting postures of each individual tester, as shown in Table 3. Furthermore, the system tests each abnormal sitting posture ten times in order to focus on recognizing postures that are close to the boundary value. By using these methods, the system can effectively reduce errors and provide accurate feedback on students' sitting posture. The recognition accuracy is shown in Table 4.

Table 4 indicates that the system has a high recognition rate for abnormal sitting postures such as prolonged learning duration, head tilt, body tilt, and head lowering, but its accuracy in recognizing reading distance that is too close is low. This is due to the limitation of the OpenPose bone node technology, which only provides a plan view of the user. The system can only use the distance change on the X axis of the left and right eyes as the reading distance degree.

Table 4
Accuracy of abnormal sitting posture recognition.

Project	Total Number	Positive Number	Recognition accuracy (%)
head tilt (left)	240	225	93.8
head tilt (right)	240	225	93.8
body tilt (left)	240	229	95.4
body tilt (right)	240	229	95.4
head lowering	240	222	92.5
reading distance too close	240	170	70.8
learning duration too long	240	240	100

7. Conclusions

To effectively monitor and address abnormal sitting postures among primary and secondary school students, a stable monitoring system based on OpenPose was implemented. The system comprehensively assesses the health of sitting postures by considering three dimensions: human posture, critical threshold value, and abnormal posture duration. The performance of the system was evaluated through experiments and practice, its effectiveness in detecting and warning students about common abnormal sitting postures was demonstrated. The design of the abnormal sitting posture monitoring algorithm is based on the feature extraction of various types of abnormal sitting posture, as well as the design of an abnormal sitting posture detection process. In addition, improvements were made to the OpenPose technology to enable it to be deployed on embedded devices with limited processing resources. The experimental process and structure of the monitoring system are described in detail, including the selection of test subjects, data collection and analysis, and system evaluation. Overall, the results showed that the monitoring system was effective in detecting and correcting poor sitting postures among students, and could potentially improve the long-term health outcomes of young students.

Acknowledgments

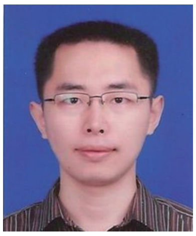
This study was supported by the special fund for electronic information engineering technology specialty group of the national double high program of Dongguang Polytechnic in 2021 (No. ZXF022), 2022 (Nos. ZXA202201 and ZXD202202), and 2023 (Nos. ZXD202308, ZXD202303, and ZXD202302), 2021 Engineering Technology Center of Colleges and Universities in Guangdong Province (No. 2021GCZX016), key social development projects of Dongguan in 2020 (No. 2020507156156), special fund for Dongguan Rural Revitalization Strategy in 2021 (No. 20211800400102), key social development projects in Dongguan (No. 2020507156156), 2022 Special Innovation Projects of Universities in Guangdong Province (No. 2022KTSCX327), Key Fields Special Project of Guangdong Provincial Department of Education in 2022 (No. 2022ZDZX1074),2023 (Nos. 2023ZDZX1088, and 2023ZDZX1087), Dongguan Science and Technology Ombudsman Project in 2022 (No. 20221800500812), Dongguan Science and Technology of Social Development Program (No. 20231800935472), Dongguan Science and Technology of Social Development Program (Nos. 20231800903592 and 20211800900252), Dongguan Science and Technology Ombudsman Project (No. 20231800500282), Special Projects in Key Fields of Colleges and Universities in Guangdong Province in 2021 (No. 2021ZDZX1093), and Dongguan Sci-tech Commissioner Program (Nos. 20231800500352, 20221800500792 and 20221800500812).

References

1 W. W. Pan, Z. Zhao, W. C. Huang, Z. Zhang, L. Y. Fu, Z. G. Pan, J. Yu and F. Wu: IEEE Trans. Neural Networks Learn. Syst. **12** (2022) 1. https://doi.org/10.1109/TNNLS.2022.3212900
2 L. Ma, Y. H. Zheng, Z. Zhang, Y. Z. Yao, X. J. Fan, and Q. L. Ye: IEEE Trans. Circuits Syst. Video Technol. **33** (2023) 2061. https://doi.org/10.1109/TCSVT.2022.3222305
3 L. Y. Fu, D. Zhang, and Q. L. Ye: IEEE Trans. Geosci. Remote Sens. **59** (2021) 8257. https://doi.org/10.1109/TGRS.2020.3042507
4 Q. L. Ye, H. Peng, Z. Zhang, Y. H. Zheng, L. Y. Fu, and W. K. Yang: IEEE Trans. Cybern. **12** (2022) 12745. https://doi.org/10.1109/TCYB.2021.3088519.
5 L. Y. Fu, Z. C. Li, Q. L. Ye, H. Yin, X. B. Chen, X. J. Fan, W. K. Yang, and G. W. Yang: IEEE Trans. Neural Netw Learn Syst. **1** (2022) 130. https://doi.org/10.1109/TNNLS.2020.3027588.
6 X. Chen, M. Li, H. Zhong, Y. Ma, and C. H. Hsu: IEEE Trans. Ind. Informatics **18** (2022) 2820. https://doi.org/10.1109/TII.2021.3075464.
7 T. M. Hoang, T. A. Pham, V. V. Do, V. N. Nguyen, and M. H. Nguyen: ATC **3** (2022) 683.https://doi.org/10.1109/atc55345.2022.9943049.
8 X. Chen, J. Q. Hu, Z. Y. Chen, B. Lin, N. Xiong, and G. Y. Min: IEEE Trans. Ind. Inf. **4** (2022) 2724. https://doi.org/10.1109/TII.2021.3076393
9 X. Chen, L. J. Yang, Z. Y. Chen, G. Y. Min, X. H. Zheng, and C. Rong: IEEE Trans. Cloud Comput. **21** (2022) 1. https://doi.org/10.1109/TCC.2022.3169157.
10 G. Huang, C. R. Luo, K. D. Wu, Y. Ma, Y. Zhang, and X. Z. Liu: IEEE Trans. Cloud Comput. **10** (2019) 1. https://doi.org/10.1109/ICDCS.2019.00166
11 W. Lan, L. Ke, and C. H. Wu: Proc. 2010 IEEE 2nd Int. Computer Engineering and Technology Conf. (IEEE,2010) 692–695.
12 M. Tariq, H. Majeed, M. Beg, F. Khan, and A. Derhab: Future Gener. Comput. Syst. **92** (2019) 745. https://doi.org/10.1016/j.future.2018.02.013.
13 S. S. Kumara, P. Chattopadhyaya, and L. Wang: arXiv **20** (2021) 1. https://doi.org/10.48550/arXiv.2110.09564.
14 Y. Y. Chen, C. H. Hsia, H. Y. Kao, Y. A. Wang, and Y. C. Hu: JIT **21** (2020) 1893. https://doi.org/10.3966/160792642020122107004.
15 C. H. Hsia, C. H. Chien, H. W. Hsu, J. S. Chiang, and H. W. Tseng: JIFS **36** (2019) 1171. https://doi.org/10.3233/JIFS-169891.
16 N. S. Pai, Y. H. Chen, C. P. Hung, P. Y. Chen, Y. C. Kuo, and J. Y. Chen: Sens. Mater. **31** (2019) 3451. https://doi.org/10.18494/SAM.2019.2481.
17 J. Qu, C. Wu, Q. Li, T. Wang, and A. H. Soliman: Sens. Mater. **32** (2020) 1209. https://doi.org/10.18494/SAM.2020.2527.
18 A. Toshev and C. Szegedy: Proc. 2014 IEEE Conf. Computer Vision and Pattern Recognition. (CVPR, 2014) 1653–1660.
19 W. Yang, S. Li, W. Ouyang, H. Li and X. Wang: Proc. 2017 IEEE Int. Conf. Computer Vision (ICCV, 2017) 1281–1290.
20 W. Tang and Y. Wu: Proc. 2019 IEEE Conf. Computer Vision and Pattern Recognition (CVPR, 2019) 1107–1116.
21 Z. Feng, X. Zhu, and Y. Mao: 2019 IEEE Conf. Computer Vision and Pattern Recognition (CVPR, 2019) 3517–3526.

About the Authors

Yanjun Zhu received his B.S. degree from Hengyang Normal University, China, in 2005 and his M.S. degree from the University of Chinese Academy of Sciences, China, in 2008. From 2008 to 2011, he was an engineer at Shenzhen Tianyuan Dike Information Technology Co., Ltd, China. Since 2011, he has been working at Dongguan Polytechnic. His research interests are in AI, smart education, and sensors.



Shi Wang received his B.S. degree from Hunan Normal University, China, in 2000 and his M.S. degree from Huazhong University of Science and Technology, China, in 2009. Since 2013, he has been working at Dongguan Polytechnic. His research interests are in AI, automatic control, and sensors.

Yong Peng received his M.S. degree in computer science from Wuhan University of Technology, Wuhan, China, in 2010. He is currently an associate professor with the School of Electronic Information Engineering, Dongguan Polytechnic. His current research interests include machine learning, pattern recognition, and expert systems.

Yuqiang Chen received his M.S. and Ph.D. degrees from Guangdong University of Technology, China, in 2010 and 2017, respectively. Since 2009, he has been working at Dongguan Polytechnic. His research interests are in big data, intelligent transportation, automatic control, and sensors.

Rongli Chen received her B.S. and M.S degrees from Guangdong University of Technology, China, in 2007 and 2009, respectively. From 2013 to 2019, she was a junior experimenter at Guangdong Innovative Technical College, China. Since 2019, she was been a senior experimenter at the School of Artificial Intelligence, Dongguan Polytechnic, China. Her research interests are in cloud computing, IoT, and intelligent manufacturing.

Gang Ke received his B.S. degree in computer science and technology from Hubei Normal University, Huangshi, China, in 2007 and his M.S. degree in computer application technology from Guangdong University of Technology, Guangzhou, China, in 2010. He is currently a Ph.D. student at Macau University of Science and Technology, Macau. His current research interests include computer vision, network security, deep learning, and IoT.

Jianxin Li received MS degree in School of computing at the University of Guangdong University of Technology in 2011. He is a Lecture at University of Dongguan Polytechnic. His research interest focuses on machine vision, behaviour recognition, deep learning and IoT.