Standardizing and Early Warning of Sewing Beginners' Posture Based on CNN Visual Recognition Technology

Zejun Huang[1], Zhen Qin[1], Hanze Ge[1]

[1] South China University of Technology, Guangzhou Guangdong, China
sang2@163.com

Abstract. This study focuses on the posture problem in the sewing process and utilizes computer vision recognition technology to identify and warn the posture status of sewing workers, thereby reducing the risks and stress during the sewing process. By analyzing actual sewing scenarios, conducting questionnaires, summarizing expert opinions, and categorizing different sewing postures, the specific requirements and technical roadmap of computer vision recognition in sewing conditions are determined. OpenPose is used to extract the posture features of novice sewers in the sewing process, and then a deep learning network is built and trained using the PyTorch framework, enabling the model to provide early warnings for incorrect sewing postures of novice sewers.

Keywords: Sewing Posture, Early Warning, Computer Vision Recognition, OpenPose, PyTorch.

1 Introduction

Sewing technology is essential to clothing design and production [1] [2]. The correctness of sewing posture affects the effectiveness of sewing production, the comfort of the body [3] [4], and the operator's safety [5]. Long-term incorrect posture also can lead to physical damage [6] [7]. Standardizing the posture of sewers in the early stages of learning can promote their development of good posture habits. Therefore, developing a system that can monitor and correct the posture of novice sewers in real time is of great significance.

In terms of sewing posture, it has always been a focus of research in human factors engineering. The literature [8] adjusted the sewing machine workstation, including table height, table tilt, and pedal position, to assess their impact on work posture and worker perception. The literature [9] used surface electromyography (sEMG) to investigate muscle load and activity patterns in the neck and shoulder muscles of female sewing machine operators and identified several key periods of muscle fatigue within 200 minutes of working time. The application of ergonomic interventions during these periods can help prevent sewing machine operators from developing work-related musculoskeletal disorders. Although these research did not directly involve the evaluation and correction of sewing posture using computer vision technology, they indirectly demonstrate the significance of this study.

In terms of posture recognition, there are two types of pose recognition methods: hardware-based methods and image-based methods. Image-based methods have lower costs and are more efficient for multi-person recognition [10]. Therefore, image-based methods have been widely applied in industrial safety and human pose detection in recent years. The literature [11] identifies abnormal behaviors such as emergencies and machine failures in workers, and the literature [12] automatically evaluates hand postures of piano players to help beginners improve their performance skills. However, there is currently relatively little research on the use of this technology for posture regulation and early warning for beginners in sewing.

In this paper, OpenPose is utilized to extract posture features of novice sewers during the sewing process [13] [14]. The extracted feature maps, marked with skeletal nodes of the students, are used as the training dataset. Subsequently, a deep learning network is constructed using the PyTorch framework, and the network is trained using the dataset. This enables the model to provide early warnings for incorrect sewing postures of novice sewers.

2 Research Organization

The research mainly consists of four steps: scene and requirement analysis, collection and classification of postures, technology organization, and finally, experimental operations.

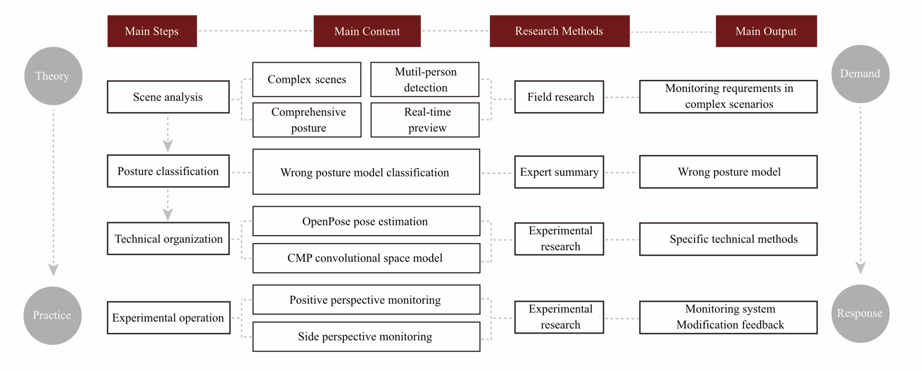


Fig. 1. The main research flowchart.

2.1 Scene and Requirement Analysis

The sewing classroom or factory is a special and complex scene. Recognizing and alerting postures in this scene requires first understanding the various factors at play, such as the layout and methods of sewing machines, which contribute to the complexity of the recognition scene. It is also important to consider whether multiple individuals need to be identified, the comprehensive nature of the postures, and whether real-time feedback is necessary for monitoring. The main research method at this stage is to conduct on-site research of the sewing machine room, and ultimately summarize the specific requirements for identifying and warning sewing postures in complex scenarios.

2.2 Collection and Classification of Postures

To enable computers to recognize sewing postures more quickly and accurately, this study will collect and classify various postures that sewing beginners encounter during the sewing process. The main research methods at this stage are questionnaire surveys and expert summaries, mainly outputting various classified sewing postures.

2.3 Technical Organization

To meet the monitoring requirements in complex scenes for this research, technical selection and improvement are conducted. The main research method in this stage is experimental research, aiming to summarize specific technical methods suitable for this study.

2.4 Experimental Operation

Based on specific requirements and technological organization, this research conducted on-site monitoring and warning experiments in a sewing scene. The experiments were mainly divided into two monitoring perspectives: frontal and lateral, resulting in the development of a practical and operable sewing posture monitoring and warning system. Additionally, issues and requirements encountered during the practical operations were collected for further modifications and improvements in the next steps.

3 Computer Vision Technology

In computer vision technology, there are two main methods: top-down approach and bottom-up approach.

Top-down approach typically require early decisions between person detection and pose estimation, which can lead to incorrect allocation and pose prediction. After detecting a person, the algorithm immediately starts estimating their pose without waiting for the detection results of other people. This early decision-making can result in erroneous pose estimation, especially in cases of person overlap or occlusion.

Bottom-up approach first detect all the body key points in the image instead of detecting individuals one by one. Then, clustering techniques are used to group these key points into person poses. This approach has the advantage of robustness in multi-person scenes and can better handle cases of overlap and occlusion.

Additionally, the operational efficiency of the two algorithms in multi person recognition is different. By comparing the runtime of the two approaches using single-person Convolutional Pose Machines (CPM), the computational complexity of the top-down method increases linearly with the total number of people in the image, while the runtime of the bottom-up method increases relatively slowly with the number of people.

This is because the top-down algorithm starts pose estimation immediately after detecting a person, so the runtime is directly proportional to the number of people in the image. The bottom-up algorithm first detects all the body key points in the image, decoupling the runtime from the number of people and making it dependent on the number of key points instead. The bottom-up algorithm has the potential to decouple runtime from the number of people in the image, allowing it to maintain relatively stable runtime even with an increasing number of people. This makes the bottom-up approach more advantageous in handling images with many people.

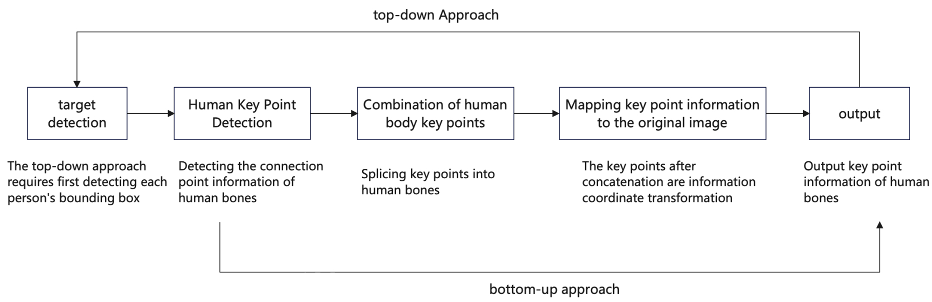


Fig. 2. The difference between top-down approach and bottom-up approach.

Based on these characteristics and factors, we consider Carnegie Mellon University's OpenPose technology for pose estimation. OpenPose is a bottom-up algorithm that first detects all the body key points in the image and then regresses all the key points using Convolutional Pose Machines (CPM), thus providing better robustness in multi-person scenes. The CPM model used in OpenPose employs large convolutional kernels to obtain a large receptive field, which can better handle the complexity and occlusion of sewing laboratory equipment and joints.

4 Experimental Research

4.1 Scene Requirement Analysis

Through specific visits to the research scene, this study summarizes the characteristics of the sewing process scene and the factors to consider in posture monitoring:

First, complex scene: Sewing areas are typically equipped with multiple sewing machines, and beginners operate around these machines. Multiple sewing equipment can cause occlusion of the human body, increasing the difficulty of computer recognition.

Second, multi-person environment: In a sewing scene, multiple individuals are usually operating multiple sewing machines simultaneously, so posture estimation needs to be able to recognize and track the poses of multiple individuals.

Third, comprehensive poses: Correct sewing poses involve not only hand and arm movements but also head, neck, and body postures. Therefore, in posture estimation, it is necessary to consider the key point information of the face and body comprehensively.

Fourth, real-time monitoring and warning: In the sewing scene, instructors of sewing techniques need to observe and guide the postures and operations of beginners. Therefore, the posture estimation system needs to be able to monitor the postures of beginners in real-time and issue warnings when incorrect postures are detected, so that instructors can intervene and guide in a timely manner.

4.2 Collection and Classification of Postures

To collect and classify sewing postures, this study conducted a survey based on field investigations. The survey respondents were mainly beginners and experienced individuals who had been involved in sewing for more than 1 hour per day on average, and 70 valid responses were collected.

The survey investigated dangerous situations and incorrect behaviors during the sewing process. Among the respondents, the behavior considered to be the most dangerous was needle pricks, accounting for over 80% of the total. The main factor causing dangerous behavior was found to be lack of concentration, accounting for 68.57%. The second factor was lack of skill, accounting for 24.29%. Therefore, it is important to maintain concentration, improve skill proficiency, avoid rushing, and reduce the likelihood of injury during sewing.

Moreover, respondents reported feeling eye fatigue or discomfort and muscle fatigue during long sewing sessions. The most common body parts to feel tired or uncomfortable during sewing were the neck and shoulders, followed by the waist and back. It is important to maintain good posture and take breaks and relax the relevant muscles to reduce fatigue and discomfort during sewing. Most respondents (92.86%) had not tried using tools or equipment to improve their working posture and reduce strain, indicating that they may not have paid enough attention to the impact of working posture and strain.

Furthermore, most participants (78.57%) had not received training or guidance on correct posture and ergonomics during sewing, which may increase the risk of incorrect postures and movements and lead to discomfort or injury during sewing.

To enable the computer to quickly recognize incorrect action patterns, this study also collected some common actions of novice sewers. These actions were rated and judged by instructors and individuals with extensive sewing experience. Based on these common actions, we classified them into three typical error types:

(a) Lack of concentration during sewing operations.

(b) Simultaneously holding or picking up other items during sewing operations.

(c) Adopting incorrect postures such as bending, hunching, or lying down during sewing operations.

These movements can often cause physical harm and deviations in production for beginners. Therefore, based on these three common risky movements, an observation and warning program based on body posture recognition was designed and produced to observe the concentration level and behavior deviations of sewers.

4.3 System Design

Based on the causes of incorrect postures obtained from the survey and the common posture forms, we designed two monitoring perspectives: a frontal view and a side view. By simultaneously using information from multiple perspectives, we can obtain more comprehensive and accurate human posture estimation results. This provides valuable information and guidance for posture correction and danger warning applications.

Frontal View. The frontal view is better for recognizing the facial and hand postures of the operator and paying attention to issues such as lack of concentration and picking up other items during sewing. The frontal view involves more complex and varied movements. We considered using OpenPose to extract key points information to draw skeletal features and then train a deep learning network model based on PyTorch for judgment.

To determine lack of concentration, we use facial orientation. Instead of training a model, we calculate whether the x-axis distance between the two eyes and the nose is equal to determine if there is a perspective relationship on the face. If there is a perspective relationship, it indicates that the operator may have a lack of concentration.

Side View. The side view provides a better observation of the overall body posture, including the head, neck, torso, and legs. The side view directly uses the key points information extracted by OpenPose to calculate angles and provide real-time reminders for similar incorrect body postures. The calculation formula is as follows:

$$\tan\theta = \frac{|y_1 - y_2|}{|x_1 - x_2|}$$

In conclusion, we have developed a warning system that combines frontal and side views. The frontal view separates the facial and body parts for detection. The logic of the detection process is illustrated in Figure 3.

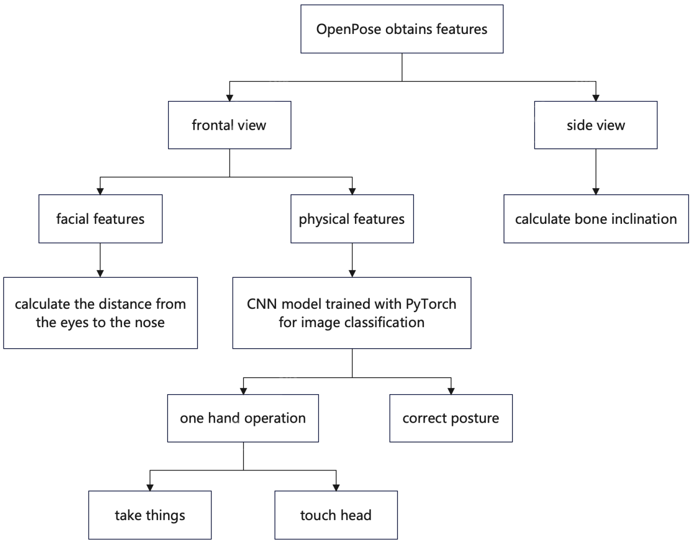


Fig. 3. The logic of early warning systems.

4.4 Construct the Convolutional Neural Network

Dataset Construction. During the construction of the dataset, we found that the number of captured photos was insufficient. Therefore, we exported each frame of the captured video as a jpeg format image. OpenPose was used to extract key points information from each image and draw skeletal features. To reduce redundant data, we made some adjustments to the images. The background of the images was set to black, and only the graphical information within the range of the minimum and maximum coordinates of x and y was retained. Partial datasets are shown in Figures 4 and 5. Before applying the entire dataset to the training process, it was preprocessed with the following steps:

(a) The images were divided into three categories: correct posture, take things, and touch head, and placed in three separate folders.

(b) All images were resized to (224, 224) pixels, with black pixels added to the edges.

(c) Data augmentation was applied to expand the dataset. Random horizontal flipping was applied to the images to increase the diversity of the dataset, improve the generalization ability of the model, make the model more robust, and reduce the risk of overfitting.

(d) The images were normalized to have pixel values ranging from 0 to 1.

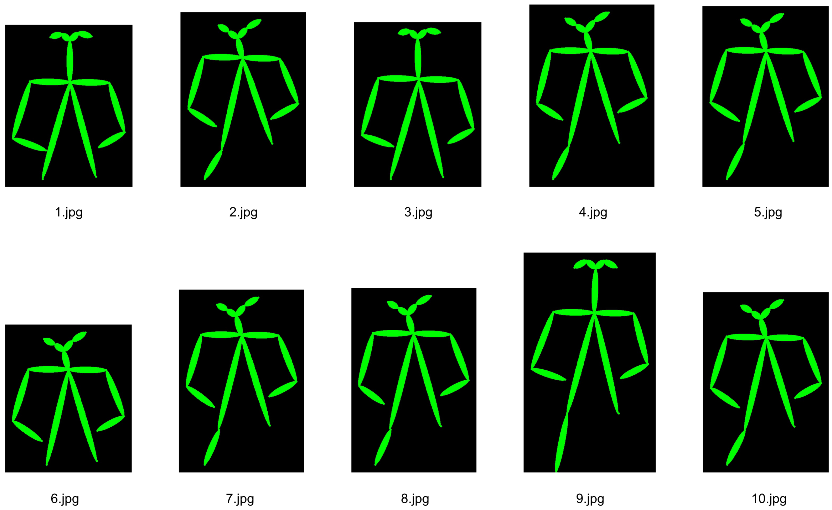


Fig. 4. Dataset of correct sewing posture.

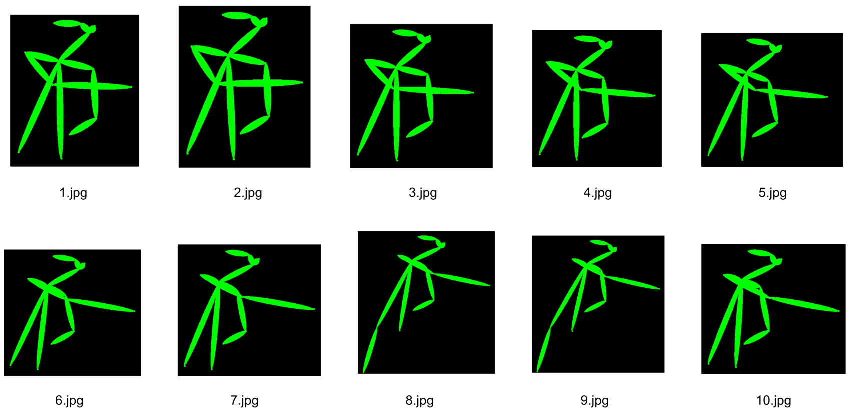


Fig. 5. Dataset of sewing while taking things.

Model Selection. The CNN model used for training was VGG16. VGG16 is a convolutional neural network model proposed by Simonyan and Zisserman in the paper "Very Deep Convolutional Networks for Large Scale Image Recognition". VGG16 has a simple structure, consisting of 13 convolutional layers, 3 fully connected layers, and 5 pooling layers. All convolutional layers in VGG16 use the same convolutional kernel parameters, with a kernel size of 3x3. All pooling layers use the same pooling kernel parameters, with a kernel size of 2x2. The pooling layers have a stride of 2 and use max pooling.

Model Training. We defined the loss function as the cross-entropy loss function and used the Adam optimizer to optimize the model's parameters. In the training loop, the number of epochs was set to 100. In each epoch, we first set the model to the training mode, then iterated through the training data loader, calculated the model's output and loss, and updated the model's parameters. At the same time, we calculated and recorded the loss value during the training process. Then, we set the model to the evaluation mode, iterated through the validation data loader, and calculated the accuracy of the model on the validation set. The accuracy was recorded in TensorBoard and the model with the highest accuracy on the validation set was saved. Finally, after training was completed, we printed out the best accuracy achieved during the training process and saved the model's parameters. The loss curve and accuracy curve are shown in Figures 6 and 7.

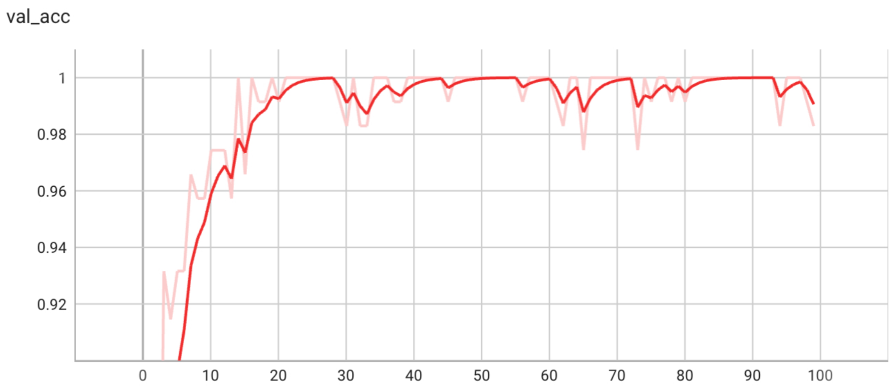


Fig. 6. Curve of training accuracy.

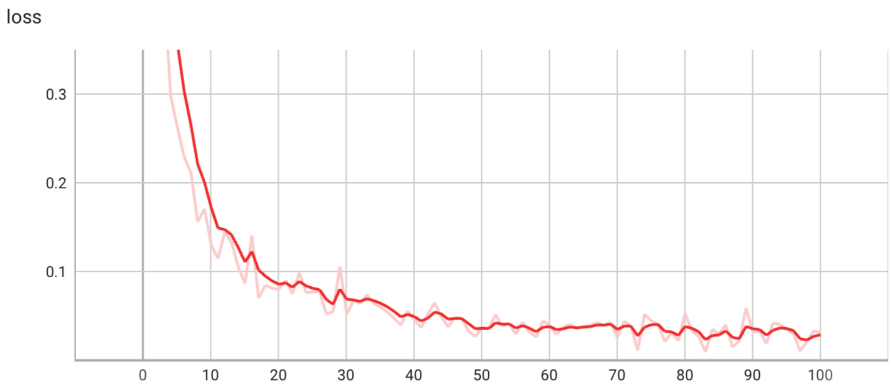


Fig. 7. Curve of training loss.

Attempt to input two bone images for testing, and the image classification results are shown in Figure 8, indicating a high recognition accuracy.

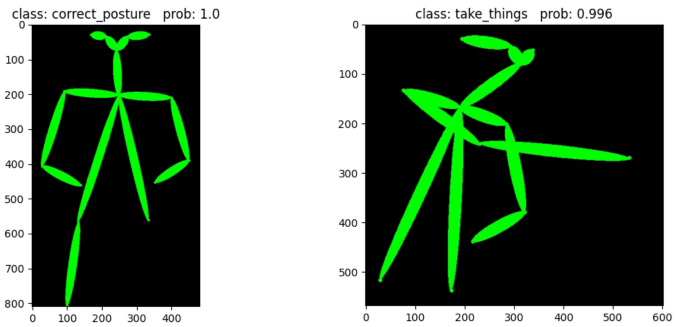
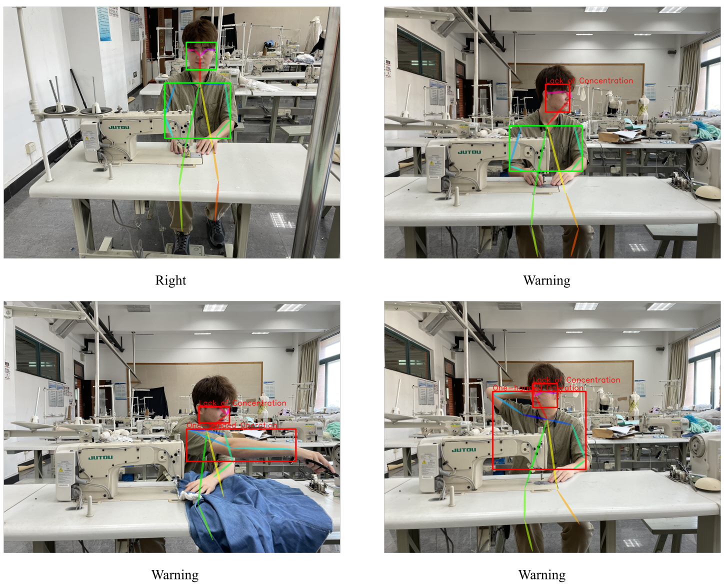


Fig. 8. Predicted Posture Results

4.5 Experimental Application

Finally, we applied the experimental techniques to specific sewing scenarios for monitoring experiments. When the seamstress's posture is correct, it is indicated by a green box. When there is a deviation in the sewing posture, a red warning box appears on the screen. In the frontal view, warnings are issued when the seamstress turns their head, reaches out, scratches their head, and other postures are detected.



Right

Warning

Warning

Warning

Fig. 9. Recognition and early warning from the frontal view.

In the side view, warnings are triggered when there is a significant deviation in the seamstress's torso, hands, and feet, such as bending over or hunching. These deviations are indicated by red warning boxes.

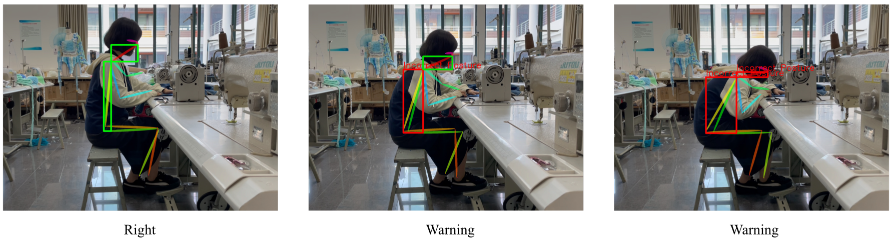


Fig. 10. Recognition and early warning from the side view

5 Conclusion

By using computer recognition of beginners' facial and body data, we can effectively observe the behavioral states of sewing beginners during the sewing process. Compared to the traditional method of manual observation and intervention in sewing instruction, computer vision recognition is more comprehensive and timelier, and can simultaneously recognize multiple samples. Computer recognition also offers convenience and operability, as it can be implemented using a webcam. It can serve as an auxiliary means for practical teaching and risk control. However, during the experiment, some deviations were observed in the OpenPose recognition. In the future, with advancements in hardware technology and computing power, the accuracy of recognition can be further improved by considering multi-view 3D human pose estimation.

Acknowledgments. This study was funded by Guangzhou Municipal Bureau of Science and Technology, Guangzhou Basic Research Program Basic and Applied Basic Research Special General Project（SL2024A04J00955）

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Kim S Y. A Comparative Study of Contents of Korean Basic Sewing Textbook[J]. Journal of the Korean Society of Costume, 2012, 62(3): 73-83.
2. Jing-Bin Y , Heng L .Comprehensive and Innovative Experimental Teaching of Industrial Engineering Based on Sewing Production[J].Research and Exploration in Laboratory, 2013.
3. Vihma T, Nurminen M, Mutanen P. Sewing-machine operators' work and musculo-skeletal complaints[J]. Ergonomics, 1982, 25(4): 295-298.

4. Tartaglia R , Cinti G , Carrara S ,et al.[Work posture and changes in the spine of sewing workers in the clothing industry][J].La Medicina del lavoro, 1990, 81(1):39-44.
5. Kirin S, Šajatović A H. Research of Working Postures in the Technological Sewing Process Using the REBA Method[C]//Proceedings of the 8th International Ergonomics Conference: ERGONOMICS 2020 8. Springer International Publishing, 2021: 111-119.
6. Jensen B R, Schibye B, Søgaard K, et al. Shoulder muscle load and muscle fatigue among industrial sewing-machine operators[J]. European journal of applied physiology and occupational physiology, 1993, 67: 467-475.
7. Zhang F, He L, Wu S, et al. Quantify work load and muscle functional activation patterns in neck-shoulder muscles of female sewing machine operators using surface electromyogram[J]. Chinese medical journal, 2011, 124(22): 3731-3737.
8. Delleman N J, Dul J. Sewing machine operation: workstation adjustment, working posture, and workers' perceptions[J]. International Journal of Industrial Ergonomics, 2002, 30(6): 341-353.
9. Fei-ruo Z, Li-hua H, Shan-shan W, et al. Quantify work load and muscle functional activation patterns in neck-shoulder muscles of female sewing machine operators using surface electromyogram[J]. Chinese Medical Journal, 2011 (22).
10. Chen K. Sitting posture recognition based on openpose[C]//IOP Conference Series: Materials Science and Engineering. IOP Publishing, 2019, 677(3): 032057.
11. Rijayanti R, Hwang M, Jin K. Detection of Anomalous Behavior of Manufacturing Workers Using Deep Learning-Based Recognition of Human–Object Interaction[J]. Applied Sciences, 2023, 13(15): 8584.
12. Johnson D, Damian D, Tzanetakis G. Detecting hand posture in piano playing using depth data[J]. Computer Music Journal, 2020, 43(1): 59-78.
13. Wei S E, Ramakrishna V, Kanade T, et al. Convolutional pose machines[C]//Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. 2016: 4724-4732.
14. Cao Z, Simon T, Wei S E, et al. Realtime multi-person 2d pose estimation using part affinity fields[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 7291-7299.