2020 IEEE 3rd International Conference on Computer and Communication Engineering Technology

Skeleton-Based Sleep Posture Recognition with BP Neural Network

Haozhou Lyu
School of Information Science and Engineering, University of Jinan, Jinan 250022, China
Shandong Provincial Key Laboratory of Network Based Intelligent Computing, Jinan 250022, China
e-mail: 20161222036@mail.ujn.edu.cn

Jinglan Tian*
School of Information Science and Engineering, University of Jinan, Jinan 250022, China
Shandong Provincial Key Laboratory of Network Based Intelligent Computing, Jinan 250022, China
e-mail: ise_tianjl@ujn.edu.cn

Abstract—**Human sleep postures are inextricably linked to health, which can be used as a pivotal indicator of disease prevention and treatment. To obtain a machine learning model for analyzing the human sleep postures, a new approach is proposed to efficiently recognize the types of sleep postures based on skeleton extraction. Four typical sleep postures, i.e., lying in the supine, prone, left lateral and right lateral, are classified with the method of extraction of key points relation feature as well as the direct coordinate feature, which can extract features of skeleton correctly and effectively. Furthermore, the presented method is applied to a specific scenario, which is utilized for monitoring sleep postures of patients who suffered from Obstructive Sleep Apnea-Hypopnea Syndrome (OSAHS) by making the detailed classification of supine posture. The effectiveness of the proposed framework was validated quantitatively and qualitatively. The performance of the extensive comparison experiments demonstrate that the proposed approach is superior and achieves the state-of-the-art.**

Keywords-sleep posture recognition; object detection; pose estimation; machine learning

I. INTRODUCTION

Sleep postures, especially for the sickly, older, convalescent patients, can predict some illnesses or physical emergencies. Choosing suitable and healthy sleep posture can effectively prevent possible illnesses [1], [2]. Hence, it is of great medical significance and application value to use computer vision technology to identify sleep postures in order to observe health status of the human body, which also plays an important role in relieving the pressure of medical resources and reducing the use of medical personnel. Therefore, considerable attention is supposed to be paid to propose approaches to classify and analyze sleep postures efficiently and accurately.

Recently, there are few researches on sleep posture recognition based on markerless methods purely depending on video sequences. Most studies are based on wearable devices [3], [4], [5] and experimental results are not entirely satisfactory. Algorithms based on wearable devices, mainly rely on several wearable sensors on body. Huang Yiqin et al. [3] stated a high-risk sleep posture multi-sensor data fusion detection system using gyroscope, pressure sensors and other devices. This system obtains spatial coordinates through real-time information collection of several sensor groups and by the data information fusion processing of specific parts through the single chip microcomputer, then uses the Naive Bayes classifier to judge the sleep postures of the target person and alarms the dangerous postures set by the user. Two problems arise due to the fact that this algorithm is based on wearable devices. One problem is that changing sleep posture during sleep may cause damage to tiny sensors they are wearing, which leads to inaccurate data transmission. The other problem is that devices need to be worn while sleeping, which may have unconscious effects on users' behaviors and cause inconvenience.

Ren Zhibin et al. [4] claimed an unconstrained fuzzy-rough set algorithm. Firstly, the geometric features are extracted from the static pressure image obtained from a flexible pressure sensor array. Then, in the fuzzy-rough set algorithm, characteristic values for recognition are analyzed as well as the continuous pressure distribution is discretized and converted to fuzzy decision table to identify sleep posture effectively by eliminating redundant information and inferring fuzzy decision rules. Experimental results show that this algorithm can achieve an accuracy of 92.90% for the typical sleep postures recognition. However, this method requires pressure sensors to cover the whole bed, and the hardware cost is unaffordable and the process is too complex. Zhang Yichao et al. [5] pointed out a Ballistocardiogram (BCG) signal based sleep position pattern recognition algorithm. BCG signals from four different sleep positions are collected with a piezoelectric film sensor in a non-contact way. Eight characteristics are extracted from the preprocessed BCG signals after wavelet-based denoising. The experimental results show that recognition accuracy of the neural network algorithm and KNN are 93.00% and 84.00% respectively. Although this method might avoid the problems mentioned above, i.e., the process is simple and the cost of equipment is low, it also has limitations. The brightness, humidity and temperature of the experimental environment will have irregular and random influence on sensors, and noises such as footsteps and electric drills will impact the BGG signal negatively.

In order to avoid problems arose from wearable devices, we tend to propose algorithms to analyze sleep postures based on monocular camera. Compared with the former algorithm, the latter has many advantages, which do always not need users to wear devices and cannot be restricted by environment. Ye Yinqiu et al. [6] proposed a sleep posture

Authorized licensed use limited to: American University of the Middle East. Downloaded on May 05,2025 at 12:18:04 UTC from IEEE Xplore. Restrictions apply.

recognition algorithm based on level set method and neural network. Firstly, level set method is used to segment the area of human body in the image, and then BP neural network model is set to train and recognize sleep postures. However, there are still a few demerits. For example, when background in images is particularly complex, recognition rate of level set method is pretty low. Thus, it is possible to have a large deviation between segmented area and targeted area, which is detrimental to recognition effects. Moreover, due to the complexity of sleep postures, recognition efficiency is low under special circumstances, such as lying in the prone but with face up. The average recognition accuracy is presented at 75.50%, which is unsatisfactory for realistic using. In order to deal with the complicated background in images, human pose estimation and object detection algorithms are utilized to gain features of complex human postures efficiently and effectively.

Yang Mingjian et al. [7] pointed out a single person sleep posture recognition algorithm based on OpenPose. By processing the sleep images collected by monocular camera, human pose estimation and sleep posture recognition can be realized by positioning key points of the face, head and various parts of the human body. Compared with previous researches' results, this algorithm has a great improvement, but still has several shortcomings. Firstly, using face key points is not a perfect way to determine the mid line of body, since face is always active during sleeping. It is unnecessary to use correction factors to get the mid line. Secondly, it is unfeasible in practice to distinguish supine and prone lying postures based on the number of key points of eyes mentioned in [7]. The number of key points of eyes extracted from supine posture is 1 or 2, and the counterpart of prone posture is 0 or 1. It is noted that there is an intersection when the number equals 1. Hence these two postures cannot be divided by this method. Thirdly, [7] did not mention how to distinguish left and right lateral postures. To differentiate four typical sleep postures practically and feasibly, we propose to extract the feature of key points relation as well as the direct coordinate feature, which are helpful to analyze and utilize the features of target's face and body to construct classifier.

In this work, we propose a skeleton-based sleep posture recognition framework with BP neural network. First, YOLO object detection algorithm [8] and human pose estimation algorithm [9] are utilized to preprocess images. Second, extracting key points relation feature as well as the direct coordinate in order to obtain the face and human body features much more accurately. Besides, the proposed framework has been applied to OSAHS patients scene, in which the supine sleep posture is further classified to two particular postures, with an emphasis on the role of the head position while lying in the supine [12].

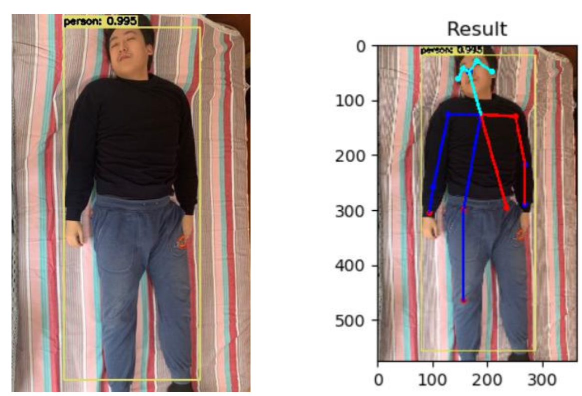
To demonstrate the performance of the proposed approach, some experiments are conducted to compare the results with above five approaches. The experimental results are shown qualitatively and quantitatively on sequences collected by ourselves. The results have demonstrated that the proposed model leads to a significant improvement in accuracy of sleep posture recognition.

II. PROPOSED APPROACH

The proposed framework in this paper can be divided into four stages. The first one is to detect the specific position of the target person in images using an object detection algorithm. The next step is to obtain the coordinate information of skeleton key points of the target by human pose estimation. After that, the proposed method presents an approach to efficiently recognize four typical type of sleep postures-supine, prone, left lateral and right lateral-based on extraction of key points relation feature as well as direct coordinate feature. Last but not the least, this paper makes further analysis on the application of specific scenario, i.e., applying this approach to monitor sleep postures of patients who suffered from OSAHS by refining the supine position to two particular types of sideways head posture and perpendicular head posture.

A. *Pre-processing*

First, images are pre-processed for better analyzing. In practice, it is found that low-resolution images can be more effective in obtaining key points of body parts. Hence, the images are rescaled to 432 x 576 before extracting the skeleton of human body. The skeleton key points are then obtained by adopting the OpenPose algorithm. There might exist false detections because of clutter background, i.e., human skeleton key points are posed in places where no human body is present [9], [10]. In this way, the accuracy of deep learning model and classification based on coordinate information of skeleton key points would be reduced. Therefore, before implementing skeleton extraction, an object detection algorithm is combined to obtain the detailed position of the target person in image as shown in Fig. 1(a). Then the bounding box location of the detection result will be compared with the key points coordinate achieved from skeleton extraction in order to assure the target person is located correctly as shown in Fig. 1(b).



(a)Human Detection Result (b) Skeleton from Human PoseEstimation

Figure 1. The example of the results of object detection and skeleton key points.

B. *YOLO Object Detection*

The images are first processed to obtain the region of interest in each frame utilizing an object detection algorithm YOLO, which is proposed by Redmon i.e., [8], [11]. It is a ready to use algorithm, in which the whole image is taken as

Authorized licensed use limited to: American University of the Middle East. Downloaded on May 05,2025 at 12:18:04 UTC from IEEE Xplore. Restrictions apply.

the input of the network and the detection results are directly returned with category and positions of the bounding boxes in the output layer. The whole object detection algorithm, in a sample way, is to firstly resize the input image to 448x488, then through the training of the convolutional neural network model, and finally divide the detected target results according to the confidence of the object.

C. Human Pose Estimation

The frames are then processed to extract skeleton by a human pose estimator, which takes the entire image as the input for VGG-19 CNN to jointly predict confidence maps for body part detection, and part affinity fields for parts association. The parsing step performs a set of bipartite matching to associate body parts candidates. It finally assemble them into full body poses for all people in the image [9].

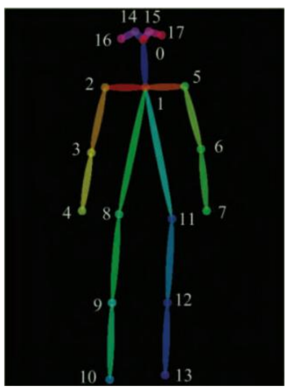
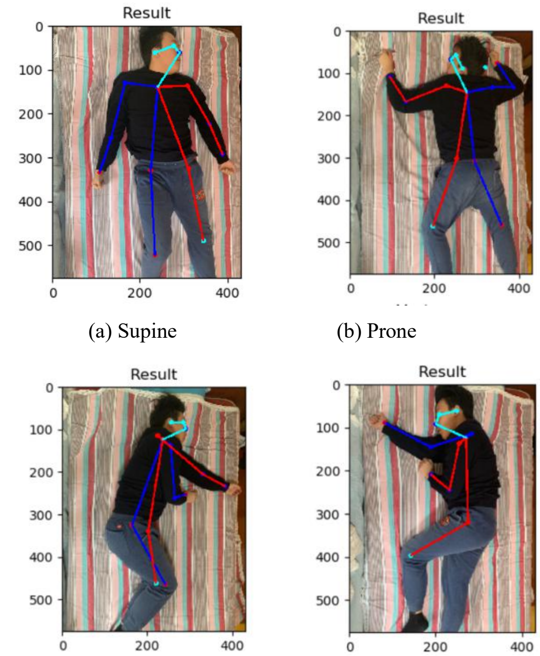


Figure 2. 18 key points of body obtained from human pose estimation.

One generic human pose estimator, OpenPose provides detection and position of 18 or 25 key points of human skeleton, 70 key points of face and 21 key points of hands. The main difference between 18 and 25 key points is that the latter has more information related to left and right foot. Since key points information of feet has little effect on classification and recognition of sleep postures, this paper uses coordinate information of 18 skeleton key points as initial input data of the algorithm. Apart from that, hands and face key points are not significant characteristics of postures as well as detection of them may reduce the operational efficiency of processing images. In summary, 18 key points as shown in Fig. 2 are selected as input of sleep posture recognition system.

D. Feature Extraction

Considering characteristics of sleep postures in frames, the key points indirect relation feature and the direct coordinate feature are selected to extract and analyse face and body features. The extraction of indirect relation feature is to extract relations among key points as eigenvectors, while the extraction of direct coordinate feature is to calculate and compares the geometry of the coordinates of skeleton key points to determine behaviors. In terms of the behaviour characteristics of four typical postures, supine, prone, left lateral and right lateral as shown in Fig. 3, we divide features into face features and body features to achieve characteristic information.



(a) Supine (b) Prone

(c) Left-Lateral (d) Right-Lateral

Figure 3. Four typical sleep postures: supine, prone, left lateral and right lateral. The red represents the left part of the body, the blue represents the right part of the body.

1) Face Feature Extraction

Face features extraction in eyes and ears of key points are particularly vital, because under the circumstance of left lateral lying posture with left face down, the human pose estimation algorithm is unable to identify left eye and left ear at the same time. There must be a status that horizontal value of one of their key points is 0, which means this key point cannot be detected. But the information of right eye and right ear is uncovering with right face up, horizontal value and vertical value of them can be detected, right lateral lying posture and vice versa. In a word, if the detected number of left eye and left ear is less than the counterpart of right eye and right ear, the posture can be recognized as left lateral, right lateral posture and vice versa. Thus, left lateral and right lateral postures can be distinguished by the detected number of eyes and ears at the same time.

2) Body Feature Extraction

In addition, body features are also essential for sleep posture classification. To distinguish prone and supine postures, the most important thing is to discriminate the front and back of the body. OpenPose itself can judge two sides of human body according to key points of face and then determine left and right orientation of shoulders, elbows and other parts. It is proved that the OpenPose framework can recognize the left and right side of body key points whether the human body is facing front or back by recording videos including one single person rotating on the same spot. When the posture is supine, the origin of coordinate system is on the upper left corner of the image, thus the horizontal

Authorized licensed use limited to: American University of the Middle East. Downloaded on May 05,2025 at 12:18:04 UTC from IEEE Xplore. Restrictions apply.

coordinate value of right shoulder's key point is less than the counterpart of left shoulder, the prone posture and vice versa. What's more, the detection accuracy of both shoulders points is high and the distance between them is relatively large, hence the distance between shoulders is completely suitable for distinguishing supine and prone postures. When human body is on supine and prone sleep posture, two arms are generally located on two sides of the body, while when the human body is on left lateral or right lateral sleep postures, both arms are located on the same side. Therefore, supine, prone, left lateral and right lateral sleep posture can be preliminarily classified into two groups according to whether both arms are located on the same side of the body. In this way, the four-classification problem is converted to a binary-classification problem, which simplifies the process.

E. Feature Correctness Verification

The correctness of face features and body features is premise of correctly recognizing sleep postures. Hence, the following experiments are performed to verify the correctness of feature extraction. In order to substantiate the correctness of face features extraction, we define the number of key points of right eye and right ear detected under the circumstance of left lateral posture as Sl_1, and the number of key points of left eye and left ear as Sl_2. The number of key points of right eye and right ear detected under the circumstance of right lateral posture is Sr_1, and the number of key points of the left eye and left ear is Sr_2. It can be easily deduced that $Sl_1 = 2$, $Sl_2 = 0$ or 1; $Sr_1 = 0$ or 1, $Sr_2 = 2$. The key points are first extracted from 200 frames, 100 frames showing left lateral posture and 100 frames showing right lateral posture respectively. The number of key points are then compared with Sl_1, Sl_2, Sr_1 and Sr_2. The equality of both number pairs denotes that the face features are extracted correctly. In order to verify the correctness of the body feature extraction, it is supposed that the body features are extracted correctly when the horizontal value of right shoulder is less than the counterpart of left shoulder, supine posture and vice versa.

The experimental results are illustrated in Table I. The accuracy of face and body features extraction achieves to 97.00%, which demonstrates that the approach for extracting direct coordinate feature is accurate. The accurate features provide an effective foundation for classifying sleep postures in later steps.

TABLE I. THE EXPERIMENTAL RESULTS OF FOUR TYPICAL SLEEP POSTURES

Sleeping posture	The number of tests	The number of correctness	Accuracy(%)
Right-Lateral	100	97	97.00
Left-Lateral	100	98	98.00
Supine	100	96	96.00
Prone	100	97	97.00
Total/Average	400	388	97.00

F. Model Building

According to the extraction method of indirect relation feature mentioned above, coordinate relation among several key points is a good clue for classification. Since the movement of head and body are relatively separate and irrelevant during sleeping, the information of face key points does not affect the judgement of the positions of both arms, which means that the key points of face can be excluded. In contrast, key points of shoulders and neck are pivotal. The position of two arms is related to the key points close to shoulders and arms. Thus, neck, shoulders, elbows and wrists, total of seven key points' coordinate information altogether 14 number (v_1, v_2, v_3, ..., v_{14}) are considered as feature vectors.

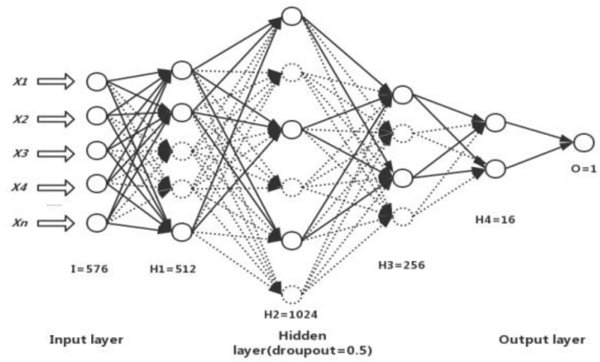
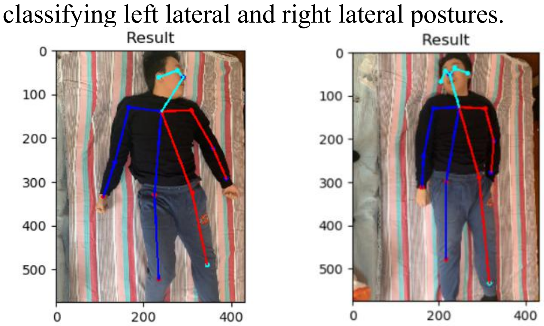


Figure 4. BP Neural Network model(1 input layer, 4 hidden layers, 1 output layer).

G. Scenario Analysis

This paper also refines sleep posture categories based on specific application scenario. Based on the observation of sleep posture in medical treatments of patients suffering from OSAHS, it has been confirmed that perpendicular head posture leads to greater risk of death compared with sideways head posture due to the reason of falling back of tongue and upper airway obstruction [1], [12]. In this scenario, supine posture can be subdivided into two categories by using the extraction of direct coordinate feature method: sideways head posture and perpendicular head posture as shown in Fig. 5. The classification procedure is the same as the above process of classifying left lateral and right lateral postures.



(a) Sideways head supine (b) Perpendicular head supine

Figure 5. The example of sideways head supine posture and perpendicular head supine posture.

Authorized licensed use limited to: American University of the Middle East. Downloaded on May 05,2025 at 12:18:04 UTC from IEEE Xplore. Restrictions apply.

III. Experiments

A. Training Details

First of all, carry out image acquisition and pre-processing, a total of 1200 images of single sleep positions were captured from overhead perspective. There were 300 images of each of four basic sleep postures, 250 images for training and 50 images for prediction. BP neural network model (as shown in Fig. 4) is constructed, including input layer, 4 hidden layers and output layer. Feature vector mentioned in subsection model building is as the input nodes of the network. The output layer has 1 node representing whether two arms are on the same side of the body. If they are on the same side, the output is 1, otherwise output is 0. The training set includes 1000 frames, 250 for each type of sleep posture. First, we shuffle the dataset to increase randomness, and then select 100 frames as test set, the rest are certified by k-fold cross validation. Experiment is conducted on the Keras deep learning framework. Adam is applied as the optimization strategy [13]. The batch size is 64. Cross-entropy is selected as the loss function to back-propagate gradients [14]. The weight decay is set to 0.0001.

B. Evaluation of Neural Network

For the binary classifier, we evaluate the BPNN by indexes as precision, recall and F1. We stipulate that the case of two arms on the same side of the body is positive and the case of two arms on the opposite side of the body is negative. Related terms are defined as follows. True positive (TP) represents the number of cases correctly identified as same-side. False positive (FP) is the number of cases incorrectly identified as same-side. True negative (TN) equals the number of cases correctly identified as opposite-side. False negative (FN) is the number of cases incorrectly identified as opposite-side. F1 is harmonic mean of Precision and Recall [15]. The evaluation results of the BPNN is shown in Table II.

TABLE II. THE EVALUATION RESULT OF THE BINARY CLASSIFIER

Index	**TP**	**FN**	**FP**	**TN**
Number	96	4	5	95
Indicator	**Accuracy**	**Precision**	**Recall**	**H-mean**
Rate	0.955	0.95	0.96	0.955

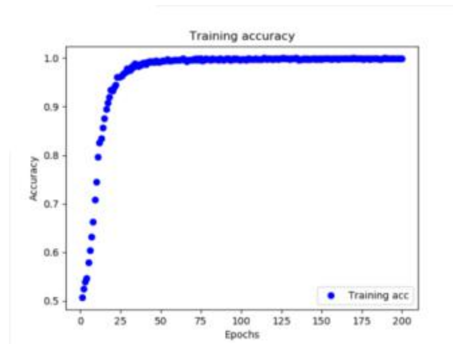
C. Experimental Results

After 200 iterations of training, the error loss of training set is 0.0838 and the accuracy is 0.9982 as shown in Fig. 6. The error loss of test set is 0.1948 and the accuracy is 0.96. In order to verify the accuracy of this sleep posture recognition algorithm, we choose 50 images of each posture, totally 200 images for the final prediction and recognition.

$$d = \frac{n}{N} \times 100\% \tag{1}$$

where n represents the number of correctly recognized samples, N indicates the total number of samples, and d means the accuracy of sleep posture recognition.

(a) Training Loss



(b) Training Accuracy

Figure 6. BP neural network training results.

The experimental results are illustrated in Table III and Table IV. The accuracy of the algorithm is 95.00%, and the recognition accuracy of the reclassification of supine posture reaches to 97.00%.

TABLE III. THE EXPERIMENTAL RESULTS OF FOUR TYPICAL SLEEP POSTURES

Sleeping posture	**The number of tests**	**The number of correctness**	**Accuracy(%)**
Right-Lateral	50	46	92.00
Left-Lateral	50	49	98.00
Supine	50	48	96.00
Prone	50	47	94.00
Total/Average	200	190	95.00

TABLE IV. THE EXPERIMENTAL RESULTS OF RECLASSIFICATION OF SUPINE POSTURE

Sleeping posture	**The number of tests**	**The number of correctness**	**Accuracy(%)**
Right-Lateral	50	48	96.00
Left-Lateral	50	49	98.00
Total/Average	100	97	97.00

D. Comparison with the State-of-the-Art

The experimental results of the proposed approach is compared with other methods of sleep posture recognition proposed in recent years as shown in Table V. This table illustrates that the accuracy of the proposed approach is higher than the counterpart of state-of-the-art methods.

Authorized licensed use limited to: American University of the Middle East. Downloaded on May 05,2025 at 12:18:04 UTC from IEEE Xplore. Restrictions apply.

TABLE V. COMPARISON WITH OTHER METHODS

Method	The number of tests	The number of correctness	Accuracy(%)
Ye Yinqiu	200	151	75.50
Huang Yiqin	600	549	91.50
Yang Mingjian	200	185	92.50
Ren Zhibin	84	78	92.90
Zhang Yichao	100	93	93.00
Ours	200	190	**95.00**

IV. CONCLUSION

In this paper, we propose a framework for sleep posture recognition. First, YOLO object detection algorithm and human pose estimation algorithm are utilized to preprocess images. Second, extracting key points relation feature as well as the direct coordinate in order to obtain the face and human body features much more accurately. Four typical sleep postures, i.e., lying in the supine, prone, left lateral and right lateral, are classified. Besides, the proposed framework has been applied to OSAHS patients scene, in which the supine sleep posture is further classified to two particular postures. Our experimental results demonstrate the effectiveness of our algorithm, and the proposed algorithm outperforms the other approaches. Moreover, it can realize the refine classification of supine posture in OSAHS patients scenario with high accuracy. This algorithm can be effectively applied to medical environment and meet the requirements of sleep posture detection and classification.

Compared with sleep posture algorithms based on wearable devices, the algorithm based on monocular camera proposed in this paper is more convenient and affordable. In addition, there are various human posture estimation models in the field of computer vision which can be explored and combined to achieve more complicated sleep posture recognition system. In conclusion, this kind of research has a broad prospect.

ACKNOWLEDGMENT

This paper is supported by the Natural Science Foundation of Shandong Province (ZR2018PF012), and National Natural Science Foundation of China (No.61603151).

REFERENCES

[1] J. Li, Z. Li, W. Li, H. Huo, J. Li, and T. Li. Effect of behavioral therapy program on patients with obstructive sleep apnea hypopnea syndrome. *Chinese Journal of Nursing*, vol. 043, no. 12, pp. 1083–1087, 2007.

[2] D. J. Ewing, J. M. Neilson, C. M. Shapiro, J. A. Stewart, and W. Reid. Twenty four hour heart rate variability: effects of posture, sleep, and time of day in healthy controls and comparison with bedside tests of autonomic function in diabetic patients. *Heart*, vol. 65, no. 5, pp. 239–244, 1991.

[3] Y. Huang, J. x. Hu, J. b. Jiang, and Z. Zhang. Dangerous posture monitoring system based on bayesian classifification. *Techniques of Automation and Applications*, vol. 37, no. 09, pp. 112–114, 2018.

[4] Z. Yichao, Y. Zhenming, and S. Xiaoyan. Sleep position recognition based on ballistocardiogram signal. *Computer Engineering and Applications*, 2018.

[5] Z. Ren, L. I. Yang, S. Guo, Z. Guo, and X. Liu. Sleep posture pressure image recognition based on fuzzy-rough set theory. *Computer Engineering and Applications*, 2018.

[6] Y. Ye. Research on sleep posture recognition system based on computer vision, Ph.D. dissertation, *Anhui University of Technology*, 2013.

[7] M. Yang, J. Li, R. Guo, and X. Tang. Implementation and research of human sleep posture recognition based on openpose. *Physical Experiment*, no. 8, 2019.

[8] J. Redmon and A. Farhadi. Yolov3: An incremental improvement. *Arxiv*, 2018.

[9] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh. Realtime multi-person 2d pose estimation using part affifinity fifields. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[10] P. Yuan and N. Liu. Elderly fall detection algorithm based on joint point extraction. *Computer Technology and Development*, 2019.

[11] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once: Unified, real-time object detection, In 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[12] W. Songjian, H. Xiaoqing, and H. Yaqi. Effects of head position on the occurrence of disordered respiratory events during supine sleep. *Beijing Biomedical Engineering*, 2018.

[13] L. Wang and R. Liu. Red blood cell recognition based on bp neural networks. *Computer and Digital Engineering*, 2018.

[14] P. Spurek, K. Byrski, and J. Tabor. Online updating of active function cross-entropy clustering. *Pattern Analysis & Applications*, no. 5, pp. 1–17, 2018.

[15] T. H. Nazer, M. Davis, M. Karami, D. Koelle, and H. Liu. Bot Detection: Will Focusing on Recall Cause Overall Performance Deterioration? *Springer, Cham*, 2019.

Authorized licensed use limited to: American University of the Middle East. Downloaded on May 05,2025 at 12:18:04 UTC from IEEE Xplore. Restrictions apply.