SHI, Zhiming, ZHAO, Zang, CHEN, Junjie and CHENG, Guangxin. Research on sit-up counting method and system based on human skeleton key point detection. Quality in Sport. 2024;24:55408. eISSN 2450-3118.
https://dx.doi.org/10.12775/QS.2024.24.55408
https://apcz.umk.pl/QS/article/view/55408

The journal has been 20 points in the Ministry of Higher Education and Science of Poland parametric evaluation. Annex to the announcement of the Minister of Higher Education and Science of 05.01.2024. No. 32553.
Has a Journal's Unique Identifier: 201398. Scientific disciplines assigned: Economics and finance (Field of social sciences); Management and Quality Sciences (Field of social sciences).
Punkty Ministerialne z 2019 - aktualny rok 20 punktów. Załącznik do komunikatu Ministra Szkolnictwa Wyższego i Nauki z dnia 05.01.2024 r. Lp. 32553. Posiada Unikatowy Identyfikator Czasopisma: 201398.
Przypisane dyscypliny naukowe: Ekonomia i finanse (Dziedzina nauk społecznych); Nauki o zarządzaniu i jakości (Dziedzina nauk społecznych).

The authors declare that there is no conflict of interests regarding the publication of this paper.
Received: 03.09.2024. Revised: 08.10.2024. Accepted: 09.10.2024. Published: 11.10.2024.

Research on sit-up counting method and system based on human skeleton key point detection

Zhiming Shi

College of Artificial Intelligence, Southwest University, Chongqing, China, 400715.

National & Local Joint Engineering Research Center of Intelligent Transmission and Control Technology, Southwest University, Chongqing, China, 400715.

jay19932014@126.com

https://orcid.org/0009-0000-1671-1657.

Hang Zhao

College of Artificial Intelligence, Southwest University, Chongqing, China, 400715.

National & Local Joint Engineering Research Center of Intelligent Transmission and Control Technology, Southwest University, Chongqing, China, 400715.

rgznhang@swu.edu.cn

https://orcid.org/0009-0009-3715-5026.

Junjie Chen

College of Artificial Intelligence, Southwest University, Chongqing, China, 400715.

National & Local Joint Engineering Research Center of Intelligent Transmission and Control Technology, Southwest University, Chongqing, China, 400715.

1031730720@qq.com https://orcid.org/0009-0006-9646-3625.

Guangxin Cheng
School of Physical Education, Southwest University, Chongqing, China, 400715.
942022264@qq.com
https://orcid.org/0009-0003-9376-1000.

Abstract

This paper presents a sit-up counting system based on skeletal key point detection, aimed at addressing the subjectivity and inefficiency of traditional manual counting methods. The system utilizes deep learning algorithms to detect and identify human skeletal key points. By optimizing the network structure and introducing skip connections, the accuracy and efficiency of skeletal key point positioning were significantly improved. The number of parameters in the improved network was reduced from 125.6M to 38.3M, greatly lowering model complexity. Simultaneously, the processing speed increased from 36.5 FPS to 63.6 FPS, demonstrating much higher processing efficiency. By integrating skeletal key point tracking with illegal motion detection, real-time and accurate counting is achieved, with the system reaching a sit-up detection accuracy of 98.57%. The system integrates real-time detection, data collection, display, storage, and query functions, providing an efficient and objective counting solution for sports and testing.

Keywords: sit-ups, skeletal key points, convolutional neural networks, counting system, sport

Introduction

Sit-ups are a common and widely practiced exercise that effectively strengthens the anterior abdominal muscles, including the rectus abdominis, external obliques, internal obliques, and quadratus lumborum. In 2014, the Ministry of Education introduced the revised "National Student Physical Fitness Standards" to establish a comprehensive physical fitness monitoring and evaluation mechanism for primary, secondary, and college students, aiming to further encourage student participation in sports activities and promote the physical and mental well-being and holistic development of youth. Sit-ups were listed as a mandatory test for female students in high school and college, with specific movement standards and test procedures. The sit-up test requires the participant to lie on their back on a soft mat, with legs slightly apart and knees bent at 90 degrees, hands clasped behind the head, while a partner holds their ankles to stabilize the lower limbs. The test administrator begins the test after visually confirming the participant has completed the initial movement setup. When the participant raises their head to an upright position, the test administrator signals with a beep, marking the movement as valid. The number of sit-ups completed within one minute is recorded as the student's test score. [1,2,3]. However, for a long time, sit-up fitness tests have mainly relied on manual timing, manual counting, and visual judgment to determine whether the movements meet technical standards. Although this method is simple and practical, it suffers from significant subjectivity and visual bias. This issue becomes especially problematic when a large number of participants are tested or when counting takes a long time, as manual counting is prone to errors and inaccuracies, which can lead to disputes over final scores. [4].

Traditional motion recognition techniques face challenges in automatic sit-up counting, often struggling with recognition difficulty and inaccurate judgments. Moreover, due to the variability in testing environments, traditional image processing techniques frequently produce errors in the sit-up counting process. Hua Zhiyuan designed a small intelligent sit-up counter based on a posture sensor. The device, installed on the upper arm and circular in shape, uses a posture sensor to establish a three-dimensional coordinate system that detects body movement angles and key parts, thereby enabling sit-up counting [5]. However, the counter shows poor performance in terms of real-time processing and detection efficiency. Liu Yile proposed a method using hand height instead of horizontal bar contours for detection. By improving the OpenPose network and designing a key point detection model, automatic monitoring of pull-ups was achieved [6]. Wang Cheng proposed a pull-up counting method based on a convolutional neural network, which solves the issue of insufficient detection accuracy for illegal pull-up movements by tracking changes in the key points of the participant's skeleton during the pull-up process [7]. Bao Ziqun proposed a sit-up test counting algorithm based on an enhanced target detection network to address the limitations in real-time performance associated with deep learning technology in the context of sit-up testing [8]. Therefore, exploring a sit-up detection and counting method that combines deep learning algorithms and traditional image processing techniques has become the optimal solution to address existing challenges.

In recent years, deep convolutional neural networks (CNNs) have demonstrated significant advantages in target detection, especially in accurately identifying key points of the human skeleton in complex environments [8]. These methods can be broadly categorized into two types: two-stage detection methods based on region proposals, such as R-CNN, Fast R-CNN, and Faster R-CNN, which offer high accuracy but poor real-time performance, and single-stage detection methods like SSD and YOLO, which predict target locations and attributes directly, providing better real-time performance and robustness. With the rapid development of the YOLO detection algorithm, numerous skeletal key point recognition techniques and methods based on YOLO and high-resolution image data have emerged. In particular, the rapid advancement of deep convolutional neural networks has significantly improved the accuracy of target detection and classification tasks. YOLO (You Only Look Once), as a representative of single-stage target detection algorithms, has significant advantages in terms of speed and accuracy compared to multi-stage detection algorithms [9,10]. To address the issue of high computational complexity caused by excessive model parameters, the YOLO network structure, based on depthwise separable convolution, improves the speed and accuracy of sit-up detection by reducing the number of parameters, reusing multi-layer features, and optimizing the loss function, thereby promoting the practical implementation of the sit-up counting system. Against this backdrop, this paper proposes a sit-up counting method based on the YOLO detection algorithm to solve the problems of high labor consumption and low accuracy in existing manual counting methods. By improving the network structure, introducing skip connections, and optimizing the loss function, the accuracy of skeletal key point positioning and counting efficiency were effectively enhanced, providing an efficient and objective counting tool for sports and physical testing [11,12].

Demand positioning and counting system construction

Combined with the needs of daily sit-up exercises and testing, the system should meet the following functional requirements:

(1) Accurate sit-up counting: The system should automatically detect and record each standardized sit-up performed by the subject, filter out non-standard invalid attempts, and ensure the accuracy of the final test results. To achieve this, the system utilizes an improved convolutional pose network, which learns and represents spatial information and texture features in the image through serial processing and a multi-stage convolutional neural network architecture. To address the common issue of gradient vanishing in deep networks, supervised training is introduced at each stage [13]. The improved convolutional pose machine network is used to locate the subject's key skeletal points during the sit-up process, ensuring both detection accuracy and real-time performance. The algorithm flow of the convolutional pose machine network is shown in Fig. 1.

(2) Data recording and management: The system should record and automatically save

the subject's basic information, including the number of valid and invalid sit-ups completed. The system should quickly filter and query any subject's test data, including valid counts, invalid counts, full test videos, and personal information, facilitating subsequent analysis and evaluation by the tester or manager [14,15].

The sit-up counting system proposed in this paper consists mainly of the following modules: data acquisition module, sit-up detection module, data display module, data storage module, and data query and printing module [16,17]. The overall architecture of the system is shown in Fig. 1. Additionally, the system hardware is supported by a CNC universal gear grinding machine (YW7232CNC) for the precision processing of gears, ensuring the system's stability and high accuracy.

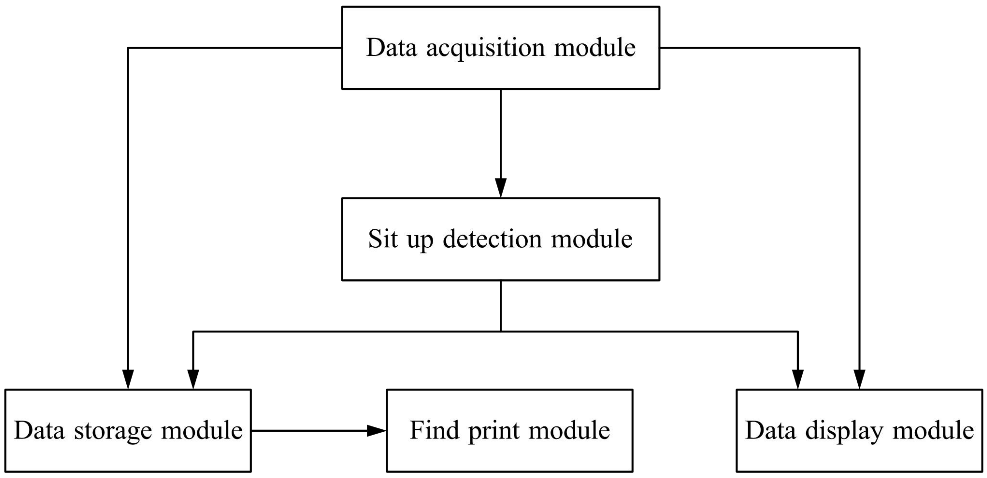

Figure 1. Overall architecture of the sit-up counting system.

Counting method and process

The sit-up counting method is based on human skeleton key point detection technology. The system defines evaluation criteria based on the changes in joint angles during the sit-up process. By extracting key points such as the shoulder, hip, knee, and ankle joints, the system calculates the sit-up angle and determines if the movement meets the standard based on the angle changes of these four parts [8,18]. Specifically, the system transforms the coordinates of human key points obtained through pose estimation and calculates the action angles θ_1 and θ_2 of the sit-up. When angle θ_1 is greater than or equal to 170° and θ_2 is greater than or equal to 120°, the system determines that the subject is in a supine position. If θ_1 is less than or equal to 90° and θ_2 is less than or equal to 30°, the system determines that the subject is in a sit-up position. When both conditions are met consecutively, the system recognizes the action as a valid sit-up and counts the number of valid sit-ups completed within one minute, outputting the final result. This sit-up counting method, based on skeleton key point detection, accurately records the number of sit-ups performed, providing an effective technical tool for physical testing, fitness training, and physical assessment [19]. The flowchart of the proposed counting method is shown in Fig. 2.

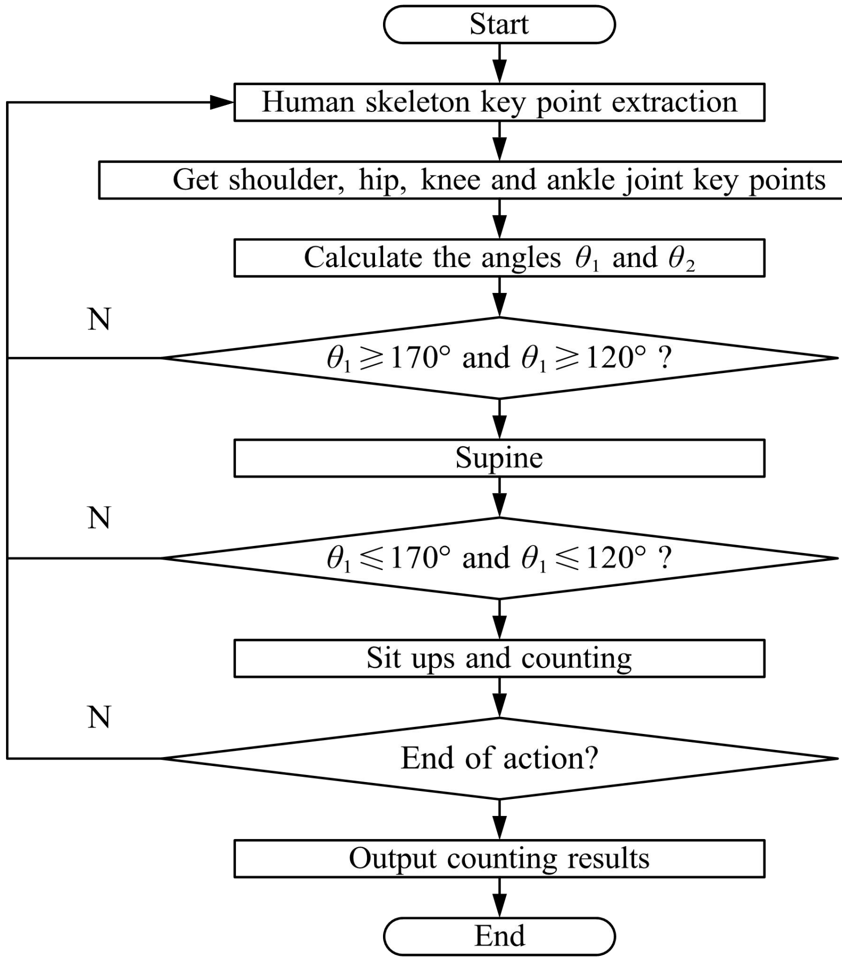

Figure 2. Flowchart of sit-up counting method based on skeleton key point detection.

Skeleton key point detection method and its evaluation index

Detection Algorithm

Early skeleton key point detection methods primarily relied on geometric prior pattern matching techniques. However, due to the variability of human postures, their effectiveness was limited. With the advancement of technology, methods based on two-dimensional human skeleton key point prediction have been widely applied across various fields. Current mainstream methods are divided into two main categories: top-down and bottom-up. The top-down approach first performs individual recognition followed by key point detection, making it suitable for scenes with a small number of individuals. The bottom-up approach detects all key points first and then clusters them, making it more suitable for scenes with a large number of individuals [7,20].

YOLOv7-Pose combines the advantages of these two methods by directly regressing to key points, significantly improving processing speed. YOLOv7-Pose is an extension of the YOLOv7 deep learning model, specifically designed for pose estimation. This method optimizes target and key point detection performance through multi-task learning and consists of three main components: a backbone network, a detection head, and a key point detection head. The backbone network (e.g., CSP Darknet) extracts image features, the detection head predicts the bounding box, category, and confidence of the target, and the key point detection head predicts the coordinates of human body key points [6,21].

The total loss function of YOLOv7-Pose consists of three components: bounding box loss, classification loss, and key point loss. Assuming we have N targets, each with K key points, the total loss function can be expressed as:

$$L = L_{bbox} + L_{cls} + L_{kpt} \tag{1}$$

The bounding box loss, L_{bbox}, consists of position and size losses, typically measured using complete intersection-over-union (CIOU):

$$L_{bbox} = \sum_{i=1}^{N} \left(1 - \mathrm{CIOU}(b_i, \hat{b}_i\right) \tag{2}$$

Here, b_i and $\hat{b}_i$ are the true and predicted bounding boxes of the i-th object, respectively.

The classification loss, L_{cls}, measures the difference between the predicted and true categories, typically using cross-entropy loss:

$$L_{cls} = -\sum_{i=1}^{N} \sum_{c=1}^{C} y_{i,c} \log\left(p_{i,\mathrm{c}}\right) \tag{3}$$

Among them, $y_{i,c}$ is the true category label of the i-th target, and $p_{i,\mathrm{c}}$ is the predicted category probability.

The key point loss, L_{kpt}, measures the difference between the predicted and true key point coordinates, typically using mean square error (MSE):

$$L_{kpt} = \sum_{i=1}^{N} \sum_{K=1}^{K} MSE((x_{i,k}, \; y_{i,k}), (\hat{x}_{i,k}, \hat{y}_{i,k})) \tag{4}$$

Where, ($x_{i,k}$, $y_{i,k}$)and ($\hat{x}_{i,k}$, $\hat{y}_{i,k}$) are the true and predicted coordinates of the k-th key point of the target, respectively.

Using these methods, YOLOv7-Pose efficiently performs target detection and human pose estimation, making it suitable for various applications requiring real-time performance and high-precision pose estimation. Additionally, the PersonLab method combines skeleton key point detection and instance segmentation technologies, further enhancing pose estimation accuracy [22,23]. The convolutional pose network uses part response maps to represent the spatial constraints between body parts and facilitates multi-stage data exchange by transferring response maps and feature maps between stages. To address the gradient vanishing problem, the network employs a multi-stage repeated supervision mechanism, expanding coverage as network depth increases and using larger convolution kernels to maintain high detection accuracy even when body parts are occluded. [9,24]. The combined application of these methods has significantly improved the accuracy and robustness of human pose estimation, driving technological advancements in this field.

Algorithm Model

The convolutional pose machine is a multi-stage sequential convolutional neural network specifically designed to extract spatial and texture features from images [7]. Using its multi-stage sequential network structure and supervised training, the convolutional pose machine has demonstrated strong performance in human pose estimation [11,17].

In this model, the pixel position of the *P*-th anatomical landmark can be represented by $\mathrm{Y}_P \in \mathrm{Z} \subset \mathbb{R}^2$, where *Z* is the set of all (u,v) positions in the image, and the goal is to predict all image locations $Y = (Y_1, \ldots Y_P)$. The pose machine consists of a series of multi-class predictors $g_t(\cdot)$ that are trained to predict the position of each part in each hierarchy [14]. At each stage $t \in \{1 \ldots T\}$, the classifier g_t predicts a position confidence $\mathrm{Y}_P = z, \forall z \in Z$ for each part based on the features extracted from the image at position *z* denoted by $X_z \in \mathbb{R}^d$, and the contextual information provided by the previous classifier near each y at stage t:

$$g_1(x_z) \rightarrow \left\{ b_1^p (Y_p = z) \right\}_{p \in \{0 \ldots P\}} \tag{5}$$

where $b_1^p (Y_p = z)$ is the prediction score of the pth component at image position *z* by classifier g_1 in the first stage. Let $b_t^p \in \mathbb{R}^{w \times h}$ denote all the confidences of the *p*-th component evaluated at each position $z = (u,v)^T$ in the image, where *w* and *h* are the width and height of the image respectively:

$$b_t^p[u,v] = b_t^p\left(Y_p = z\right) \tag{6}$$

To simplify the representation, the mapping set of all parts is represented as $b_t \in {}^{w\times h\times(P+1)}$.

In subsequent stages, the classifier predicts a belief $\mathrm{Y}_P = z, \forall z \in Z$ about the location of each part based on (1) the features of the image data $x_z^t \in {}^{d}$, and (2) contextual information from the previous classifier in the neighborhood around each Y_p:

$$g_t\left(x'_z, \psi_t\left(z, b_{t-1}\right)\right) \to \left\{b_t^p\left(Y_p = z\right)\right\}_{p\in\{0\ldots P+1\}} \tag{7}$$

Where a is $\psi_{t>1}(\cdot)$ mapping of belief b_{t-1} to context features. At each stage, the computed belief provides increasingly accurate estimates of the location of each part. Note that the image features used in subsequent stages are allowed to be different from those used in the first stage. The prediction machine proposed by V. Ramakrishna uses boosted random forest to predict $\left(\{g_t\}\right)$, uses fixed hand-crafted image features $\left(x' = x\right)$ at all stages, and uses fixed hand-crafted context feature maps $\left(\psi_t(\cdot)\right)$ to capture spatial context information at all stages [17].

This paper optimizes the original convolutional pose machine network for human targets captured by a camera in single-person measurement scenarios. While the original network performs well in detecting skeletal joints during sit-ups, its detection speed fails to meet real-time requirements when run on standard computing devices. To address this, the proposed improved algorithm optimizes the network in two ways: First, the network architecture is simplified from six stages to four, reducing iterations, the number of parameters, and model size, thereby improving computational efficiency. Second, skip connections are introduced between stages to enhance data-sharing capability [7,25]. In the original network, each stage relied solely on the output of the immediately preceding stage and the features extracted from the original image. The addition of skip connections allows the network to retain more low-level details from the original image and effectively utilize these details when processing higher-level features, thereby improving the network's expressive power [17]. As shown in Fig. 3, the improved structure demonstrates how the deep convolutional architecture replaces the prediction and image feature calculation modules of the pose machine, allowing direct learning of image and context feature representations from data. Additionally, the convolutional architecture's fully differentiable nature allows for end-to-end joint training of all pose machine stages [26].

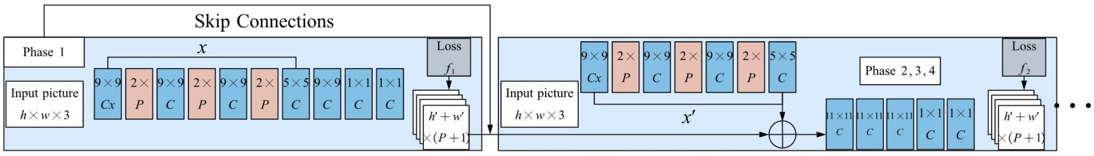

Figure 3. Flowchart of the improved convolutional pose machine algorithm.

Evaluation Metrics

This paper presents a sit-up counting method based on the angle variation of key skeletal joints. The method evaluates whether the sit-up is performed to standard by detecting changes in the angles of the shoulder, hip, knee, and ankle joints. As illustrated in Fig. 4, given the unique nature of the sit-up movement, a customized joint angle detection method is used. A sit-up is deemed valid only when the angles of the shoulder, hip, knee, and ankle joints simultaneously meet specific criteria. Specifically, when angle θ_1 is greater than or equal to 170° and θ_2 is greater than or equal to 120°, the system classifies the tester as being in the supine position. If θ_1 is less than or equal to 90° and θ_2 is less than or equal to 30°, the system classifies the tester as being in the sitting position [11,27]. A valid sit-up is only recorded when both conditions are met in sequence. These angle values are calculated by processing the key point coordinates obtained from pose estimation. The system counts the number of sit-ups performed by the tester within one minute, excluding actions that do not meet the standard from the final count [28].

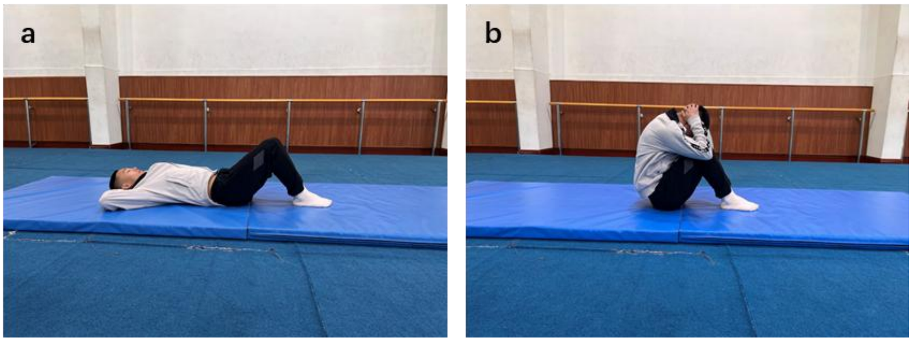

Figure 4. Sit-up technique diagram: (a) Lie flat on your back with your upper body touching the mat; (b) Elbows touching knee joints.

The formula for calculating joint angles is as follows: Let the positions of the shoulder, hip, knee, and ankle joints be represented by points A(x_1,y_1),B(x_2,y_2),C(x_3,y_3) and D(x_4,y_4) respectively. The angle to be measured is calculated based on the coordinates of these points.

$$\theta_1 = arccos\left(\frac{\vec{AB}\cdot\vec{BD}}{|\vec{AB}|\times|\vec{BD}|}\right) \qquad (8)$$

$$\theta_2 = arccos(\frac{\vec{AB} \cdot \vec{BC}}{|\vec{AB}| \times |\vec{BC}|}) \tag{9}$$

Where, $\vec{AB}$, $\vec{BC}$ and $\vec{BD}$ are the vectors from point A to B, B to C, and B to D respectively.

(1) Average Precision (AP): The AP calculation for single-person pose estimation is based on Object Keypoint Similarity (OKS), a specific metric used to assess the accuracy of keypoint detection. OKS provides a standardized method for comparing predicted keypoints with true keypoints, accounting for factors such as scale, unlabeled keypoints, and ambiguity in annotations. The AP calculation formula is:

$$AP = \frac{\sum_p \delta\left(oks_p > T\right)}{\sum_p l} \tag{10}$$

Where T is a given threshold set.

(2) Mean average precision (mAP): mAP is a commonly used indicator for evaluating the overall performance of multi-category or multi-task detection models. It is the result of averaging the average accuracy (AP) of all relevant nodes. The calculation formula is as follows:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{11}$$

Where N is the number of all key points.

(3) Count accuracy (CA): Count accuracy (CA) is an evaluation indicator used to measure the accuracy between the actual number of sit-ups measured and the true number. The calculation formula for counting accuracy is:

$$CA = \frac{TestNum}{TrueNum} \times 100\% \tag{12}$$

Where, *TrueNum* represents the actual number of sit-ups, and *TestNum* represents the measured number.

Experiments and Analysis

Experimental Test

To validate the performance of the sit-up counting method proposed in this paper in practical applications, a series of field experimental tests were conducted. The test subjects included students of different grades and ages, comprising 6 first-year junior high school girls, 6 first-year high school girls, and 7 first-year undergraduate boys. The primary objective of the experiment was to compare the accuracy of sit-up counting between the proposed method and manual referees.

In the experiment, the number of sit-ups performed by the subjects was classified into two categories: valid completions and invalid fouls. A valid sit-up required strict adherence to the prescribed rules: the shoulder blades must touch the mat when lying down, hands must be held behind the head, elbows must touch the knees when sitting up, and the buttocks must remain on the mat. Failure to meet any of these criteria—such as shoulder blades not touching the mat, hands not behind the head, elbows not touching the knees, or buttocks lifting off the mat—would result in the action being judged as a foul and counted as invalid by the system. To ensure the accuracy and fairness of the test results, three professionally qualified and experienced referees were present to make on-site judgments. The core of the experimental design was to compare the results assessed by manual referees with those obtained using the sit-up counting method proposed in this paper. This comparison aimed to evaluate the accuracy and reliability of the proposed method in real-world scenarios. By testing students of different grades and physical conditions, the applicability of the counting method could be verified, providing data support for future method improvements.

Results and Analysis

Based on the model reliability evaluation metrics, Tab. 1 compares the changes in network parameters before and after the optimization. The improved network demonstrates significant optimization in several key performance indicators: the number of parameters has been substantially reduced from the original 125.6M to 38.3M. This reduction highlights the effectiveness of the optimization in decreasing the model's complexity and lowering the consumption of computational resources. Furthermore, the processing speed of the network increased significantly, from 39.5 FPS (frames per second) to 63.6 FPS. This improvement indicates that the optimized network performs well in real-time processing, enabling faster data processing and thus enhancing the system’s response speed and overall efficiency. Overall, although the accuracy decreased slightly after optimization, the performance remained at a high level. The optimized network achieves faster processing speeds and reduced complexity while maintaining sufficient accuracy to meet practical application needs.

Table 1. Comparison of network parameters before and after improvement.

Algorithm Model	Model Size/MB	m AP@0.5	m AP@0.5:0.95	FPS	CA/%
Convolutional pose machine network	125.6	95.3	76.5	39.5	99.5
Improved network	38.3	94.5	74.4	63.6	98.57

The comparison of system detection performance and network training efficiency with existing methods is presented in Tab. 2. The results of the sit-up test experiment and manual referee judgments are shown in Tab. 3. The results indicate that, although the algorithm used in this study is low-cost and based on a small self-constructed dataset, it achieves high accuracy and meets the expected objectives. The sit-up counting method proposed in this paper effectively addresses two common foul issues often encountered in subjective judgments by manual referees: failure to have the shoulder blades touch the pad or hands hold the head while lying down, and failure to have the elbows touch the knees or the hips stay on the pad while sitting up. In 207 valid sit-up tests performed by students across different grades, the system recorded only 3 counting errors, achieving an overall accuracy of 98.57%. This demonstrates that the sit-up counting method based on machine vision can meet the accuracy requirements necessary for field testing. Additionally, the data in Tab. 3 indicate that the standardization of body posture and movement speed significantly affect the accuracy of sit-up test results. Differences in the judgment of the same test results among different manual referees reflect the subjectivity inherent in manual assessment. Moreover, slight discrepancies between the system records and manual judgments may be attributed to technical factors or the subjectivity of the referees.

Table 2. Evaluation results of different network algorithms.

Network Algorithms	Dataset	Accuracy /%
Reference 3	Self-built	94.00
Reference 4	MPII	99.21
Reference 5	Self-built	96.40
This article	Self-built	98.57

The sit-up counting system not only demonstrates high accuracy but also effectively illustrates the key aspects of sit-ups, the muscle groups involved, and the impact of strengthening exercises. The system's image and video playback features, along with its data analysis modules, play a critical role in sit-up testing. It reduces both the time required for score entry and the likelihood of errors, while eliminating the subjectivity of manual refereeing, allowing users to review their individual scores and overall performance at any time. The system allows for clear analysis of performance at specific stages and enables longitudinal comparisons of repetitions and movement quality across different stages. These features ensure the system's excellent stability and reliability, making it well-suited to meet the diverse requirements of sit-up testing.

Table 3. Experimental test results of sit-up counting method and manual referee judgment results.

Category	Serial number	Total times	Effective completion times	Invalid completion times	Counting system test results	Manual referee 1	Manual referee 2	Manual referee 3
First grade	1	10	8	2	8	8	8	8
	2	10	9	1	9	9	9	9
	3	11	10	1	10	10	10	10
	4	8	5	3	5	5	5	4
	5	12	10	2	10	11	10	10
	6	7	7	0	7	7	7	7
Senior Year	1	13	10	3	10	10	10	11
	2	11	10	1	10	10	10	10
	3	15	12	3	13	12	12	12
	4	13	12	1	12	12	12	11
	5	16	15	1	15	15	15	15
	6	9	8	1	8	8	8	8
First-year undergr-aduate	1	18	15	3	15	15	15	15
	2	16	16	0	16	16	16	16
	3	15	12	3	13	12	12	12
	4	20	19	1	19	19	19	20
	5	13	10	2	11	10	9	10
	6	9	8	1	8	8	8	8
	7	11	11	0	11	11	11	11
Total		237	207	29	210	208	206	207

Exercise recommendations

First, it is crucial to recognize the sensitive period for core muscle strength development in students, particularly those in primary and secondary schools. Strengthening core muscles, including the abdomen and back, establishes a foundation for students' overall physical development. Second, health education should be further deepened, guiding students to develop scientifically informed living habits. Encouraging daily core exercises, such as sit-ups and push-ups, along with diverse training activities and competitions, can stimulate students' interest in physical fitness and promote greater participation. Third, the dissemination of theoretical knowledge and practical methods for core muscle training, such as sit-ups, should be expanded. Special lectures, campus radio, official school accounts, and social media platforms should be used to educate students on the correct posture, breathing control, intensity, and frequency of training. Additionally, students should be taught techniques to safely and effectively strengthen core muscles, including how to avoid neck strain and coordinate breathing to enhance performance.

Conclusion

(1) This article presents a sit-up counting method utilizing bone key point detection to enhance the objectivity and efficiency of traditional manual counting methods. By optimizing the network structure and incorporating skip connections, the system achieves a sit-up test accuracy of 98.57%, thereby meeting both real-time and precision requirements.

(2) The optimization of the network reduced the number of parameters from 125.6 million to 38.3 million, substantially decreasing model complexity. Concurrently, the processing speed improved from 36.5 FPS to 63.6 FPS, demonstrating enhanced processing efficiency. These advancements boost speed and optimize the system's practical performance while preserving high accuracy.

(3) The system integrates skeletal key point tracking with illegal action detection to enable real-time counting. It features capabilities for real-time detection, data collection, display, storage, and search, offering an efficient and objective tool for sports and testing applications.

Authors' contribution

Conceptualization: Zhiming Shi, Hang Zhao
Methodology: Zhiming Shi, Guangxin Cheng
Check: Zhiming Shi, Hang Zhao
Data Curation: Zhiming Shi, Junjie Chen
Writing-Rough Preparation: Zhiming Shi, Hang Zhao, Junjie Chen
Writing-Review and Editing: Zhiming Shi, Hang Zhao
Visualization: Zhiming Shi, Junjie Chen

All authors have read and agreed with the published version of the manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding statement

This paper is funded by the Ministry of Education Industry-University Cooperation Collaborative Education Project (230902371264520) and Chongqing Education Commission Science and Technology Research and Technology Youth Fund Project (KJQN202400208).

Institutional review board statement

Not applicable.

Informed consent statement

Not applicable.

Data availability statement

Not applicable.

Acknowledgments

Special thanks to Southwest University for providing high-end scientific research equipment such as the CNC universal gear grinding machine (YW7232CNC) and the five-axis linkage vertical machining center (I-600). These equipment laid a solid foundation for the smooth progress of the experiment and the accurate acquisition of data.

References

1.Li Chaoyang, Zhang Xinping. Discussion on optimization of sit-up test item in the National Student Physical Fitness Standards[J]. Curriculum Teaching Research, 2016, (12): 68-72. https://doi.org/10.3969/j.issn.2095-2791.2016.12.016.

2.Bin-Bin H, Li-Qian H, Wen-Bo S H U, et al. Influence of fat and muscle on sit-ups in female college students[J]. Acta Anatomica Sinica, 2021, 52(5): 822. https://doi.org/10.16098/j.issn.0529-1356.2021.05.023.

3.Lemes V B, Brand C. Disclosing core strength: reliability and correlations before and after COVID-19 pandemic in plank and sit-ups tests[J]. Sport Sciences for Health, 2024, 20(1): 249-257. https://doi.org/10.1007/s11332-023-01136-x.

4.Dr. Rohit A Tambe. Establishment of norms for sit ups test of higher secondary students of Maharashtra state. Int J Phys Educ Sports Health 2022;9(4):27-30.

5.Hua Zhiyuan, Wang Yizhi, Li Quanbin. Design of portable intelligent sit-up counter[J]. Internet of Things Technology, 2024, 14(01): 73-76. https://doi.org/10.16667/j.issn.2095-1302.2024.01.020.

6.Liu Yile, Tu Yaqing. Key point detection and monitoring method for pull-ups in video images[J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2022, 34(06): 995-1004. https://doi.org/10.3979/j.issn.1673-825X.202204300098.

7.Wang Cheng. Design and implementation of pull-up intelligent counting system based on human skeleton key point detection[D]. Nanjing University of Posts and Telecommunications, 2021. https://doi.org/ 10.27251/d.cnki.gnjdc.2021.001106.

8.Bao Ziqun. Improving sit-up test counting for target detection networks[J]. Intelligent Computers and Applications, 2022, 12(08): 102-105+109.

9.Wen Qian. Research and implementation of human posture estimation system based on convolutional neural network[D]. Xidian University, 2019. https://doi.org/ 10.27389/d.cnki.gxadu.2019.002472.

10.Yan Xiaobo, Pan Feng, Zhang Yansha, et al. Pull-up intelligent detection technology based on image vision[J]. Industrial Control Computer, 2023, 36(06): 46-47+49.

11. Jing Qian, Tao Qingchuan. Sit-up counting algorithm based on computer vision[J]. Modern Computer, 2023, 29(07): 69-72.

12.Zhang Yuxin. Design and implementation of pull-up test system[J]. Electronic Test, 2014, (21): 24-25.

13.Liu Shidong. Common problems and solutions in the test of physical education middle school entrance examination instruments: taking pull-ups and sit-ups as examples [J]. Physical Education, 2022, 42(06): 81.

14.Yue Liying. Design and implementation of human behavior recognition system based on posture estimation and graph convolutional neural network[D]. Shihezi University, 2023. https://doi.org/ 10.27332/d.cnki.gshzu.2023.000587.

15.Rätsch G, Onoda T, Müller K R. Soft margins for AdaBoost[J]. Machine learning, 2001, 42: 287-320. https://doi.org/10.1023/A:1007618119488.

16.Papandreou G, Zhu T, Chen L C, et al. Personlab: Person pose estimation and instance segmentation with a bottom-up, part-based, geometric embedding model[C].Proceedings of the European conference on computer vision (ECCV). 2018: 269-286. https://doi.org/10.1007/978-3-030-01264-9_17.

17.Lv X, Ta N, Chen T, et al. Analysis of gait characteristics of patients with knee arthritis based on human posture estimation[J]. BioMed research international, 2022, 2022. https://doi.org/10.1155/2022/7020804.

18.Chen Y, Wang Z, Peng Y, et al. Cascaded pyramid network for multi-person pose estimation[C].Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7103-7112. https://doi.org/10.1109/CVPR.2018.00742.

19.Zhang Lijie, Li Xiuting. Research and development of campus smart sports equipment-design of intelligent sit-up counter[J]. China Modern Educational Equipment, 2022, (10): 37-39. https://doi.org/10.13492/j.cnki.cmee.2022.10.010.

20. Zhu Z, Gao X, Shi Y, et al. Standardized detection and counting of sit-ups based on CenterNet[C]. AIIPCC 2022; The Third International Conference on Artificial Intelligence, Information Processing and Cloud Computing. VDE, 2022: 1-7.

21.Zhao W D, Zhang Q Q, Xue Q, et al. Lightweight sit-ups recognition and counting method based on openpose[C].2022 4th International Conference on Frontiers Technology of Information and Computer (ICFTIC). IEEE, 2022: 681-685.

22.Hutahaean A D. Evaluation of the performance of five pre-trained deep learning networks for automatic counting of sit-ups in Indonesia civil servant technical tests. Masters thesis, Universitas Pelita Harapan[D]. Universitas Pelita Harapan, 2023.

23.Anders C, Ludwig F, Sänger F, et al. Eight Weeks Sit-Up versus Isometric Abdominal Training: Effects on Abdominal Muscles Strength Capacity[J]. Arch Sports Med, 2020, 4(1): 198-204. https://doi.org/10.36959/987/252.

24.Rifki M S, Farma F, Komaini A, et al. Development of Sit Up Measuring Tools Based on Arduino and Ultrasonic Sensors With Android Applications[J]. Int. J. Interact.Mob.Technol,2022,16(8):182-189.https://doi.org/10.3991/ijim.v16i08.30673.

25.Griban G, Yahupov V, Svystun V, et al. Dynamics of the students' physical fitness while studying at higher educational institutions[J]. International Journal of Applied Exercise Physiology, 2020 9 (9): 147-156.

26.Mardela R, Irawan R, Laksana A A N P, et al. Development of a Digital-Based Push Up and Sit Up Test Counter[J]. Halaman Olahraga Nusantara: Jurnal Ilmu Keolahragaan, 2023, 6(1): 287-298. https://doi.org/10.31851/hon.v6i1.10723.

27.Wang H, Chen G, Yang F, et al. The Design and Simulation of Sit-Up Counter Based on MCU[C]. Advancements in Mechatronics and Intelligent Robotics: Proceedings of ICMIR 2020. Springer Singapore, 2021: 459-464. https://doi.org/10.1007/978-981-16-1843-7_53.

28.Jang S Y, So W Y, Jeong T. Intelligent system of training data sets for current reported normality levels and physical fitness analysis[J]. Science & Sports, 2021, 36(4): 310. e1-310. e6. https://doi.org/10.1016/j.scispo.2020.10.001.