Heliyon 10 (2024) e30055

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Design of sports achievement prediction system based on U-net convolutional neural network in the context of machine learning

Guoliang Wang, Tianping Ren *

College of Sport, Henan Polytechnic University, Jiaozuo, Henan, 454003, China

ARTICLE INFO

Keywords:
Machine learning
U-Net convolutional neural network
Achievement prediction
Dense connection
Attention module
Residual learning

ABSTRACT

Sports plays a pivotal role in national development. To accurately predict college students' sports performance and motivate them to improve their physical fitness, this study constructs a sports achievement prediction system by using a U-Net Convolutional Neural Network (CNN) in machine learning. Firstly, the current state of physical education teachers' instructional proficiency is investigated and analyzed to identify existing problems. Secondly, an improved U-Net-based sports achievement prediction system is proposed. This method enhances the utilization and propagation of network features by incorporating dense connections, thus addressing gradient disappearance issues. Simultaneously, an improved mixed loss function is introduced to alleviate class imbalance. Finally, the effectiveness of the proposed system is validated through testing, demonstrating that the improved U-Net CNN algorithm yields superior results. Specifically, the prediction accuracy of the improved network for sports performance surpasses that of the original U-Net by 4.22 % and exceeds that of DUNet by 5.22 %. Compared with other existing prediction networks, the improved U-Net CNN model exhibits a superior achievement prediction ability. Consequently, the proposed system enhances teaching and learning efficiency and offers insights into applying artificial intelligence technology to smart classroom development.

1. Introduction

Sports achievement prediction is crucial for improving the pertinence of college students' sports training and aiding educational administration departments in formulating sports training plans [1-3]. Therefore, achieving high-performance sports achievement prediction methods has become a hot research issue. Initially, students' sports achievement prediction relied on physical education (PE) teachers' subjective assessments and basic statistical methods, resulting in low prediction efficiency and accuracy, thus failing to meet practical needs [4,5]. This study is motivated by the pressing need to enhance PE quality and accurately predict students' sports performance. Given the increasing emphasis on sports in holistic education, precise prediction of students' sports performance is essential. It not only assist teachers in better assessing students' physical fitness and athletic skills but also facilitates formulating personalized training plans and improves teaching effectiveness. However, traditional assessment methods often rely on subjective judgment and simple statistical analyses, lacking precision and objectivity. Additionally, existing prediction models face limitations in handling complex sports data and addressing class imbalance issues.

Predicting students' sports achievement involves modeling a regression problem, with numerous approaches proposed to address it. Jaiswal et al. (2021) [6] emphasized the importance of timely interventions based on students' academic records to accurately

* Corresponding author.
E-mail address: rtp@hpu.edu.cn (T. Ren).

https://doi.org/10.1016/j.heliyon.2024.e30055
Received 14 January 2024; Received in revised form 16 April 2024; Accepted 18 April 2024
Available online 1 May 2024

forecast their upcoming achievements. They established a machine learning system to evaluate students' graduation results. Joshi et al. (2021) [7] highlighted educational data mining (EDM) as a major field in modern education research, emphasizing the significance of data quality in determining the success of a data mining strategies. Their approach improved prediction accuracy by identifying and explaining factors influencing student achievement. Rodríguez-Hernández et al. (2021) [8] explored the application of Artificial Neural Networks (ANNs) in predicting higher education performance. They found it outperformed other machine learning algorithms in terms of F1-score and recall. They also identified previous academic achievements, high school characteristics, and socio-economic conditions as vital predictors. In order to find out the extent to which mental intelligence and anxiety contribute to predicting the academic performance of college students, Singh (2021) [9] investigated the contribution of mental intelligence and anxiety to 100 college students' academic achievement, finding a positive correlation between mental intelligence and academic achievement. Meanwhile, they discovered that anxiety was not correlated with mental intelligence but played a role in predicting it. Monteverde-Suárez et al. (2021) [10] compared students' academic achievement prediction using ANN and Naive Bayes in a medical undergraduate program's first year. They identified attributes associated with predictions, with Naive Bayes highlighting students' prior knowledge as the most vital prediction attribute. In summary, scholars have explored various approaches to achievement prediction, yet challenges persist. Few studies focus on the academic performance prediction of college students, with primary and secondary school students often receiving more attention. Additionally, some researchers overlook multifactorial influences on performance prediction.

College students' sports achievement prediction poses a typical uncertainty challenge. This study proposes a sports achievement prediction method for students by combining U-Net Convolutional Neural Network (CNN) with DenseNet. Initially, a U-Net CNN model for predicting students' sports achievement is constructed. Subsequently, to address model limitations, residual learning mechanisms and densely connected neural networks are incorporated. Finally, experimental data validates the effectiveness of the prediction model.

1.1. Study on sports achievement prediction scheme using neural network

1.1.1. Challenges confronting PE teachers in the context of machine learning

PE teachers face various challenges throughout their professional journey, spanning instructional complexities, student-related concerns, curriculum content, and external environmental factors.

Firstly, there is an increasing demand for personalized teaching in PE. PE teachers must tailor training plans to each student's physical fitness, interests, and learning styles. However, achieving effective personalized teaching without sufficient data support poses significant challenges. Teachers need to understand each student's characteristics and possess the ability to access and analyze relevant data to support their teaching decisions. Secondly, data-driven decision-making is increasingly crucial in PE. Utilizing data analysis to guide improvements in students' sports performance and maintain their health represents a major trend in modern teaching. However, PE teachers may lack experience and tools for effectively collecting, analyzing, and applying this data. Furthermore, performance prediction and goal setting are crucial components of PE teaching. Teachers need to set reasonable training goals and accurately predict student performance improvement. Without precise and scientific prediction tools, these goals and predictions often rely on teachers' experiential judgment, which may not be accurate or effective. Moreover, the challenge of technological integration should not be underestimated. While advanced machine learning technologies offer new possibilities for PE, teachers may face difficulties integrating these technologies into their daily teaching practices. This includes a lack of necessary technical equipment, the need for relevant technical knowledge, and strategic considerations on how to blend new technologies with traditional PE teaching methods. Lastly, training and resource limitations are also a significant challenge. Effectively utilizing technologies such as U-Net requires teachers to have relevant technical knowledge and may demand additional training and resources. Obtaining such support in resource-constrained environments can be challenging.

To resolve the issues faced by PE teachers, enhancing teaching capabilities and integrating modern technology are essential. Firstly, data collection methods such as surveys, interviews, and observations can be employed to analyze students' sports performance, interests, physical fitness levels, as well as PE teachers' teaching methods, technical skills, and challenges. Secondly, machine learning models can conduct in-depth analysis of the collected data, identifying patterns and trends in students' sports performance and assessing the effectiveness and limitations of teachers' teaching methods. Based on this data analysis, key issues in PE teaching can be identified, such as insufficient personalized teaching, slow curriculum updates, low student engagement, and inadequate integration of resources and technology.

Through this systematic approach, combining machine learning technology with improved teaching capabilities, PE teachers can more effectively identify and address current challenges, thus enhancing teaching quality and fostering students' sports performance and overall health.

1.1.2. Analysis of machine learning technology and probabilistic neural networks

Over recent years, deep learning, a subset of AI and machine learning, has tremendously simplified the workflow of machine learning. The concept of deep learning was introduced by American scholars in the second half of the 20th century [11-13], originally aimed at exploring learners' knowledge acquisition and learning processes. In learning, diverse strategies are employed by learners to master knowledge, categorized into deep learning and shallow learning. Deep learning involves understanding, critical thinking, and problem-solving during the learning process, while shallow learners rely on passive memorization [14,15]. Deep learning surpasses shallow learning in effectiveness, and their comparison is depicted in Fig. 1, portraying the relationship between machine learning and deep learning [16,17].

Deep learning represents a prominent pathway to AI, stemming from machine learning, a technology that uses computer systems to analyze vast datasets and derive underlying patterns [18,19]. As a specialized form of machine learning, deep learning excels in capturing complex hierarchical structures of within data, enabling it to model the world effectively. Its applications span various domains, including speech, video, image processing, and natural language processing, catalyzing significant advancements in science and technology and produced profound influence. Among deep learning architectures, CNNs stand out as pivotal models, particularly since 2012 when they became pivotal in image processing [20-23]. When combined with other technologies, CNNs find utility across diverse fields. Fig. 2 illustrates the development timeline of deep learning.

The basic CNN structure is presented in Fig. 3, showcasing its sequential layers including input, convolution, pooling, and fully connected layers [24-26].

The Probabilistic Neural Network (PNN) is a feedforward neural network that operates based on principles of probability statistics, employing a radial basis function network as its core model. PNN has many merits, such as rapid training, accurate classification, simple learning procedures, and high fault tolerance. Bayes' rule plays a pivotal role in guiding data selection and offers enhanced decision-making capabilities within the input space, thus serving as the theoretical foundation for PNN implementation.

A posterior density $\pi(\theta|x)$ in Bayes' rule is expressed as Eq. (1):

$$\pi(\theta|x) = \frac{\pi(\theta)f(x|\theta)}{\int \pi(\theta)f(x|\theta)d\theta} = \frac{h(x|\theta)}{m(x)} \tag{1}$$

In Eq. (1), $f(x|\theta)$ implies the sample density; $h(x|\theta)$ stands for the joint distribution; $\pi(\theta)$ refers to the sample distribution; $m(x)$ means the marginal distribution.

Bayes in neural networks can be written as Eq. (2):

$$\pi(\lambda|X) = \frac{\pi(\lambda)f(\mathrm{X}|\lambda)}{\int \pi(\lambda)f(\mathrm{X}|\lambda)\mathrm{d}\lambda} \tag{2}$$

In Eq. (2), λ refers to model parameters, and X represents the dataset.

Bayes' rule of decision function d(x) is defined by Eq. (3), with a value range of [−1, 1].

$$\mathrm{d} : X(X \subset R^n) \to y = \{-1, 1\} \tag{3}$$

The loss function is a critical method to evaluate prediction accuracy, closely tied to the decision function and expressed by Eq. (4) and Eq. (5):

$$l(x, y, \mathrm{d}(x)) = \mathrm{l}(y - \mathrm{d}(x)) \tag{4}$$

$$l(\beta) = \begin{cases} 0 & \beta = 0 \\ 1 & \text{others} \end{cases} \tag{5}$$

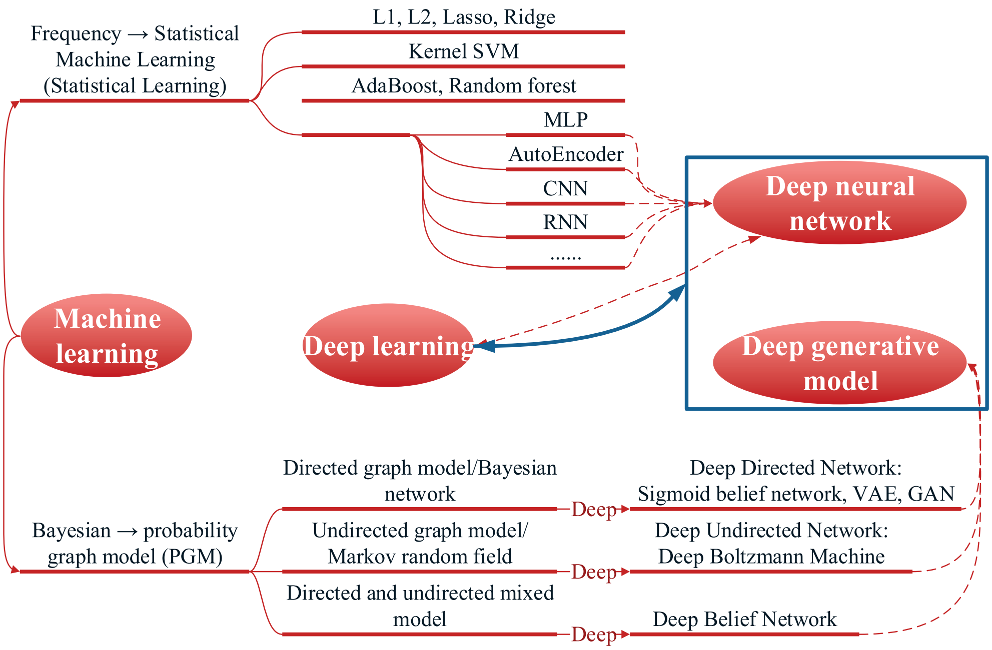

Fig. 1. Relationship between machine learning and deep learning.

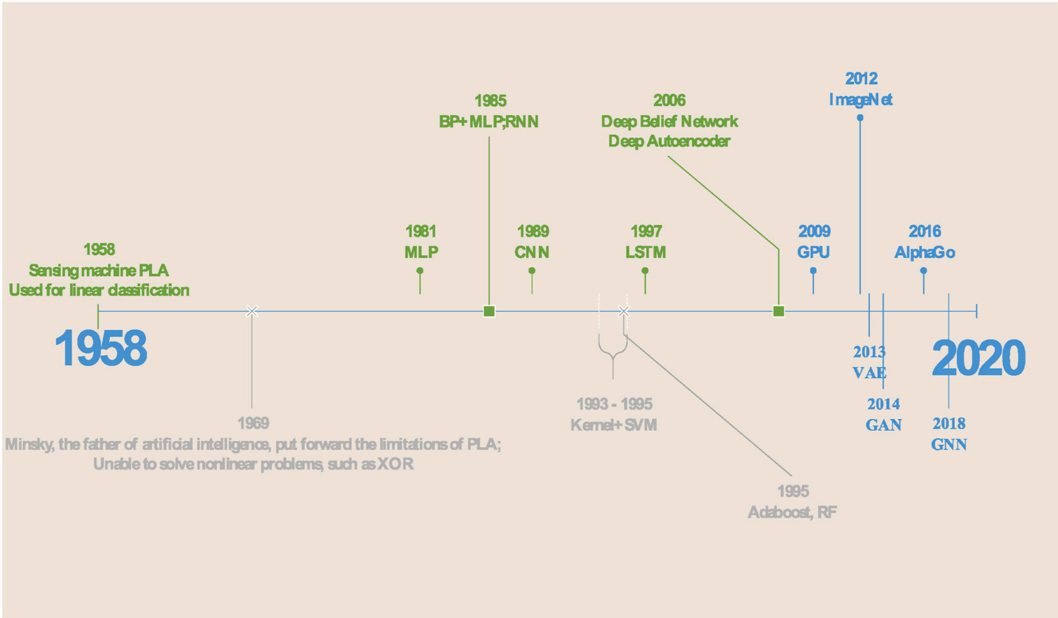

Fig. 2. Timeline of deep learning development.

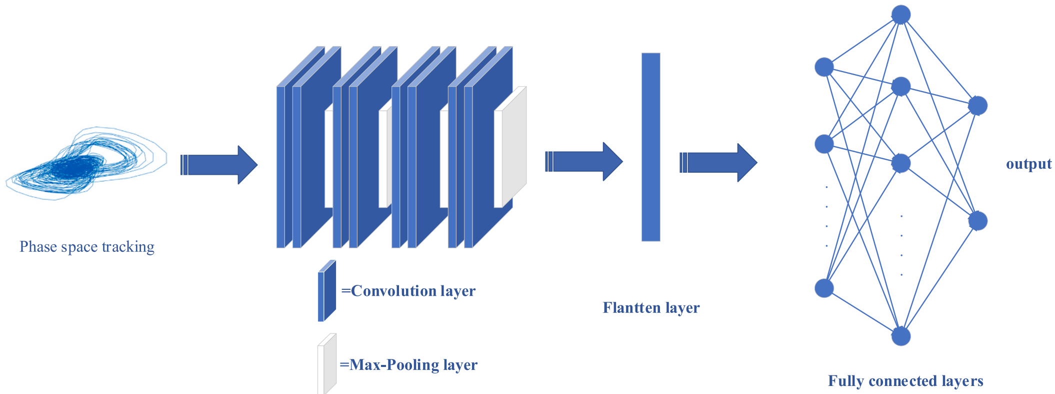

Fig. 3. CNN structure.

The output function, denoted as y, holds significance in various machine learning and deep learning applications, particularly in image processing. These applications often encounter notable class imbalance within datasets, where certain classes have significantly more samples than others. Such imbalance can lead to biases towards the majority class during model training, thereby reducing its ability to recognize minority classes [27]. To mitigate this issue, an improved hybrid loss function is employed, which combines the strengths of various loss functions to handle class imbalance. By assigning different weights to the losses of each class, with increased loss weights for minority classes and decreased weights for majority classes, this weighting mechanism directs the model to focus more on minority classes during training, thereby enhancing its ability to recognize them [28].

The expected risk is an indicator to assess the generalization ability of the decision function. It represents the Riemann-Stieltijes integral of the loss function over a probability distribution, as illustrated in Eq. (6):

$$R[d] = E[l(x, y, d(x))] = \int l\Big((x, y, d(x))dP(x, y) \tag{6}$$

1.1.3. EDM

Data mining involves various disciplines such as statistical analysis, information retrieval, machine learning, and database technology, aimed at extracting knowledge from data to support decision-making. This knowledge can be widely applied to various industries, including banking, insurance, telecommunications, retail, education, and medicine, spawning new research topics. In education, the proliferation of campus information technology and the Internet propelled the use of data mining to enhance education outcomes, giving rise to the interdisciplinary subject of EDM. EDM involves data extraction, screening, conversion, mining, and model analysis using mathematical techniques and mining algorithms to derive actionable insights for teaching and management, thereby adding meaningful value [29-31].

EDM technology uses theories and technologies from social psychology, computer science, pedagogy, and statistics to address various educational issues. It aids educational decision-making, fostering student initiative, and refining teaching methods. EDM shares common processes with most data mining endeavors, such as data preparation, screening, pre-processing, and pattern evaluation [32]. However, its distinctiveness lies in utilizing educational data to discover knowledge and iteratively optimize the educational environment, constituting a circular process of generation, testing, and continuous improvement.

EDM extracts valuable knowledge from the education system, benefiting educators, learners, administrators, and system developers. It serves diverse roles for different stakeholders. For learners, personalized course recommendations enhance learning enthusiasm and outcomes, fostering effective teaching activities. Educators benefit from objective teaching feedback, course evaluation, and early identification of students needing support [33]. For administrators, combining EDM into educational administration facilitates comprehensive analysis of curriculum, teaching processes, and teaching resources, providing managers with objective decision-making insights to improve teaching management.

1.1.4. Research on U-Net

U-Net, a classic CNN introduced in 2015, has spurred extensive research into its network structure, leading to numerous publications focused its structure improvement. Tailored for image data, U-Net's unique symmetric structure excels in feature extraction and utilization [34-36]. This capability makes U-Net well-suited for processing complex patterns and structures in data like athlete movement data or videos. By efficiently extracting key features from motion data, U-Net can accurately predict athletic performance. Furthermore, its encoder-decoder architecture efficiently integrates context information across different scales, enabling it can to identify both local features and their broader contextual significance. This contextual understanding is crucial for predicting athletic performance, encompassing an athlete's performance at specific moments and their overall performance throughout the entire motion process.

The U-Net architecture comprises symmetrical U-shaped structure, composed of a contracting path and an expanding path. The contracting path is a classic CNN architecture, featuring repeated structures with two convolutional operations and a pool each time. Convolutional layers employ 3×3 convolution kernels and Rectified Linear Unit (ReLU) activation function, alongside maximum pooling. The number of feature maps doubles with each down-sampling step. Each step in the expanding path begins with up-sampling, halving the number of feature maps and doubling the map size. The upsampled results are then concatenated with corresponding feature maps from the contracting path, cropped to match size, and subjected to two 3×3 and one 1x1 convolution operations. However, the network faces two significant challenges. Firstly, it is computationally inefficient, running separately for each neighborhood and repeating for overlapping neighborhoods. Secondly, there is a trade-off between accurate positioning and obtaining context information [37]. Larger patches require more maximum pooling layers, reducing positioning accuracy, while smaller neighborhoods yield less contextual information [38]. Balancing these factors is crucial. The structure of the U-Net is illustrated in Fig. 4.

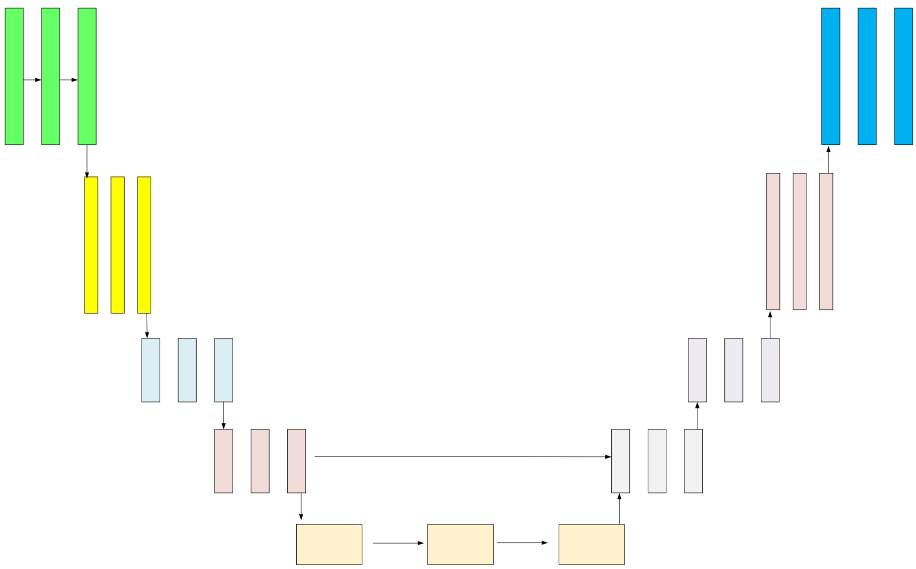

Fig. 4. U-Net structure.

1.1.5. Improved U-Net

To sort out the super-resolution problem, researchers proposed the Residual Dense Network (RDN), comprising three components. Part 1 is a shallow feature extraction module (SFEM), and part 2 comprises a residual dense block (RDB) integrating residual and dense blocks to extract intensive features from SFEM outputs. Part 3 involves long-link global residual learning, facilitating smoother network training and deeper feature extraction. Fig. 5 displays the structure of RDN.

In U-Net, features extracted from each layer are learned once, and there is no inter-layer connection, leading to low feature utilization and segmentation accuracy. In contrast, DenseNet employs dense connections to combine current layer features with all previous layers and transfer the generated features to subsequent layers. This cascading approach enables each layer to learn from previous layers, enhancing feature transmission and addressing gradient vanishing. Let x_l denote the output of the l-th layer in the DenseNet, as defined by Eq. (7).

$$x_l = H_l([x_{l-1}, x_{l-2}, \cdots, x_0]) \tag{7}$$

In Eq. (7), function H_l represents the nonlinear transformation of the l-th layer, and [...] means the fusion operation of all layers' characteristics.

Therefore, this study introduces a dense connection module based on DenseNet principles to enhance model performance by refining feature utilization and transmission. Specifically, each dense connection module consists of two 3 × 3 convolutional layers followed by two feature fusion operations, to augment feature extraction and integration. This design ensures that the feature maps generated after each convolution operation are fused with the original input's feature map, enriching it. Such fusion enhances feature complexity and depth and promotes the model's comprehensive understanding of information. Additionally, by transmitting the fused feature maps between modules, this optimization strategy facilitates information flow in the network, addressing potential issues of information loss and gradient vanishing. This dense connection mechanism means allows each layer direct access to features from all preceding layers, significantly enhancing feature extraction and learning. Therefore, this structure not only deepens the network's hierarchy but also improves its accuracy and efficiency in handling complex data. The improved network model's structure is implied in Fig. 6.

Moreover, a batch normalization (BN) layer and a modified Linear Unit activation layer are added after each convolutional layer to enhance network performance. The BN layer solves the problems related to the network's sensitivity to initial data distribution and poor generalization ability. It first normalizes the input data x_i, represented mathematically as follows:

$$\mu_B = \frac{1}{m}\sum_{i=1}^{m} x_i \tag{8}$$

$$\sigma_B^2 = \frac{1}{m}\sum_{i=1}^{m} (x_i - \mu_B)^2 \tag{9}$$

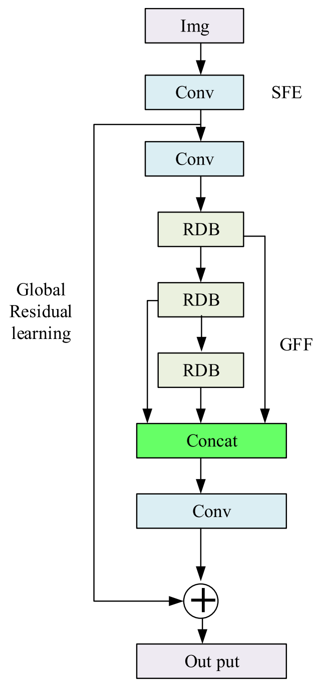

Fig. 5. The structure of RDN.

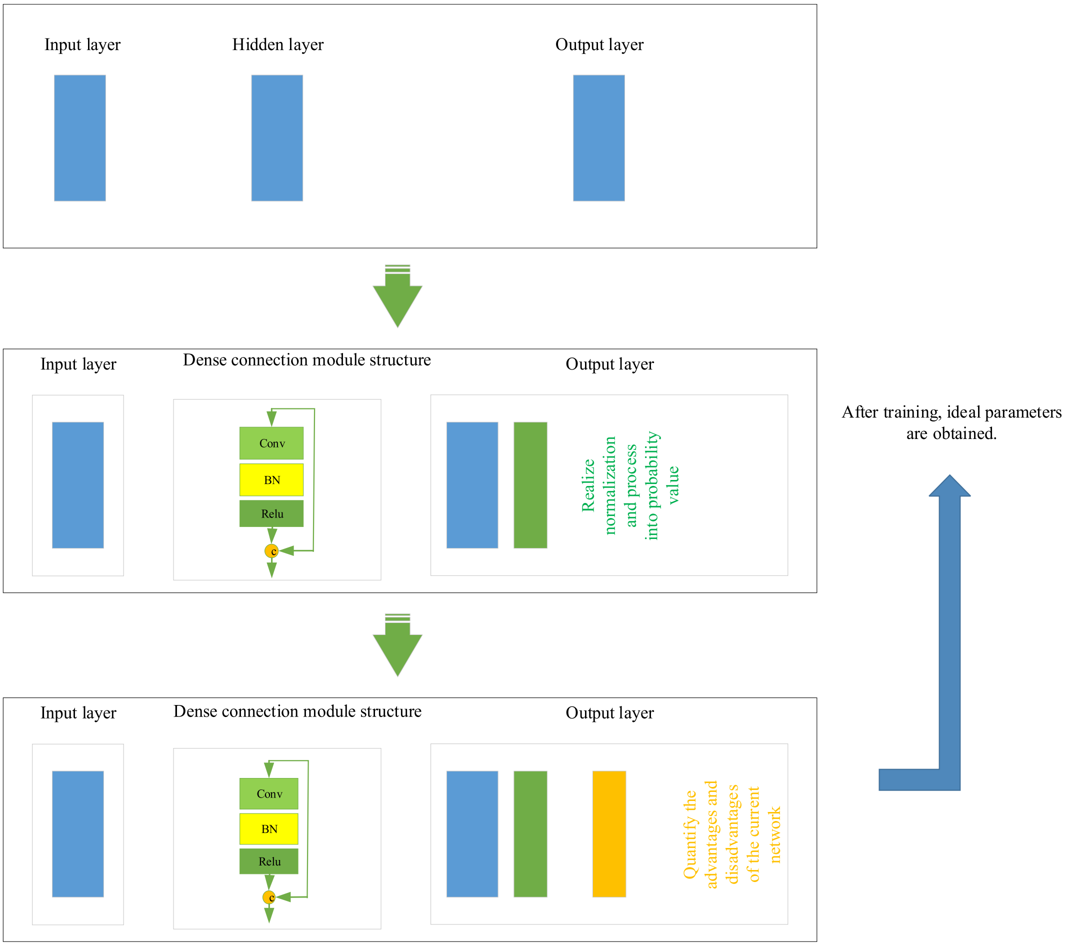

Fig. 6. Structure of the dense connection module.

$$\widehat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}} \tag{10}$$

In Eqs. (8)–(10), m refers to the batch size of the input data; μ_B means the mean value; σ_B^2 signifies the variance; ε stands for the smoothing factor, preventing division by zero. This alteration reshapes the distribution of data features learned from the network. To mitigate the impact of changes in the effect of online learning, the normalized data $\widehat{x}_i$ needs to be transformed and reconstructed, as implied in Eq. (11).

$$y_i = \gamma \widehat{x}_i + \beta \tag{11}$$

In Eq. (11), γ and β represent the learnable reconstruction parameter, and y_i denotes the output value of the network after BN of the input data x_i. In the improved U-Net, each layer is directly connected to all preceding layers, enabling the output of each layer to serve as input for the subsequent layer and to be directly transmitted to all subsequent layers. This dense connection design fosters seamless information flow within the network, effectively addressing issues of gradient vanishing in deep networks. Because gradients can propagate more effectively through these direct connections, the learning speed and quality of lower-level features are enhanced. Additionally, dense connections facilitate feature reuse, allowing the network to directly utilize features from preceding layers. This not only amplifies the network's feature learning capabilities but also mitigates the issue of gradient vanishing during training, as gradients can propagate directly from the output layer to nearly all input layers.

The improved Linear Unit activation is responsible for mapping the input of neurons in the neural network to the output, thereby introducing nonlinear factors into the network and improving the nonlinear expression ability of the network. Its definition can be written as Eq. (12):

$$f(x) = max(0, x) \tag{12}$$

The attention mechanism facilitates the network in selecting important features in the data while suppressing irrelevant ones. The depth feature selection method of the Convolutional Block Attention Module (CBAM) is embedded in the feature extraction stage of the improved U-Net model, which prioritizes regions with the most category differentiation. CBAM, a lightweight universal attention module, enhances the network's feature extraction ability without significantly increasing parameters or computational complexity. It is composed of two sub-modules: the spatial attention module (SAM) and the channel attention module (CAM). SAM focuses on spatial relationships, while CAM explores inter-channel dependencies, treating feature map on each channel as a feature detector. To integrate the attention mechanism module into U-Net, the residual structure and 3 × 3 convolution kernel size are combined to facilitate convolutional separation and improve network efficiency. Subsequently, transpose convolution is added to capture local features to gather global information, while BN processing is performed to enhance overall network generalization. The encoding of the attention mechanism is portrayed in Eq. (13).

$$C_m = \sum_{m=1}^{n'} a_{nm} h_m \tag{13}$$

In Eq. (13), h_m represents the conversion function of the coding stage, a_{nm} refers to the weight, n' and indicates the length of feature information.

Down-sampling reduces the probability of accurately identifying pixels, while up-sampling in U-Net has a limited capacity to recover feature information. Therefore, this study extends the U-Net framework by incorporating a fusion layer of pooling and convolution layers at the bottom of the U-shape architecture. This extension aims to improve network speed, reduce computational complexity, and expand the feature map to capture local and detail information. The fusion module comprises a convolution-pooling layer consisting of a maximum pooling layer with a stride of 2, a depth-wise separable convolution with a kernel size of 1 × 1, and a 2 × 2 up-sampling layer. The maximum pooling layer reduces parameters in subsequent convolutional layers and prevents overfitting while decreasing image resolution. After pooling, the up-sampling operation restores lost feature information, independently inputting the output from the previous layer into each convolution-pooling layer. Through layer fusion, channel dimensions are spliced to expedite the training process of the network model.

1.2. Experimental design

1.2.1. Dataset collection

The experiment utilizes the Microsoft Common Objects in Context (COCO) dataset, an open dataset widely used for object recognition, detection, and segmentation tasks. Released by Microsoft in 2014, the COCO dataset comprises over 330,000 images covering 80 common object categories, including people, animals, vehicles, and furniture. Each image includes multiple annotations, including object bounding boxes, instance segmentation masks, and image descriptions. Additionally, the dataset presents various challenging scenarios such as occlusion, small objects, and irregular shapes. In addition to image-level annotations, the COCO dataset provides detailed annotations for objects, specifying the category, bounding box shape, and instance segmentation mask for each object instance. This allows researchers to perform more fine-grained object detection and segmentation tasks. The COCO dataset is widely employed for training and evaluating the performance of computer vision models, and it serves as a benchmark for state-of-the-art algorithms for object detection, instance segmentation, and image captioning. It is available for download from the official website (http://cocodataset.org/#download). While primarily applied in image recognition and object detection, techniques developed using the COCO dataset can be transferred to sports performance prediction, particularly in the analysis and recognition of athletes' movements and performances. Processing methods applied to the COCO dataset, such as data cleaning, annotation, and augmentation, are highly valuable for handling sports performance datasets. Models demonstrating strong performance on the COCO dataset indicate their generalizability and robustness, instilling confidence in their application for sports performance prediction. In summary, although primarily used in image processing, the experience and techniques gained in model development and validation on the COCO dataset are entirely applicable to the domain of sports performance prediction.

1.2.2. Grouping normalization

To accelerate model convergence and alleviate overfitting, normalization is added to the model. BN operates on batches of data, with the batch size N impacting its performance. Given the limited experimental data, setting a small batch size can lead to inaccurate variance calculation. To address this, this model adopts grouping normalization to calculate the mean and variance for each group within channel C. This ensures accurate normalization even with small batch sizes. The equation is as follows:

$$\widehat{x}_a = \frac{1}{\sigma_a}(x_a - \mu_a) \tag{14}$$

$$\sigma_a = \sqrt{\frac{1}{n}\sum_{k\in S_a}(x_k - \mu_a)^2 + \varepsilon} \tag{15}$$

$$\mu_a = \frac{1}{n}\sum_{k\in S_a} x_k \tag{16}$$

$$y_a = \gamma \widehat{x}_a + \beta \tag{17}$$

$$S_a = \left\{ k \middle| k_M = a_M, \left\lfloor \frac{k_C}{C/G} \right\rfloor = \left\lfloor \frac{a_C}{C/G} \right\rfloor \right\} \tag{18}$$

Here, $a = (a_N, a_C, a_H, a_W)$, ε refers to a constant; γ and β represent learnable variables; C and G stand for the number of channels and groups, respectively.

1.2.3. Data cleaning and pre-processing

The quality of the sample library significantly impacts the performance of deep learning models. Higher-quality samples facilitate faster and more precise training of neural network models. In this study, data cleaning focuses mainly on the sample base. First, an initial model is trained using the uncleaned samples. Then, the initial model is employed to evaluate all similar samples. Samples with low scores undergo direct elimination or correction of mislabeling. This process iterates through reprocessing steps until the desired result is achieved, typically involving two or more cycles.

1.2.4. Experimental training process and environment

The improved network architecture is built in Torch, and the hardware configuration for network training is detailed in Table 1. The training process unfolds as follows.

1) Input the training sample into U-Net simultaneously for processing. The network conducts forward propagation to predict pixel categories and calculates the loss of the cost function by comparing predictions with marked images.
2) Utilize stochastic gradient descent for backpropagation learning to optimize network parameters.
3) Cease training when the cost function converges and stabilizes, iterating the entire training process 50 times.
4) Conclude the training.

2. Analysis and discussion of experimental results

2.1. Experimental performance comparison of the algorithm on microsoft COCO dataset

To validate the model's efficacy, multiple tests are conducted on the prediction capabilities of the original U-Net network, improved U-Net, Resnet U-Net, Attention ResUNet, and DUNet using the COCO dataset. Normalize the data from the COCO dataset and conduct experiments afterwards. The original U-Net network serves as the baseline model, providing a crucial comparative foundation. Given U-Net's renowned success in medical image segmentation and its versatility across various image processing tasks, its inclusion as a reference model facilitates clear visualization of performance gains of the enhanced model. Additionally, evaluating the original U-Net's performance helps ascertain the genuine effectiveness of any architectural modifications or enhancements. Choosing Resnet U-Net as a comparative model aims to assess the impact of residual learning on enhancing network performance, especially concerning complex sports performance prediction tasks. Introducing an attention mechanism into Resnet U-Net, represented by Attention ResUNet, further enhances the model's ability to focus on pivotal features. This attention mechanism enables the model to prioritize information relevant to the prediction task, thereby improving accuracy and efficiency. DUNet, a variant U-Net incorporating dense connections inspired by DenseNet's dense connection strategy, aims to enhance feature propagation and reuse. Selecting DUNet as a comparative model reflects an endeavor to address deep learning challenges by optimizing feature flow.

The prediction performance of each model is analyzed and compared across key evaluation indexes such as accuracy, precision, recall, and F1-score. Table 2 presents a comprehensive comparison of five algorithms' prediction performance on the COCO dataset.

In Table 2, the improved U-Net demonstrates superior performance across all key performance metrics, achieving 93 % accuracy, 94 % precision, 93 % recall, and an F1 score of 93 %. Conversely, the original U-Net and DUNet show slightly inferior performance, with accuracies of 88 % and 87 %, respectively. While Attention Resnet U-Net matches with the improved U-Net in accuracy, it exhibits slightly lower values in other metrics.

Table 1
List of experimental parameters.

Items	Parameters
Number of CPUs in servers	3
CPU	40 address sizes: 46 bits physical, 48 bits virtual 20 physical id: 0 20 physical id: 1
CPU model	Intel(R) Xeon(R) Silver 4210 CPU @ 2.20 GHz
Number of running bits of CPU	64-bit operating system
Kernel information of the operating system	Linux ubuntu 5.4.0–94-generic #106~18.04.1-Ubuntu SMP Fri Jan 7 07:23:53 UTC 2022 x86_64 x86_64 x86_64 GNU/Linux
Python	3.8

Table 2
Comparison of prediction performance of five algorithms on the COCO dataset.

Model	Accuracy	Precision	Recall	F1-score
Original U-Net	0.88	0.87	0.87	0.88
Improved U-Net	0.93	0.94	0.93	0.93
Resnet U-Net	0.88	0.90	0.88	0.89
Attention ResUNet	0.93	0.91	0.90	0.90
DUNet	0.87	0.88	0.87	0.87

2.2. Analysis of experimental results of network performance

The experiment's results are analyzed based on six repeated tests, enhancing the reliability and stability of the findings. Single tests may be influenced by random factors, leading to non-representative results. Multiple tests enable the aggregation of results, facilitating in-depth analysis through statistical indicators such as mean, median, and standard deviation. This approach ensures a comprehensive and accurate evaluation of the model's performance. Additionally, multiple tests aid in assessing the model's generalization ability across different datasets or conditions, offering insights into its stability and generalization capability. Fig. 7 outlines the change curve of network error during training, comparing the performance between the original and the improved U-Net models.

Fig. 7a and b shows the training performance of the improved U-Net model at different learning rates. Upon reaching 3500 iterations, the loss curve exhibits minor peaks at 4550 and 6800 iterations. However, the overall trend indicates a decrease in network loss, approaching stability. By 8000 iterations, the error remains consistently below 0.2, indicating stabilization. Conversely, the original U-Net stabilizes at below 0.25 loss by 7800 iterations. Notably, the training error of the improved network consistently outperforms that of the original network, highlighting the latter's superior training effect.

The convergence speed scores of the model under varying learning rates is indicated in Fig. 8.

The results of Fig. 8a and b shows that at lower learning rates, the model requires more iterations to achieve convergence. This phenomenon arises because excessively small lr values compromise the regular term's constraint ability. For lr values of 0.0001 and 0.3, it takes 60 and 25 iterations, respectively, for the diffusion process to stabilize.

In the sports achievement prediction problem, feature extraction is carried out for various types of data features. The accuracy of model feature extraction is revealed in Fig. 9.

Fig. 9a and b suggest that the optimized U-Net demonstrates enhanced feature extraction accuracy compared with the original U-Net. However, this improvement comes at the cost of extended extraction time. Particularly for numerous extraction tasks, employing the optimized U-Net entails a relatively higher feature extraction time cost.

2.3. Analysis of experimental results of sports achievement prediction

Fig. 10 depicts the actual results of PE subjects alongside the predicted scores by the original U-Net, the improved U-Net, Resnet U-Net, Attention ResUNet, and DUNet.

Fig. 10a and b shows that the actual result of PE subjects in the first test is set as 74 points. In Fig. 10, the original U-Net, Resnet U-Net, Attention ResUNet, DUNet, and the improved U-Net predicted the PE subject scores of 67, 61, 67, 66, and 71, respectively. Comparative analysis of the predicted results against the actual outcomes reveals that the improved network achieves a 5.22 % higher accuracy in sports performance prediction than DUNet and a 4.22 % higher accuracy than the original U-Net. These findings indicate that the predictive performance of the improved U-Net is closer to the target results, with smaller prediction errors.

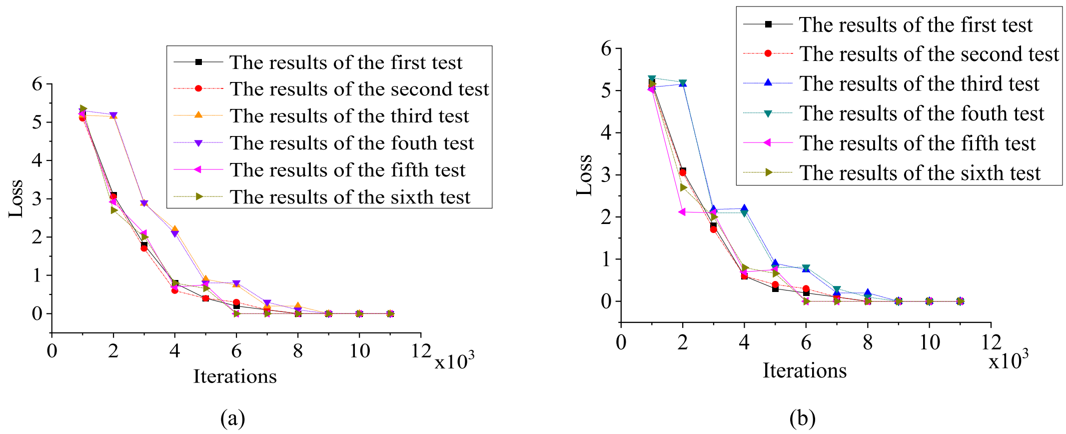

Fig. 7. Change curve of network error
(a. Learning rate: 0.1; b. Learning rate: 0.2).

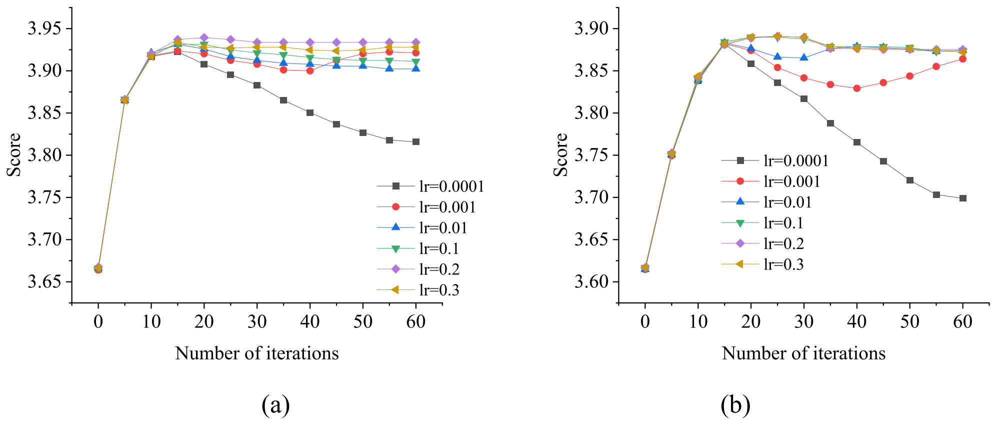

Fig. 8. Convergence speed scores of the system under diverse learning rates
(a. maximum pooling operation score; b. sum pooling operation score).

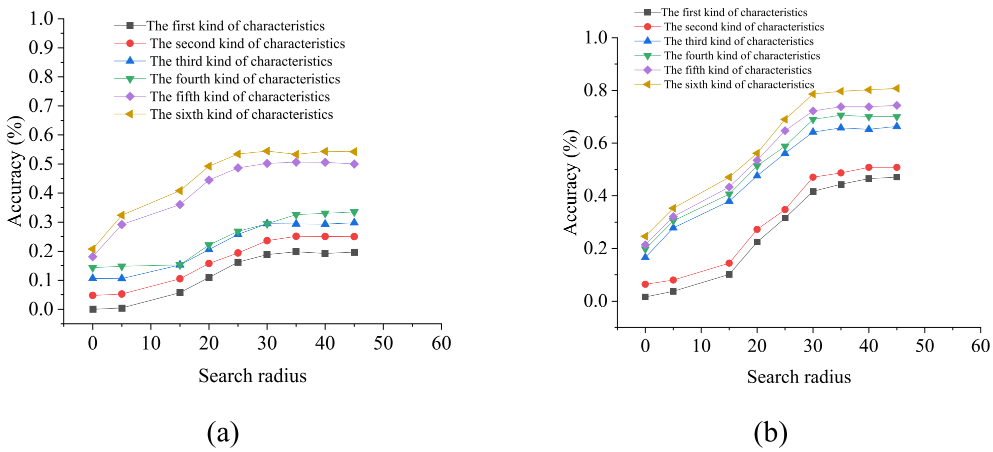

Fig. 9. Accuracy of model feature extraction
(a: Original U-Net; b: Improved U-Net).

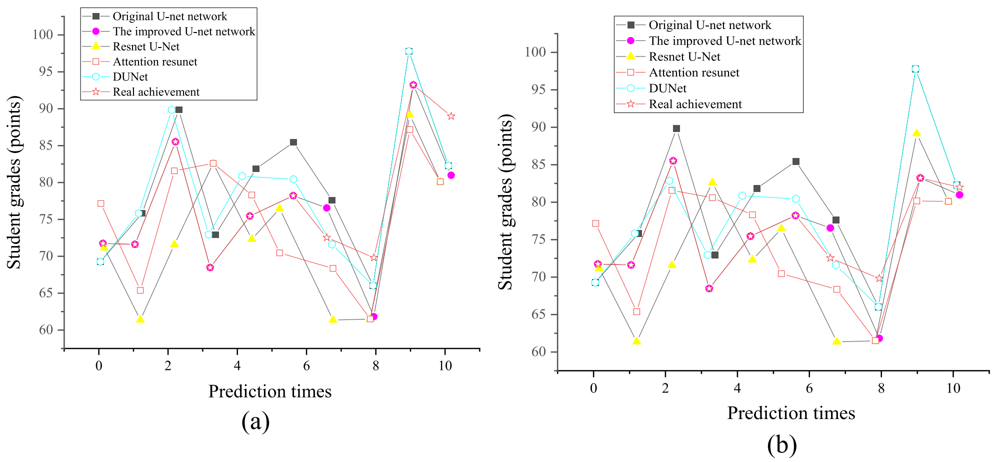

Fig. 10. Experimental results of model prediction
(a. Class 1; b. Class 2).

3. Discussion

In the performance comparison experiment, the improved U-Net integrates the dense connection mechanism of DenseNet, enhancing feature propagation and reducing information loss during transmission. This enhancement significantly improves the model's accuracy and recall. Additionally, employing a mixed loss function effectively addresses class imbalance issues, enhancing recognition rates for minority classes and improving precision. Through these strategies, the improved U-Net not only enhances performance but also strengthens adaptability to different sports and diverse data distributions, exhibiting higher generalization ability. The study aims to develop a sports performance prediction system based on the improved U-Net model, with the goal of optimizing the quality and efficiency of sports teaching through machine learning techniques. Throughout this process, various challenges encountered include data accessibility and diversity, model generalization ability, and high computational resource requirements. Nevertheless, the improved U-Net model, with its outstanding feature extraction capabilities and high adaptability to different data types, demonstrates good generalizability across various sports and performance scenarios. In real-world educational environments, especially in smart classrooms, the predicted system is expected to significantly enhance teaching interactivity and personalized learning experiences. The system can monitor and analyze students' sports performance in real-time, enabling teachers to provide more precise guidance based on each student's specific needs. However, when developing and applying such systems, careful consideration must be given to ethical impacts and potential biases, especially ensuring data representativeness and fairness, as well as safeguarding student privacy and data security. This study has profound implications for the field of PE, advancing education modernization and showcasing the immense potential of artificial intelligence (AI) technology in promoting teaching efficiency and personalized learning. Furthermore, it contributes to the ongoing dialogue on the adoption of AI technology in education, offering new perspectives and practical cases for AI applications in PE. For researchers or educators intending to implement similar AI systems into teaching practices, it is recommended to prioritize data quality and diversity, maintain model transparency and interpretability, and rigorously consider ethical and privacy protection issues. Comprehensive testing and evaluation should be conducted before implementation to ensure the effectiveness and applicability of the system. Through these efforts, this study not only advances research in the technical aspects of sports performance prediction but also provides valuable experience and insights for the broader application of AI in education. In comparison to the research by Sun et al. (2023) [39], this study examines the theoretical performance of the model and explores its practical deployment in real-world educational environments, especially within intelligent classrooms. By continuously monitoring students' athletic performances, this system can offer teachers real-time feedback and personalized guidance suggestions, an application potential often overlooked in previous studies. Furthermore, compared to the standard CNN used by Migliaccio et al. (2023) [40], the improved U-Net model in this study demonstrates superior adaptability to different sports and performance scenarios, showcasing enhanced generalization capabilities. This is attributed to its dense connection mechanism and feature fusion strategy, enabling effective capture and utilization of deep features within complex data, thereby achieving efficient performance predictions across various sports disciplines.

4. Conclusion

Sports achievement prediction plays a pivotal role in aiding educational administration departments to devise tailored sports training plans and effectively improve the focalization of sports training. Consequently, high-performance methods for sports achievement prediction have become a hot issue. This study delves into the sports achievement prediction problem and proposes a prediction model based on the U-Net CNN. Firstly, in the feature extraction stage of the improved U-Net model, the depth feature selection method of CBAM is embedded, enabling the model to focus on regions with the most category differentiation. Secondly, the U-Net framework is extended by adding a fusion layer of convolution and pooling layers at the bottom of the U-shape. This extension thus reduces computation, improves network speed, and expands the feature map to capture local and detail information. The actual test proves the proposed model's high predictive performance, demonstrating its efficacy in effectively predicting sports achievements with accuracy meeting actual requirements. Nonetheless, this study has certain limitations. Although the improved U-Net model performs exceptionally well in the current task, there may be insufficient assessment of its generalization capabilities to other related tasks. Future research could explore the model's applicability and effectiveness in different types of sports performance prediction tasks. Additionally, while the introduction of dense connections and mixed loss functions has enhanced the model's performance, it may have increased the complexity and computational costs of the model. Subsequent studies need to explore more efficient model optimization techniques to strike a balance between performance improvement and computational resource consumption.

Data availability statement

All of the data generated or analyzed during this study are included in this published article [and its supplementary information files].

CRediT authorship contribution statement

Guoliang Wang: Writing – review & editing, Writing – original draft, Supervision, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Tianping Ren:** Writing – original draft, Resources, Project administration, Methodology, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Project source: General Humanities and Social Sciences Research Project of the Ministry of Education of China; Project name: Research on the long-term governance mechanism of school sports informatization; Project No.: 21YJA890030.

Project source: Science and Technology Research Program of Henan Province of China; Project name: Sports fitness training system based on 3D somatosensory technology; Project No.: 222102320179.

Project source: Higher Education Teaching Reform Research and Practice Project of Henan Province of China; Project name: Research and Practice on the Implementation Path of Physical Education Informatization in Colleges and Universities; Project No.: 2019SJGLX243.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e30055.

References

[1] F.J. Agbo, I.T. Sanusi, S.S. Oyelere, Application of virtual reality in computer science education: a systemic review based on bibliometric and content analysis methods, INT J EDUC SCI 11 (3) (2021) 142–144.
[2] N. Pellas, A. Dengel, A. Christopoulos, A scoping review of immersive virtual reality in STEM education, IEEE T LEARN TECHNOL 13 (4) (2020) 748–761.
[3] M. Sattar, S. Palaniappan, A. Lokman, et al., Motivating medical students using virtual reality based education, iJET 15 (2) (2020) 160–174.
[4] E. McGovern, G. Moreira, C. Luna-Nevarez, An application of virtual reality in education: can this technology enhance the quality of students' learning experience, J. Educ. Bus. 95 (7) (2020) 490–496.
[5] O. Halabi, Immersive virtual reality to enforce teaching in engineering education, MULTIMED TOOLS APPL 79 (3) (2020) 2987–3004.
[6] K.C. Jaiswal, P.D. Pathak, V.J. Pawar, Prediction of degree student achievement analysis, IOP Conf. Ser. Mater. Sci. Eng. 1070 (1) (2021) 012057.
[7] A. Joshi, P. Saggar, R. Jain, CatBoost - an ensemble machine learning model for prediction and classification of student academic performance, ADV DATA SCI ADAPT 1 (3/4) (2021) 13.
[8] C.F. Rodríguez-Hernández, M. Musso, E. Kyndt, Artificial neural networks in academic performance prediction: systematic implementation and predictor evaluation, MULTIMED TOOLS APPL 9 (3) (2021) 29.
[9] D.K. Singh, A study of spiritual intelligence and anxiety as a predictors of academic achievement among undergraduate students, Computer-Aided Design and Applications 1 (2021) 12–20.
[10] D. Monteverde-Suárez, P. González-Flores, R. Santos-Solórzano, Predicting medicine students' achievement and analyzing related attributes with ANN and Nave Bayes, PUBLIC POLICY ADMIN 5 (3) (2021) 27.
[11] S. Grek, J. Ozga, Re-inventing public education: the new role of knowledge in education policy making, PUBLIC POLICY ADMIN 25 (3) (2019) 271–288.
[12] W. Vincent, S.K.S. Shanmugam, The role of classical test theory to determine the quality of classroom teaching test items, PEDAGOGIA Jurnal Pendidikan 9 (1) (2020) 5–34.
[13] S. Li, T. Shi, J. Kuang, Exploration and practice of informatization means in the quality supervision of college classroom teaching, J. Phys. Conf. 1345 (4) (2019) 042032.
[14] H. Li, H. Zhao, Improvement of intelligent computer aided Chinese teaching system, Computer-Aided Design and Applications 18 (2020) 12–24.
[15] L. Wu, Student model construction of intelligent teaching system based on Bayesian network, PERS UBIQUIT COMPUT 24 (3) (2020) 419–428.
[16] M. Li, L. Xie, A. Zhang, Reinforcement emotion-cognition system: a teaching words task, COMPUT INTEL NEUROSC 2019 (4) (2019) 1–13.
[17] S. Dong, Intelligent English teaching prediction system based on SVM and heterogeneous multimodal target recognition, INT J FUZZY LOG INTE 38 (153) (2020) 1–10.
[18] H. Lin, S. Xie, Z. Xiao, Adaptive recommender system for an intelligent classroom teaching model, iJET 14 (5) (2019) 51.
[19] X. Gao, H.Z. Xiong, F. Xu, Design and implementation of Wisdom classroom management system based on voice control, J. Phys. Conf. 1651 (1) (2020) 012038.
[20] Q. Huang, Research on the hybrid teaching Reform of online open course three-dimensional classroom, Learn. Teach. 2 (9) (2019) 55–61.
[21] J. Protopopova, S. Kulik, Educational intelligent system using genetic algorithm, Procedia Comput. Sci. 169 (2020) 168–172.
[22] P. Sokkhey, T. Okazaki, Development and optimization of deep belief networks applied for academic performance prediction with larger datasets, IEIE Transactions on Smart Processing and Computing 9 (4) (2020) 298–311.
[23] J. Mirarabrazi, I.H. Navrodi, I. Ghajar, Identifying optimal location of ecotourism sites by analytic network process and genetic algorithm (GA): (Kheyroud Forest), INT J ENVIRON SCI TE 17 (5) (2020) 2583–2592.
[24] S. Lai, B. Sun, F. Wu, Automatic personality identification using students' online learning behavior, IEEE T LEARN TECHNOL 13 (1) (2020) 26–37.
[25] S. Hachmoud, A. Hachmoud, Analysis of students online learning behavior in a pedagogical model combining blended learning and competency based approach, Int. J. Adv. Trends Comput. Sci. Eng. 8 (6) (2019) 3389–3395.
[26] N. Wongtongkam, Influence of coping, self-esteem and social support on undergraduate students' emotional distress, Health Educ. 119 (3) (2019) 187–201.
[27] M.P. Gronberg, S.S. Gay, T.J. Netherton, D.J. Rhee, L.E. Court, C.E. Cardenas, Dose prediction for head and neck radiotherapy using a three-dimensional dense dilated U-net architecture, Med. Phys. 48 (9) (2021) 5567–5573.
[28] D. Li, X. Chu, Y. Cui, J. Zhao, K. Zhang, X. Yang, Improved U-Net based on contour prediction for efficient segmentation of rectal cancer, Comput. Methods Progr. Biomed. 213 (7) (2022) 106493.
[29] T. Gangopadhyay, S. Halder, P. Dasgupta, U. Mtse, Net: an architecture for segmentation, and prediction of fetal brain and gestational age from MRI of brain, Network Modeling Analysis in Health Informatics and Bioinformatics 11 (1) (2022) 50.
[30] Y. Ni, Z. Xie, D. Zheng, Y. Yang, W. Wang, Two-stage multitask U-Net construction for pulmonary nodule segmentation and malignancy risk prediction, Quant. Imag. Med. Surg. 12 (1) (2022) 292.
[31] N. Sun, Z. Zhou, Q. Li, X. Zhou, Spatiotemporal prediction of monthly sea subsurface temperature fields using a 3D U-Net-Based model, Rem. Sens. 14 (19) (2022) 4890.

[32] J.S. Hong, Y.H. Tzeng, W.H. Yin, K.T. Wu, H.Y. Hsu, Automated coronary artery calcium scoring using nested U-Net and focal loss, Comput. Struct. Biotechnol. J. 20 (11) (2022) 1681–1690.

[33] M. Dabass, J. Dabass, S. Vashisth, R. Vig, A hybrid U-Net model with attention and advanced convolutional learning modules for simultaneous gland segmentation and cancer grade prediction in colorectal histopathological images, Intelligence-Based Medicine 7 (1) (2023) 100094.

[34] A. Halder, D. Dey, Atrous convolution aided integrated framework for lung nodule segmentation and classificat, Biomed. Signal Process Control 82 (2023) 104527.

[35] P. Asthana, M. Hanmandlu, S. Vashisth, Brain tumor detection and patient survival prediction using U-Net and regression model, Int. J. Imag. Syst. Technol. 32 (5) (2022) 1801–1814.

[36] Y.P. Wang, Y.C. Jheng, K.Y. Sung, H.E. Lin, I.F. Hsin, Use of U-Net convolutional neural networks for automated segmentation of fecal material for objective evaluation of bowel preparation quality in colonoscopy, Diagnostics 12 (3) (2022) 613.

[37] L. Nava, K. Bhuyan, S.R. Meena, O. Monserrat, F. Catani, Rapid mapping of landslides on SAR data by attention U-Net, Rem. Sens. 14 (6) (2022) 1449.

[38] M.L. Taccari, J. Nuttall, X. Chen, H. Wang, B. Minnema, P.K. Jimack, Attention U-Net as a surrogate model for groundwater prediction, Adv. Water Resour. 16 (3) (2022) 104169.

[39] H.C. Sun, T.Y. Lin, Y.L. Tsai, Performance prediction in major league baseball by long short-term memory networks, International Journal of Data Science and Analytics 15 (1) (2023) 93–104.

[40] G.M. Migliaccio, L. Russo, M. Maric, J. Padulo, Sports performance and breathing rate: what is the connection? A narrative review on breathing strategies, Sports 11 (5) (2023) 103.