Feng Liu

Design of intelligent assistive system for physical education: based on personalized training plan

Abstract: Instead of viewing different components in solitude, teachers' opinion with respect and pedagogical knowledge demands a comprehension of the relationships connecting them. The instructor must be knowledgeable about the subject matter, comprehend the most effective delivery methods for the information, be acquainted with the traits of the children, and be aware of the academic setting to educate in that field effectively. Sport is still undergoing economic, cultural, and ethical change. On the other side, throughout the past century, science has been the sport's most pervasive transformation. Players can now jump higher, run faster, and, importantly—retain their health due to scientific understanding. Even though scholars, organizations, and governments have indeed encouraged physical education instructors to incorporate software in their classes, technology is often employed for routine duties, such as tracking enrollment and assessing, recording, and reporting children’s progress. In order to estimate posture in physical training, this research proposes a continuous filter convolution neural network. The approach also assesses the learners' understanding, memory, and accomplishments and offers suggestions for enhancements and remedial actions. The framework and conventional teaching-learning methodologies are then compared characteristics per aspect for output criteria. Finally, the classification algorithm is contrasted with other deep learning algorithms, and it is found that the proposed ConFil_CNN achieves 98.5% accuracy, 96.7% of precision, 93.5% of recall, 95.2% of F1-score and 12.4ms of response time.

Keywords: Physical education; pedagogical concept; posture estimation; neural network; frame selection.

I. INTRODUCTION

Around 165 million children are enrolled in higher education programs globally, and it is expected that number will rise by 98 million within the next ten years. Organizations are constantly implementing strategies for successfully teaching to specific massive enrollments to educate the best, such as the growing university education market [1]. Children’s who are referred to as "Millenials," "NextGen'ers," "Digital natives," or "Gen Y" of the next century and were born after 2002 are enrolling in higher education courses. Because these children were born once the computer system was first launched, a large portion of modern culture in the early class possesses computers, and exposure to electronic content has become a regular part of their surroundings. These pupils have inclinations to use technologies [2]. Institutions are boosting the number of courses they offer electronically to meet such burgeoning enrollments and to engage those who are "digital natives actively." While enrolling in online education courses, most learners are either unable to attend on-campus classes or seek "convenience" and "flexibility" in their education. If not appropriately managed, children’s learning online may have fewer possibilities for social, intellectual, and developmental work than those participating in conventional face-to-face classes [3]. Instructors were looking for novel methods to integrate educational materials and other kinds of student resources into the courses as courses gained in popularity. Despite the numerous chances for student and instructor contact in distance learning, it is more difficult to replicate the same level of conversations that occur in face-to-face settings, particularly in "action-packed" practical classes like physical education (PE) [4]. Virtual pedagogy is one well approach that can

1College of Art and Physical Education, Fuyang Preschool Teachers College, Fuyang, Anhui, 236015 China

*Email: Anhui1ufeng2025@163.com

Copyright © JES 2024 on-line: journal.esrgroups.org
mimic the intense exchanges that occur in real-world PE sessions. Modeling is a tool for teaching and learning that replicates the features and look of a more real-world experience [5]. As an alternate, models are defined as a setting created to improve a person's involvement with real situations, which can help a person build skills and learning goals. Institutions, secondary and elementary campuses, and other training centers were forced to solely offer educational software when the new monarch pandemic broke out in 2020 and in response to the cry for "rescinding courses and not learning"[6]. Online learning has so far progressed and evolved into a brand-new educational paradigm for the Age of the internet. It is also clear how online learning has evolved in higher education. Digital marketing is the ideal fusion of contemporary educational design and communications technologies.

Online desktops, smartphones, and e-readers offer a high level of interactivity, expression, simplicity, and coverage [7]. It does not impose time or location restrictions and can compensate for the drawbacks of conventional educational approaches. University physical education can overcome obstacles and enable children to gain knowledge anytime and anywhere. Big data and machine intelligence capabilities are now significant engines propelling scientific innovation [8]. eLearning is evolving into a new education mode due to the deep convolution network's full integration into artificial learning and education. In order to attain this goal, a system for online media instruction is set up [9]. This platform allows children's and educators to engage while learning across time and distance. The eLearning platform uses the Web to bring technology, innovation, education, and instruction together. It eliminates the restrictions on the teaching environment and the uneven educational resources seen in conventional education. It makes the instructional task of ensuring and distributing a reality.

Additionally, as network resources expand exponentially, so do the resources that users can access. Users benefit from a better learning environment as a result. Consequently, this work's contributions are as described in the following:

- Developing a continuous filter convolution neural network helps estimate posture correctly in the online learning setting for intercollegiate athletics.
- The network testing procedure is conducted to confirm the consistency and effectiveness of the processor speed in this article as well as the degree of satisfaction among teachers and children with the technology.

The structure of the current paper is as follows: A relevant collection of research for video-based pedagogy teaching using neural networks is provided in section 2 of the presentation with a table. In section 3 suggested feature extraction and estimation model is given. In part 4, the performance of the suggested model is shown along with benchmark methods. The fifth section presents the general conclusion for the suggested method.

II. RELATED WORKS

Online education is a flexible, accessible, and connected network learning strategy. Some have the chance to radically alter the nature of learning delivery and competitiveness. With some, it offers the potential for learning for young viewers. With the growth of online learning, academics from around the country have already been concentrating on how satisfied children's, guardians, and teachers are with virtual learning. The forecasting and advanced detection of children's educational performance based on machine learning has been suggested in [10], along with enhanced K-nearest neighbor clustering predicated on frequent patterns. The methodologies of a response variable, stochastic tree, back-propagation Bayesian network, and deep learning models are mainly in comparison. [11] highlights the shortcomings of the evaluation and review of conventional physical education technical skill course work and suggests a procedure that utilizes convolutional neural networks and limited sample-acquiring knowledge. It is analyzed based on the undertone and qualities of machine learning. Multiple RNN-based designs have been proposed in [12] investigation: the first is a recommendation system model for a child's physical education curricula called the LSAPR (long short attention point of interest recommendation) prototype, while the other is a suggestion prototype for a child's PE coursework sequence the LSTM-RNNSR (long- and short-term memory-RNN sequence recommendation).

For determining the quality of primary training on campuses, [13] suggests a Joint Neural Network (JNN) made up of an Improving Support Vector Machine (ISVM) and an Enhanced Back Propagation (BP) network. The settings utilized when employing a typical SVM to categorize extracted features affect the SVM's categorization. In order to undertake in-depth investigations regarding the use of deep neural networks in physical training and its character traits, drawbacks, and developments, [14] embraces the literary criticism method and Joint Neural Network (JNN)
Then it built a Snapchat cellphone based on deep learning in physical training strategic planning overhaul instructional media. IoT-DPARS, an Internet of Things-driven system for recognizing physical activity, has just been suggested in higher education [15]. This module receives pertinent information from the Web of Things and connects with the mobile station through the internet using the information it has access to in tangible.

The goal of [16]'s hybrid physical training teaching approach is to provide pupils with individualized instruction. Initially, the voice control system is built around the three factors of natural language processing, management and key, and automatic speech, and algorithmic improvements are made to increase the recognition rate. Secondly, a brand-new hybrid effective teaching for physical education has been developed. The benefits of conventional physical education instruction are blended with the capabilities of intelligent computer technology to increase the effectiveness of the physical education curriculum in the classroom and the potential for learners to receive customized instruction. An administration paradigm for physiological education in schools is put forth in [17], and a distributed processing mechanism for managing school physiological information is built. Primary school training administration is optimized through large-scale data gathering, and fuzzy C is utilized to make the most of the information or evidence. This research aims to categorize the factors that predict satisfaction and use of massive open online courses (MOOCs) [18]. This paper proposes a behavioral model to illustrate intentions for using via several components, supported by a scholarly literature review process. In order to do this, the researchers conducted a study using an online poll of Spanish-speaking Online consumers.

An internet-supervised training method might be straightforwardly used in various real-world applications, particularly for real-time computer vision from streaming data in which traditional packet-monitored learning approaches could experience severe drawbacks. Supervised Web-based learning is an instinctual extension of traditional batch reinforcement methods.

<table>
<thead>
<tr>
<th>Author/year</th>
<th>Method</th>
<th>Merits</th>
<th>demerits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tao et al. (2022)</td>
<td>K-nearest neighbor clustering</td>
<td>More accuracy</td>
<td>Less complex</td>
</tr>
<tr>
<td>Guo et al. (2022)</td>
<td>Review work</td>
<td>Less cost</td>
<td>More cost</td>
</tr>
<tr>
<td>Hao et al. (2022)</td>
<td>RNN</td>
<td>Computationally active</td>
<td>Less accuracy</td>
</tr>
<tr>
<td>Li et al. (2022)</td>
<td>Joint Neural Network (JNN)</td>
<td>Microscopic samples are needed for re-training.</td>
<td>time-consuming &amp; expensive</td>
</tr>
<tr>
<td>Ba et al. (2021)</td>
<td>IoT-DPARS</td>
<td>The simple and straightforward approach</td>
<td>For dealing with tags, a sizeable amount of labeled information is required.</td>
</tr>
<tr>
<td>Wang et al. (2021)</td>
<td>natural language processing</td>
<td>Less response time</td>
<td>Network training is challenging</td>
</tr>
<tr>
<td>Yang et al. (2020)</td>
<td>distributed processing mechanism</td>
<td>High precision</td>
<td>Different stages of processing are necessary.</td>
</tr>
<tr>
<td>Fang et al. (2019)</td>
<td>fuzzy C</td>
<td>More complex</td>
<td>Cost-effective</td>
</tr>
</tbody>
</table>

### III. SYSTEM MODEL

A structured framework of performance assessment criteria for primary training is required for an objective review. The usual method for assessing teaching quality involves personal education material and intensive educational and academic performances. These two types are combined to generate the instruction and evaluation scores [19-25].
The fundamental characteristics of the categories above of traditional evaluation system indications are practicality, generalizability, and autonomy. When implementing the principle of being less, not being more, selecting observable and measurable that is trustworthy, practical, and genuine is of utmost importance. The four components of teaching content, teaching technique, teaching attitude, and teaching effect are used in this study to develop a multi-index evaluation method for college physical education (PE), as shown in Fig. 1.

Figure 1. Behavioral model of physical education (PE) pedagogical approach

3.1 Pedagogical Approach in Sports Education

The foundation for the interactions is the manager's, trainer's, or professor's pedagogical competency, outlined in the performance standards for education. A professor who exhibits strong didactic skills need not also be a skilled educator. A teacher must possess pedagogical competence alongside professional competency and specific topic instructional competency (trying to transfer the individual's content to his children in a manner they can comprehend as specified in curriculum and educational targets). Someone can establish a positive learning environment for his children's by being pedagogically competent [26-32]. They keep tabs on their children’s’ progress in terms of their cognition and memory and then adjust their actions accordingly. It aids in their children’s’ growth on the interpersonal, psychological, and ethical fronts. This one is transcribed to the qualifications (e.g., "he is acquainted with theoretical approaches as well as behaviorism that are pertinent for his teaching process and that he is capable of applying those to his pedagogy actions"). It is also transcribed toward the skills needed. Physical education (PE) academic interventions frequently concentrate more on the professor's teaching knowledge than on their abilities in terms of qualities and instructional measures [33-39].

3.2 Display Style Preprocessing and Frame Extraction of Video from Instructor

Therefore, the gray-level co-occurrence matrix, or grey-level size zone matrix (GLSZM), is being utilized to investigate the texturing, which includes spatial relationships between pixel elements. The place to start is the gray-level SZM innovative concept, which is predicated on every flatness zone's dimension co-occurrence (linked pixel that has the same gray level) [40-46]. However, second-level metrics only consider a bitmap image, although it examines the connection among pairs of pixels (often neighboring). Assume that the visual image being examined is a rectangle, with short and quality pixels in the longitudinal and transverse directions. The gray tone present in each layer is quantified to quant levels. Letting hor_spax = {1,2,...,horx} ver_spay = {1,2,...,very} where horx is the horizontal spatial domain. If gr = {1,2,...,grayg} and let exceptionally be the vertical spatial domain, then grayg is the collection of quantized gray tones with the prefix "gray." We suggest building a numerous scheme with several matrices and then combining them into a single matrix rather than maximizing the
gray-level numbers $K$. The GLSZM method involves creating eight SZM for eight different gray-level quantizations, where $k = 1, 2, \ldots, 8$, then merging the resultant matrix with a weighted sum for an image file.

$$
\text{gauss\_fun}(\text{hor}_{x}, \text{ver}_{y}) = \sum_{k=1}^{8} \text{gauss\_fun}(f(\text{hor}_{x}, \text{ver}_{y})).
$$

(1)

Where $T$ is quantized in $N_k$ gray levels, and $\text{gauss\_fun}$ is computed from there. A Gaussian function with a center between $N_4 = 16$ and $N_5 = 32$ gray levels determine the rewards dispersion; this allocation penalizes the extreme values of gray-level numbers because low levels have little structural analysis and elevated amounts are dependent on noise. The amount of brightness regions $g_{m}$ at a distance $d$ further from the space support boundary is produced by the new statistical $\text{gauss\_fun}(\text{hor}_{x}, \text{ver}_{y})$ matrix element. Between the flat zone and the shape border, this length is the lowest Distance measure. In reality, the proximity function is computed for the entire textured supporting area to speed up computational effort.

$$
dist(a, b) = \inf\{dist(a, b) : b \in K\}
$$

(2)

where $\text{dist}(a$ and $b)$ is are frequently calculated by applying a discontinuous metric approximation to the Distance function (Chamfer or Montanary). The corresponding data is utilized and then calculated for each area reg as its lowest value in the proximity map [47-55]. The suggested method employs the GLCM to calculate just 4 texture features using the Thibault matrices approach.

Contrast

$$
f_1 = \sum_{i,j} |i - j|^2 p(i, j)
$$

(3)

Correlation

$$
f_2 = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)}{\sigma_i \sigma_j} p(i, j)
$$

(4)

Angular second moment

$$
f_3 = \sum_{i,j} p(i, j)^2
$$

(5)

Diagonal distribution

$$
f_4 = \sum_{i,j} \frac{p(i-j)}{1+|i-j|}
$$

(6)

Where $\mu_i$ and $\mu_j$ stand for the average value and $\sigma_i$ and $\sigma_j$ for the confidence interval, respectively. Feature A texture’s change in gray levels can be identified using the $f_1$ measure of contrast between a pixel and its neighborhood over the $W_{m,n}$ subbands. A texture can also be identified using the $f_2$ measure of correlation between one pixel and its neighbors over the $W_{m,n}$, the $f_3$ measure, which denotes the sum of squared elements over the $W_{m,n}$ and is also known as uniformity [56-63].

3.3 Frame Selection Using Histograms of Oriented Optical Flow

The frames retrieved in the first phase of the suggested method must be chosen in the second stage to fit the HAR system’s input size. Additionally, picking the cropping area randomly could be a better and more straightforward strategy. The optimal method is to locate the human subject in the frame and then use human location and action-related information to choose the cropped area. However, the significant computational burden of this strategy makes it undesirable. To reduce the required computations, all the estimated discrepancies between standardized frames are carried over to this stage and used to create an energy map of the condensed video. Every pixel’s number on this mapping corresponds to the total amount of pixel measurements throughout all frames. The created radiation map is then subjected to an average filter. The ultimate image is generated when the input length of the employed system is specified as equal to the average filter’s window size. The center of the cropping region is then chosen to be the pixel with the highest value in the filtered image. It is simple to argue that this window contains more pertinent motion information than any other window that can be defined in the film. This information was discovered based on the gradient change between frames.
Assume the feature extracted data set contains J lecture frames $a_1, a_2, ... a_J$ and $b_1, b_2, ... b_J$. In the meantime, the M parameters correlating to the sound trademark feature are characterized as $\text{sig}_{\text{f}}(a_m, b_n), \text{sig}_{\text{f}} = 1, 2, ... M$, and the $\text{Ker}_\text{fun}(G)$ kernel functions correlating to the Histogram of Oriented Gradient (HOG) characteristic are described as $\text{Ker}_\text{fun}(G)(a_m, b_n)$ the $\text{Ker}_\text{fun}(F)$ kernel features correlating to the Histogram The following equations can be used to define the linear model of the kernel function combining the three properties above:

$$S = -\sum_{k=1}^{M} p_k \log_2 p_k$$  \hspace{1cm} (7)$$

Equation (6) meets the conditions $\phi_g \geq 0 \forall g, \phi_f \geq 0 \forall f, \phi_m \geq 0 \forall m, \sum_{g=1}^{G} \phi_g + \sum_{f=1}^{F} \phi_f + \sum_{m=1}^{M} \phi_m$. The values of the associated basis functions are expressed as $m$. When isolating crucial frames, conservative principles will convey the shot's substance as completely as possible. If all of the multiple images at each instant are used when evaluating a video, there will be an excessive amount of duplicate consecutive frames.

3.4 Estimation of Postures Using Continuous Filter Convolution Neural Network

This process seeks to calculate the frame movements for CNN information fusion and postural identification. The surface-camera proximity and the sensor range of vision were unchanged throughout the sessions. Thus, the clips featured translation. In order to calculate template matching between subsequent image pairings, this strategy uses a block-based motion prediction model. The interpolation $\text{mot}_{\text{vec}}(i)$, based on template matching, is the distance between a central inner block region in a frame (i) and its best match within a searching range in frame (i) + 1. In order to do this, the pairing requirement is the sum of absolute errors (SAD) of image intensity. The displacement $\text{mov}_i$ from frame (i) to frame (i) + k equals $\text{mot}_{\text{vec}}(i) + \text{mot}_{\text{vec}}(i+1) + ... \text{mot}_{\text{vec}}(i+k-1)$ for $k > 0$ when all motion vectors are present. At every stage, regions of each image sequence of 120 by 120 pixels are scanned in raster scan order with steps of eight pixels. Then, every posture is estimated by CNN. Every scanning has a 2-D offset that starts at (0, 0) and ends at (7, 7). A 120 x 120 R, G, and B channel image patch serves as the input for the 3-D data. Every component is linearly scaled during image standardization to achieve a mean of zero and unit L2 norm. Following the application of several layers, a SoftMax layer forecasts the poses.

Figure 2. The architecture of Continuous filter convolution neural network (ConFil_CNN)

The test error was increased until it stopped improving, at which point the rules adjusted the hyperparameters. The convolution layer extracts image edges and corners of various frequencies and orientations by performing a 3-D convolution with several kernels (i.e., filters) that work as an impulse response reaction filtration with such a specified stride (i.e., step size). The specific kernel with sections of black lines serves as an edge detector to identify crack features. Some irregularly patterned kernels function as textural feature extractors that can aid in separating splits from the backdrop. For down-sampling data and performing a transform locally, max pooling executes a 3-D maximum filtration with such a predetermined stride. An average pooling layer is a regularization term that facilitates faster training by allowing more excellent training data and enhancing Performed. Each network's data is transformed linearly to produce a distribution with zero mean and random values. The following is a mathematical formula for the convolution in layer 1. Fig.2 shows the architecture of a Continuous filter convolution neural network (ConFil_CNN).

$$C_j = f \left( \sum_{m=1}^{M} m \times k_j + b_j \right)$$  \hspace{1cm} (8)$$

Where $K_j$ and the bias of the j-th convolution kernel are, correspondingly, the convolutional kernels for $j=1, 2$, and F1. The input layer is $f$, too. This convolutional layer's convolution kernel size is F1. F1 kernels can generate F1 local features. The layer l's pooling procedure is displayed in

$$S_j = \beta_j \downarrow \text{c}(j) + b_j$$  \hspace{1cm} (9)$$
where down(n) is the subsampling operation, blue is the bias, and j (j=1, 2, ..., F1) is the multiplicative bias of the j-th pooling. The observed values and predicted value of the Classification algorithm at epoch tk of stainless steel I am, correspondingly, the mean squared error with L2 normal and the CNN model's transfer functions. The quantity of practice data is M. Additionally, an L2 regularization constraint is included to stop the overfitting issue. An ELU performs better than any other input signal as an activation function layer. The information is flattened by the initial completely connected layer, while the second one acts as the highest classification. The final two scores—i.e., decision values—are provided by the SoftMax layer. The two scores add up to one and vary from 0 to 1. If the probability of being a crack (signified as sc) is more than 0.5, the CNN classifies the input as a split patch; else, it classifies it as a frame correct patch. A dropout layer periodically disconnects some interconnections throughout the training stage to avoid overfitting.

IV. PERFORMANCE ANALYSIS

The performance of our proposed Continuous filter convolution neural network (ConFil_CNN)) is carried out compared with three state-of-art methods as Recurrent Neural Network (RNN) [12], Joint Neural Network (JNN) [13] and Internet of Things driven Physical Activity Recognition System (IoT-DPARS) [15] in terms of parameters such as accuracy, precision, recall, F1-score and response time.

<table>
<thead>
<tr>
<th>Number of users</th>
<th>Accuracy (%)</th>
<th>Precision (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RNN</td>
<td>JNN</td>
</tr>
<tr>
<td>100</td>
<td>86.5</td>
<td>76.5</td>
</tr>
<tr>
<td>200</td>
<td>87.6</td>
<td>76</td>
</tr>
<tr>
<td>300</td>
<td>86.5</td>
<td>77.8</td>
</tr>
<tr>
<td>400</td>
<td>86</td>
<td>77.9</td>
</tr>
<tr>
<td>500</td>
<td>87.4</td>
<td>77</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>number of users</td>
</tr>
</tbody>
</table>

**Figure 3. Comparison of accuracy**

Fig.3 depicts the accuracy comparison for the existing RNN, JNN, and IoT-DPARS with the proposed ConFil_CNN. X-axis and Y-axis show the number of users and the values obtained in percentage, respectively. When compared, existing RNN, JNN, and IoT-DPARS methods achieve 87.3%, 77.8%, and 83% of accuracy, respectively. In comparison, the proposed ConFil_CNN method achieves 98.5% accuracy, which is 11.2% better than RNN, 21.3% better than JNN, and 15.2% better than the IoT-DPARS method.
Figure 4. Comparison of precision

Fig. 4 depicts the precision comparison for the existing RNN, JNN, and IoT-DPARS with the proposed ConFil_CNN. X-axis and Y-axis show the number of users and the values obtained in percentage, respectively. When compared, existing RNN, JNN, and IoT-DPARS methods achieve 72.4%, 82.3%, and 88.5% of precision, respectively. In comparison, the proposed ConFil_CNN method achieves 96.7% of precision which is 11.2% better than RNN, 24.3% better than JNN, and 15.2% better than IoT-DPARS method.

Table 2. Analysis of recall and F1-score

<table>
<thead>
<tr>
<th>Number of users</th>
<th>Recall (%)</th>
<th>F1-score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RNN</td>
<td>JNN</td>
</tr>
<tr>
<td>100</td>
<td>66</td>
<td>87</td>
</tr>
<tr>
<td>200</td>
<td>67.5</td>
<td>86.6</td>
</tr>
<tr>
<td>300</td>
<td>67.4</td>
<td>87.9</td>
</tr>
<tr>
<td>400</td>
<td>66.6</td>
<td>86.5</td>
</tr>
<tr>
<td>500</td>
<td>67</td>
<td>86</td>
</tr>
</tbody>
</table>

Figure 5. Comparison of recall

Fig. 5 depicts the recall comparison for the existing RNN, JNN, and IoT-DPARS with the proposed ConFil_CNN. X-axis and Y-axis show the number of users and the values obtained in percentage, respectively. When compared, existing RNN, JNN, and IoT-DPARS methods achieve 67.8%, 87.4%, and 72.4% of recall, respectively, while the proposed ConFil_CNN method achieves 93.5% of recall which is 26.2% better than RNN, 14.5% better than JNN, and 15.2% better than IoT-DPARS method.
Fig. 6 depicts the F1-score comparison for the existing RNN, JNN, and IoT-DPARS with the proposed ConFil_CNN. X-axis and Y-axis show the number of users and the values obtained in percentage, respectively. When compared, existing RNN, JNN, and IoT-DPARS methods achieve 75.6%, 84%, and 75% of the F1-score, respectively, while the proposed ConFil_CNN method achieves 95.2% of recall which is 10.4% better than RNN, 11.2% better than JNN, and 10.2% better than IoT-DPARS method.

Table 3. Analysis of response time

<table>
<thead>
<tr>
<th>Number of users</th>
<th>RNN</th>
<th>JNN</th>
<th>IoT-DPARS</th>
<th>ConFil_CNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>34.5</td>
<td>41.3</td>
<td>23.4</td>
<td>12.3</td>
</tr>
<tr>
<td>200</td>
<td>35</td>
<td>41</td>
<td>22</td>
<td>11</td>
</tr>
<tr>
<td>300</td>
<td>35.6</td>
<td>43.2</td>
<td>22.1</td>
<td>13.2</td>
</tr>
<tr>
<td>400</td>
<td>35.7</td>
<td>45</td>
<td>22.7</td>
<td>12.2</td>
</tr>
<tr>
<td>500</td>
<td>36</td>
<td>43.4</td>
<td>22.6</td>
<td>12</td>
</tr>
</tbody>
</table>

Fig. 7 depicts the response time comparison for the existing RNN, JNN, and IoT-DPARS with the proposed ConFil_CNN. The X and Y axes show the number of users and the values obtained in milliseconds. When compared, existing RNN, JNN, and IoT-DPARS methods achieve 36.7ms, 45.3ms, and 22.5ms of response time, respectively, while the proposed ConFil_CNN method achieves 12.4ms of response time which is 34.5ms better than RNN, 33.1ms better than JNN, and 12.1ms better than IoT-DPARS method. Table 4 shows the overall comparative analysis.

Table 4. Overall comparative analysis

<table>
<thead>
<tr>
<th>Parameters</th>
<th>RNN</th>
<th>JNN</th>
<th>IoT-DPARS</th>
<th>ConFil_CNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1-score (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>number of users</td>
<td>100</td>
<td>200</td>
<td>300</td>
<td>400</td>
</tr>
<tr>
<td>response time (ms)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
V. CONCLUSIONS

This study discusses the continuous filter convolution neural network-based online learning system for college physical education teaching. The first part of this essay provides background information on online learning and its research significance. Previous techniques for estimating human posture typically involve nonlinearly mapping the image to be analyzed to the locations of various components or joints. We merged with the deep learning algorithm, managed to sort out the statistical model of sporting events footage or photos, and especially in comparing them with both the precise control of genuine training to achieve the benchmark of the motions of educators in PE class in order to safeguard the nonstandard sporting events motions in Physical education class. The above article develops and implements the conceptual model and layout of the system based on the planning fundamentals of the online learning platform for physical education in schools, which enhances the system's overall performance and satisfies the requirements for the delivery of physical education in colleges, and has positive social effects. To greatly increase reliability, the focus must be placed on methods for selecting features using metaheuristic modeling techniques.

REFERENCES


| Accuracy (%) | 87.3 | 77.8 | 83 | 98.5 |
| Precision (%) | 72.4 | 82.3 | 88.5 | 96.7 |
| Recall (%) | 67.8 | 87.4 | 72.4 | 93.5 |
| F1-score (%) | 75.6 | 84 | 75 | 95.2 |
| Response time (ms) | 36.7 | 45.3 | 22.5 | 12.4 |


[49] Li, A.; Yu, S.; Qinggang, F.; Yongmei, Z. Exploration of the mixed teaching mode of “three classes” under the intelligent teaching-based on computer programming course. Education Study, 2020, 2, 33–42.


