Research Article

Comparative Analysis of Volleyball Serve Action Based on Human Posture Estimation

Bo Zhou¹ and Omer Saeed²

¹School of Physical Education, Southwest Minzu University, Chengdu 610000, Sichuan, China
²Institute of Management, Kyrgyz International Universal College, Bishkek, Kyrgyzstan

Correspondence should be addressed to Omer Saeed; omersaeed@mail.cu.edu.kg

Received 20 June 2022; Revised 7 August 2022; Accepted 9 August 2022; Published 30 September 2022

Copyright © 2022 Bo Zhou and Omer Saeed. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Serving is one of the most crucial techniques in volleyball. Serving is a method that does not require team interaction and is difficult for the opponent to immediately interfere with. The feature migration module with a fixed offset is suggested in this work. This module can be thought of as a cross-channel dilated convolution approximation of dilated convolution. The reason cross-channel dilated convolution is not worse than standard dilated convolution with few parameters is discussed in this article. An improved random forest model is put forth to address the issue of the human pose estimation system’s high memory consumption when utilizing random forest as the classifier. This model presents the Poisson process and incorporates it with the depth data to create a filter before using Bootstrap sampling. In order to optimize and reconstruct the training dataset, a portion of the feature sample points that do not contribute positively to subsequent classification is removed from the original training dataset. This allows the training dataset to better account for the repeated sampling of the random forest during the sampling process. Resampling has some drawbacks, but they are not very representative. The effectiveness of the optimization model, which significantly lowers the system’s time and space complexity and increases the system’s applicability, is demonstrated by experiments.

1. Introduction

Taking an overview of today’s volleyball world, the techniques and tactics in volleyball games are developing in the direction of “comprehensive, high, fast, flexible, and changeable.” Volleyball is one of the better-developed sports in my country. With the Chinese women’s volleyball team standing on the championship podium again after many years at the Rio Olympics in 2017, the Chinese people’s attention to volleyball has increased significantly [1]. Mass sports are carried out in various forms. While participating teams continue to use new technologies, new tactics, and new styles of play, they also put forward higher requirements on the overall abilities of volleyball players [2]. Excellent volleyball players can maximize the accuracy and offensive characteristics of serving, which is undoubtedly an effective supplement to the overall offensive strength of the team [3]. The precision of the athlete’s serve technique must therefore be continually improved during the volleyball training process if they are to win the match. Technical and psychological training are the sources of the elements impacting volleyball serve accuracy. The effectiveness of an athlete’s serve and his degree of expertise are closely related, and it also depends on subjective elements like athletes’ psyche and is influenced by objective factors like the venue’s equipment, environment, and setting [4].

Human body pose estimation is the process of detecting a given image or video and calibrating the key parts or main joint points of the human body in the video or picture [5]. Since the concept of pose estimation was put forward, scholars at home and abroad have conducted a lot of research for this [6, 7]. When estimating the human body pose, two main solutions are used, namely, the traditional image processing method and the method based on deep learning...
Although the traditional method has high time efficiency, it mainly extracts SHIFT features and HOG features, and these features are manually set, which will lead to different viewing angles, appearances, occlusions, and their inherent geometry in the image [9]. Fuzziness will have a certain impact on the performance of the algorithm, so that the full use of image information cannot be realized. Aiming at the influencing factors of the serving technique, while training the athletes’ serving technique, the combination of sports psychological indicators and the accuracy of volleyball serve for investigation and analysis is not only conducive to the improvement of the athletes’ serving skills but also can be used for coaches and teachers in volleyball players [10]. It applies the depth perception ability and muscle exertion of volleyball students on the accuracy of volleyball serve, and improves the technical level of volleyball students and the application of psychological principles [11]. It can play an active role in the training of volleyball technical skills, and at the same time, it can provide a theoretical basis for solving problems in the service of volleyball students in training, provide a theoretical basis for future volleyball players’ psychological selection indicators, and improve the psychological principles [12]. The theoretical system applied to the learning of sports technical skills provides research inspiration.

Translating other spatial location features to the location is an easy and direct way to increase the model’s receptive field. We suggest a method for translating each channel feature map by a predetermined offset, enabling the integration of each point with other location features at predetermined periodic intervals. We categorize this technique as a module, a module for long-distance relationships. This module may be thought of as a particular variant of dilated convolution. In order to improve the random forest model in the experiment and analysis sections, this research highlights the drawbacks and inadequacies of the random forest classifier in the human pose estimation system. The system’s time and space usage before and after the forest model optimization are compared. According to the experimental data, this optimization is quite successful in decreasing the system’s time and space complexity without significantly affecting the accuracy of system identification. Analyzing the test results reveals that there was no statistically significant variation in the movement skill difficulty between the three groups and that the subject’s serving motion essentially took the shape of an automatic frontal overhand serve. For the jumping technique where the skill difficulty is medium and the subjects are still in the stage of differentiation to automation, the three training groups also showed no significant differences. Among them, sequence exercises have the most obvious improvement in accuracy and total score, while random exercises are still only the best for improving motor skills. Jumping ball is a technique between basic and difficult serving techniques in volleyball. It is not easy to send high-quality jumping ball, even for volleyball students. The sequence practice mode with moderate background interference can enable the subjects to correct some wrong actions in time during the whole training process and establish the correct action mode, so that the subjects can be allowed to perform under the premise of moderate background interference.

The rest of this article is organized as follows. Section 2 discusses related work. Section 3 analyzes the global feature extraction method based on feature map migration. Section 4 designs the human pose estimation strategy. In Section 5, a comparative analysis of the experimental results of the volleyball serve action was carried out. Section 6 summarizes the full text.

2. Related Work

Generally, two solutions are mainly used when estimating the human body pose, that is, the traditional image processing method and the method based on deep learning [13]. Among them, the method based on deep learning is the mainstream method used now. In the traditional method, some tree diagram models can be used to express the joint structure of the human body, and these models have dynamic a priori. They divide the joints of the human body, which can be divided into the head, trunk, upper and lower left arm, upper and lower right arm, upper and lower left leg, and upper and lower right leg. Related scholars have proposed a very classic graph structure algorithm [14]. The author uses a collection of multiple parts to express the human body, and there are certain spatial constraints between these parts. The graph structure mainly includes two parts: a space model and a component model. The component model is used to describe the various components that make up the human body.

Related scholars have proposed a greedy parts allocation algorithm, which can utilize the inherent structure of the human body, thereby reducing the complexity of the graphic model [15]. In addition, in order to fuse feature information at different scales, the existing network structure often performs repeated upsampling of pictures, resulting in low spatial sensitivity of the final high-resolution features [16]. In response to this phenomenon, related scholars have proposed a high-resolution representation network, which uses a distinctive parallel structure to ensure that the resolution is not reduced in different stages of the network [17]. Researchers have proposed a method of posture refinement, which is model-independent and does not require prior knowledge and code support from other algorithms, which makes it easy to complete related operations during post-processing [18, 19]. Experiments show that, compared with our commonly used multistage network structure, this method has a better detection effect [20]. In addition, related scholars have proposed a new method using bottom-up thinking to complete the estimation of multiperson 2D human body pose [21]. This method uses the “parts intensity field” to locate the various joint points of the human body. In addition, in order to obtain a complete body posture, the
aforementioned “part affinity field” is used to associate the various joint points belonging to the same person [22]. This method has very good applicability for application scenarios such as delivery robots and unmanned driving.

Relevant scholars have summarized the inadequate completion of the volleyball player’s serve technique and smash technique in sports and believe that the higher the athlete’s bounce when smashing the ball, the better, because the higher the jump, the smashing will have more lethality [23, 24]. However, the fact is that athletes often do not jump high when doing the above two technical actions. This is mainly related to the lack of muscle ability when jumping. The buckle and serve skills need to have the early approach speed. There are certain requirements for athletes’ waist and leg muscle abilities, and they need to have strong coordination and contractility. In response to this shortcoming, scholars also proposed that short-distance running and weight load can be used to exercise this part of the muscle strength [25–27]. The second aspect is that the error of the spiking action is due to the premature squatting when hitting the ball, which cannot be normal. Under normal circumstances, when volleyball takes off, the coordination of each link has a great influence on the completion of the take-off. Relevant scholars have mentioned in the research that the extension of the knee cannot be fully exerted when the human body is stepping and jumping [28, 29]. At this time, the cooperation of the ankle joint is urgently needed. In short, the cooperation of the lower limbs of the human body is crucial to the effect of jumping.

3. Global Feature Extraction Method Based on Feature Map Migration

3.1. Baseline Model. Instead of using regression models in our approach, we employ categorization models to forecast where the important body parts will be located. The location of important sites in the human body is predicted by this model using categorization rather than regression. For each pixel classification, it tries to ascertain the existence and type of key points. The deep CNN needs to maintain an output scale that is similar to the input scale in order to accomplish pixel-level prediction. In comparison to conventional CNNs, this characteristic can be obtained with fewer pooling layers, although less pooling necessitates higher computing and storage costs in the deep layers of the network. U-shaped networks are increasingly being used in computer vision techniques to attain this goal [30]. The U-shaped network can be divided into two stages. Its shallow layer is the first stage, the downsampling stage. The network is convolved and downsampled like the traditional CNN network. The feature map is downsampled to a relatively small size to extract features and reduce the optimization space [31, 32]. This stage is followed by an upsampling stage; some upsampling layers and convolutional layers expand and shrink the feature maps together. Subsequently, the upper-level sampled feature map is added to the downsampling stage feature map of the same size. This jump addition method is usually called a shortcut connection [33–35].

3.2. Dilated Convolution. Let S denote the discrete 2D spatial location on a feature map, and let F denote the feature map defined on SE. Let \( \Omega r = [-r, r] \times [-r, r] \) represent the offset space of a convolution kernel element, that is, the horizontal or vertical spatial offset of each convolution kernel should be between \( -r \) and \( r \) (including boundary). We define \( \Omega r \rightarrow R \) as a convolution kernel with size \( (2r + 1) \times (2r + 1) \). Then, the dilated convolution can be expressed as follows:

\[
F_{k,t}(p) = \prod_{p=lt-s} F(s)Zk(t),
\]

where \(*1\) is called the l-dilated convolution (l-DConv), and \( l \in Z^+ \) is the dilation step.

We extend this definition to the concept of convolutional neural networks. For a convolutional layer with \( M \) input feature maps and \( N \) output maps, the expanded convolutional layer can be written as follows:

\[
G_n(p) = \prod_{m=0}^{M-1} k_{m,n}F_{m,t}(p)n = 0, 1, 2, \ldots, N - 1.
\]

3.3. Cross-Channel Dilated Convolution. Assuming that \((2r + 1)^2\) is a factor of \( M \), we divide a total of \( M \) feature maps into \( W \) groups so that \( M = W \cdot (2r + 1)^2 \) is satisfied. Assume that adjacent feature maps are grouped into the same group. Note that the concept of “group” here is orthogonal to the “offset group” mentioned in the feature migration process, that is, the \( w \)th group is equivalent to the \( w \)th element in all \( H \) offset groups. Our \( w \)th feature group can be expressed as follows:

\[
F^{(w)}_i(p) = F_{(w-2)(i+1)}(p) \quad i = 0, 1, 2, \ldots, H - 1.
\]

We define the cross-channel expansion convolutional layer as follows:

\[
G_n(p) = \prod_{w=0}^{W-1} [k_{nw}F^{(w)}_{wl}](p),
\]

where \( \otimes 1 \) is the l-cross-channel dilated convolution (l-XDCov) operation, assuming that the finite set \( \Omega r \) is ordered, and \( i(t) \) is the \( i \)th element of \( \Omega r \). Note that \( l(i) \) is equivalent to \( u \) used in the feature migration process. Unlike traditional convolution or dilated convolution, the cross-pass dilated convolution operation applies the \((2r + 1) \times (2r + 1)\) convolution kernel to \( H = (2r + 1) \) channels.

\[
[k_{nw}F^{(w)}_{wl}](p) = \prod_{p=x-lt(i)} k(l(i))F^{(w)}_{i}(s).
\]

The traditional convolution kernel and the traditional expansion convolution kernel focus on extracting the local features in each feature map and weighting the sum of the extracted features. They can be regarded as a broader version of cross-channel dilated convolution because they can achieve the same operation as cross-channel dilated convolution by setting the weight of each convolution kernel to...
only one nonzero, that is, the in-channel operation of the traditional convolution is turned off.

4. Human Pose Estimation Strategy

4.1. Deep Feature Selection and Poisson Process. Poisson distribution is a common discrete probability distribution in statistics and probability. Poisson distribution is suitable for describing the number of random events occurring in unit time, and its probability distribution function is as follows:

\[ P(X = k) = \frac{(\lambda)^k e^{-\lambda}}{k!} \quad \lambda > 0. \quad (6) \]

The parameter \( \lambda \) represents the average occurrence rate of random events in unit time (area). If the random variable \( N(t) \) represents the total number of "events" that have occurred until time \( t \), then the random process \( \{N(t), t \geq 0\} \) is called the counting process.

The number of events in any interval of length \( t \) obeys Poisson distribution with rate \( \lambda t \) as the mean, that is, for all \( s, t \geq 0 \) satisfies

\[ P[N(s) − N(s − t) = n] = (n!)^{-1}e^{-\lambda t} (\lambda t)^n. \quad (7) \]

Then, the above counting process is called the Poisson process with rate \( \lambda (\lambda > 0) \). It can be seen that the Poisson process is the partial unitization of Poisson, that is, the Poisson process describes the count of the number of occurrences of events (number of occurrences of particles) in a unit time interval.

4.2. Poisson Optimized Random Forest. In this study, the Poisson process is introduced and used as a filter to filter the original sample dataset. First, we scan the image and calculate the filter factor \( L \) of all feature points in the image. Suppose the depth value of the feature point \( X \) to be operated is \( D_e \), and the depth values of the feature point \( X_1 \), the feature point \( X_2 \) on the left, and the feature point \( X_3 \) on the right, which are adjacent to and directly above the feature point \( X \), are \( D_{e1}, D_{e2}, \) and \( D_{e3} \), where the Euclidean distance is integrated. We calculate the Euclidean distance \( L' \) between the feature point \( D_e \) and \( D_{e1}, D_{e2}, \) and \( D_{e3} \) as in (8) and then use (9) to obtain the characteristic point \( X \) filtering factor \( L \). Then, we sort the filtering factors of all feature points in the image. Finally, the sorted filtering factor is regarded as a counting process, and the Poisson value of each feature point is calculated by formula (10), and the filtering threshold is set at the same time. According to the filtering threshold and the calculated value of each feature point, the size of the Poisson value is selected and left, so as to reconstruct the training dataset.

\[ L' = \frac{2}{i=0} (D_{e_i} - 2D_e)^2, \quad (8) \]

\[ L = D_e \bullet L', \quad (9) \]

\[ \text{Poisson}[L_u = N(u) − N(u - 1)] = (L_u)^\lambda e^{-\lambda}. \quad (10) \]

In formula (10), \( N(u) \) is the value of the \( u \)th filtering factor after sorting, and \( \lambda \) is the average of all filtering factors in the image.

4.3. Experiment and Result Analysis. In Figure 1, the first row represents the experimental sample pictures. There are 600 pictures in the sample library, which are divided into 3 categories according to the characters, and each category of pictures is a continuous action made by the same person. The second line is the experimental effect diagram before random forest optimization, and the third line is the experimental effect after optimization. The optimized random forest model did not cause the classification accuracy to drop too much. The specific accuracy comparison is shown in Figure 2.

Following optimization, the random forest’s recognition accuracy increases consistently as the number of decision trees increases. The more decision trees there are, the less the accuracy rate will fluctuate, making the system more stable. Prior to optimization, the system will run out of memory when there are 700 decision trees, causing it to crash. Therefore, once there are 700 decision trees, the matching graph from before optimization will stop. In order to verify the influence of the optimization of the random forest model on the accuracy of human pose estimation, this study designs a comparison of the accuracy of the model before and after optimization, as shown in Figure 2. Figure 2 shows that the optimization of the model has increased the recognition accuracy of the system to a certain extent, and at the same time, it has further improved the robustness of the system.

While ensuring that the accuracy of system identification will not be too reduced, this study designs a comparison of the system time costs before and after model optimization, as shown in Figure 3. Before model optimization, the time cost of the system increases with the increase of decision trees. When the number of decision trees reaches a certain number, the system cannot run on ordinary PCs. However, after the model is optimized, the time cost of the system has been greatly reduced compared with that before the optimization, and it also shows a more stable trend. The optimized random forest model has made great progress in the time cost of the original human pose estimation system.

This work designs a comparison of the memory used by the system operation, as shown in Figure 4, in order to demonstrate the contribution of the improved random forest model to the space overhead of the original system. Prior to model optimization, as the number of decision trees grew, the system’s memory overhead grew along with it. When the number of decision trees reached a certain point, the memory footprint reached roughly 1900MB, rendering standard PCs impossible to function. The optimized model has been kept between 750 MB and 1050 MB, with minimal adjustments, indicating a consistent trend. This prevents the system’s memory cost from rising as the number of decision trees increases. This demonstrates how the optimized random forest model has significantly reduced the original system’s space overhead.
5. Comparative Analysis of Experimental Results of Volleyball Serve

5.1. Subjects Maintain Test Scores. 72 hours after the end of the last training arrangement, the straight ball skill retention test was performed on the frontal overhand serve and jump ball practiced in this experiment. In order to discuss in detail the training effects of each training method for each serving technique and the overall training effects of each practice group’s serving skills after three weeks of practice, under the conditions of fixed technical difficulty factors, the front-hand serve and jump ball analysis of variance was performed.

5.1.1. Test Results of Frontal Overhand Serve in Different Practice Groups. First, the homogeneity test of the frontal overhand serve of the movement skills with lower operational difficulty is carried out, and the result is shown in Figure 5. Subsequently, a one-way analysis of variance was performed on the scores of the retention test of frontal overhand serve skills, as given in Table 1. The results showed that serving accuracy $F = 1.33$, $P = 0.176$, action technique $F = 0.178$, $P = 0.645$, and total score $F = 0.415$, $P = 0.507$; $P$ values were all greater than 0.05, and there was no significant difference between the three groups. Among them, the fixed practice group has the best serving accuracy and total score. The average serving accuracy reached 9.14, and the average
Figure 3: Comparison of system time. (a) Poisson optimized random forest. (b) Random forest.

Figure 4: Comparison of system memory usage. (a) Poisson optimized random forest. (b) Random forest.

Figure 5: The homogeneity test of frontal overhand serve in different practice groups.
total score reached 16.0, followed by the sequential practice group, with an average serving accuracy of 8.64. In comparison, the random practice group has the worst performance, with an average accuracy of 8.14 and an average total score of 15.22. However, in the performance of the action technique, the random practice group achieved the best performance with an average of 7.07, followed by the fixed practice group with an average of 6.74, and the sequential practice group has a poorer performance with an average of 6.70. In addition, the standard deviations of the three performances in the fixed practice group are all the lowest. The serving accuracy is 0.40, the action technique is 0.06, and the total score is 0.31. In the random practice group, the standard deviation of the serving accuracy is the same as that of the sequence practice group. The standard deviations of the technique and total score were both in the middle, 0.176 and 1.05, respectively, and the standard deviations of the sequence practice group were the highest, 0.77 and 1.34, respectively. It shows that the performance fluctuations in the fixed practice group are relatively small and stable, followed by the random practice group, and the sequential practice group has the largest fluctuations. Overall, the fixed practice group, followed by the sequential practice group and the random practice group, has relatively poor effects on low-difficulty front-hand serve practice. The frontal overhand serve performance during interference practice with varied backgrounds varies, but there is no statistically significant difference between them, according to the difference test of the three groups of measured data.

5.1.2. Jumping Ball Test Results of Different Practice Groups. Similarly, the homogeneity test of jumping ball with medium-difficulty action skills is carried out, and the result is shown in Figure 6. Subsequently, a one-way analysis of variance was performed on the retention test scores of jumping ball skills, as given in Table 2. The results showed that the $P$ value is greater than 0.05, and the three groups are not significantly different. Among them, the sequence practice group performed the best in serving accuracy and total score. The average serving accuracy was 9.64, and the total score was 16.57. The movement technique was second to the random practice group and higher than the fixed practice group, reaching 6.82; the mean value of serving accuracy is the same as that of the random practice group, which is 8.64, while the random practice group has the highest mean value of movement skills, reaching 7.04, and the total score is higher than that of the sequence practice group and higher than the fixed practice group, with an average value of 15.8, so the score is in the middle. Compared with the fixed practice group, the performance is the worst, with the average value of the movement technique and total score being 6.72 and 15.47, respectively. In addition, the standard deviation of serve accuracy in the sequence practice group is 0.4, the standard deviation of movement technique is 0.04, the standard deviation of the total score is 0.36, and the standard deviation of the three items is the lowest, indicating that the performance fluctuations in the group are relatively stable; the random practice group excludes the service. The standard deviation of accuracy is the same as that of the fixed practice group, both are 0.85. The standard deviations of the movement technique and total score are both higher than the sequence practice group and lower than the fixed practice group, which is in the median. The standard deviations are 0.06 and 1.01, respectively. The standard deviation of the action technique was 0.32, and the standard deviation of the total score was 1.25. The standard deviations of the two performances were the highest, indicating that the jumping skills in this group fluctuate greatly and are the most unstable. On the whole, the sequence practice group has the best results for the medium-difficulty movement technique jumping ball practice, followed by the random practice group. In comparison, the fixed practice effect is the least satisfactory. After the difference test of the three groups of measured scores, it can be seen that under different levels of background interference practice, the jumping ball with medium skill difficulty maintains scores, but there is no significant difference between them.

5.2. Subject Transfer Test Results

5.2.1. Test Results of Frontal Overhand Serve of Subjects in Different Practice Groups. We test the homogeneity of the frontal overhand serve of the movement skills with lower operational difficulty, and the results are shown in Figure 7. Table 3 shows the single-factor variance of the test scores for the transfer of frontal overhand serve skills. The $P$ values are all greater than 0.05, and there is no significant difference. Among them, the fixed practice group has the highest serving accuracy and average total score of 9.40 and 16.04, respectively. The action technique is the same as the sequence practice group, both are 6.54, next to the random practice group, which has an average of 7.07, and the random practice group has a relatively middle score. The average total score is 15.47, second only to the fixed practice group; the total score of the sequential practice group is poor, with an average of only 15.04. In addition, the three results of the fixed practice group have relatively small fluctuations within the group, and the standard deviations are the lowest, respectively, 1.00, 0.30, and 1.17. In the sequential practice group, except for the movement technique standard deviation of 0.63, which is higher than the random practice group of 0.34, the accuracy and the standard
deviation of the total score are both in the middle, 1.81 and 1.31, respectively; the random practice group has relatively unstable fluctuations in the group, the standard deviation of serve accuracy is as high as 3.00, and the standard deviation of the total score is as high as 3.16. On the whole, the fixed practice group has the best results for the low-difficulty technique of frontal overhand serve, followed by the random practice group. In comparison, the sequential practice effect is the least satisfactory. After the difference test of the three groups of measured scores, it can be seen that under different levels of background interference practice, there are high and low scores for the transfer of frontal overhand serve with lower skill difficulty, but there is no significant difference between them.

5.2.2. Floating Ball Test Results of the Subjects in Different Practice Groups. Figure 8 shows the outcome of the homogeneity test of the leaping ball using a medium-difficulty action skill. The results of the leaping ball skill transfer test’s single-factor variance are displayed in Table 4. According to the data, there was a fairly noticeable difference between the three groups that practiced jumping and floating. Among them, the fixed practice group has the best results in three performances, with a serve accuracy of up to 10 points, a movement technique...
of 7.04, and a total score of 17.04, followed by the sequence practice group, except that the average movement technique is slightly inferior to the fixed practice, which is 6.64. The other two average scores are in the middle, the serving accuracy is 9.00, and the total score is 15.64; the fixed practice group scores are the most unsatisfactory, except for the average movement technique only 0.02 points higher than the sequence practice, the average scores of the other two scores ranked last. The average serve accuracy is 7.64 and the average total score is 14.41. In addition, the results in the random practice group are relatively stable. Except for the standard deviation of the movement technique, which is only higher than the fixed exercise by 0.02, the other two grades have the lowest standard deviation; the sequence exercise also has the highest standard deviation of the movement technique. It is higher than the random practice group, and the score distribution within the group is in the middle; the fixed practice group has a lower standard deviation of movement techniques than the other two groups, and the service accuracy and the standard deviation of the total score are both the highest, indicating that the results within the group are fluctuating largely.

### 6. Conclusion

The relationship between this module and the expansion convolution is demonstrated in this article using a fixed offset translation feature map. This module’s operation method is referred to as cross-channel expansion convolution. The purpose of the cross-channel dilated convolution is discussed in this article, and it is demonstrated that it performs similar to the dilated convolution while having significantly smaller parameters. Prior to the sampling operation in the training of the random forest classifier, the depth information and the Poisson process theory are combined to create a filter net. The filter’s purpose is to exclude other training data that will not be helpful for the follow-up work while keeping the sample data that will be kept from the training dataset that will be sampled. To compensate for the random forest classifier’s repetitive sampling and resampling samples’ inadequate representativeness, the training sample dataset can be rebuilt.

The purpose of the experiments was to assess the accuracy of recognition and the time and space consumption before and after model optimization. The results demonstrate that this optimization is highly efficient, drastically lowering the time and spatial complexity of the entire system while maintaining a relatively high level of identification accuracy and increasing the system’s applicability. The frontal overhand serve and jumping ball indications did not significantly differ between the fixed practice group, sequence practice group, and random practice group in the skill maintenance test \( (P > 0.05) \). The three groups differ significantly in serve accuracy and overall score during the ferocious leap serve \( (P < 0.05) \). The sequence practice group is...
in the middle, the random practice group is less effective, and the fixed practice group has the best results. The action technical indicators do not significantly differ from each other \( (P > 0.05) \). The three groups’ serving accuracy was significantly different in the overall score \( (P = 0.05) \). The fixed practice group performed the best, followed by the sequence practice group, which performed averagely, and the random practice group, which performed poorly.

**Data Availability**

The data used to support the findings of this study are included within the article.

**Conflicts of Interest**

The authors declare that they have no conflicts of interest.

**Acknowledgments**

This study was supported by the Southwest Minzu University (1231119059).

**References**


