
Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 6, pp. 5302–5312, DOI 10.12694/scpe.v25i6.3249

THE APPLICATION OF DEEP LEARNING IN SPORTS TRAINING

FANGLING YAN* AND QIUPING PENG†

Abstract. In response to the problems of overfitting, susceptibility to interference information, and insufficient feature expression ability in existing deep learning methods for sports action recognition, the author proposes a deep learning sports action recognition method that integrates attention mechanism. This method proposes a video data augmentation algorithm in data preprocessing to reduce the risk of model overfitting. Then, during the video frame sampling process, the existing sampling algorithms are improved to effectively suppress the influence of interference information. In the special section, the network residual consolidation is proposed to improve the feature extraction capacity of the structure. The Long Term Time Transform (LSTM) network is used to solve the problem of the global correlation of the spatial correlation, and the classification algorithm is achieved by Softmax. and the classification algorithm is proposed. The experimental results show that the recognition rates of this method on UCFYouTube, KTH, and HMDB-51 data are 96.73%, 98.07%, and 64.82%, respectively.

Key words: Deep learning, Sports training, Residual network, Attention mechanism, Long short-term memory network

1. Introduction. Identification and prediction of human body mass can provide useful supporting information for sports training [1]. By obtaining the relevant information of human movement and combining it with the data structure, adjusting the content of sports can improve their sports level. With the continuous development of computer technology, intelligence technology is also gradually growing.Deep neural networks (DNNs) can study data features to distinguish and distribute large-scale data, but the latest techniques still can not be directly used to capture human data. In order to capture human motion data, the resolution filters have to cover all human joints, so the resolution only occurs in the temporal direction [2,3]. DMP is a special hardware of IMU equipment which can calculate quaternion data by reading sensors. IMU directly receives the data from the auxiliary system, allowing the system design process to process the sensor data fusion without the interference of the system application process.

The purpose of the recognition algorithm is to determine the action taken by the human body in the video. Because of its application in intelligent building, intelligent security, human-computer interaction, video retrieval and so on, it is a very challenging research topic. In the field of behavior recognition, people have written a lot of textbooks, and carried out a lot of experiments on human contour, human node, space-time interest point and motion parameters. The traditional teaching material extraction method is too dependent on the content of the teaching material, and its anti-interference and generalization ability are not strong, so it is difficult to popularize. In contrast, deep learning can study individual personality traits more effectively. On this basis, a feature extraction method based on deep learning is proposed.

Sports is an activity with profound cultural heritage, which can not only exercise the body, but also improve the spiritual quality. The sports humanistic spirit is a kind of moral concept and humanistic spirit in sports, and it is an indispensable factor in sports. Among them, personality cultivation, as an important aspect of sports humanistic spirit, has a unique connotation and value.Personality cultivation refers to the behavior habits, ideas and moral level formed by an individual in social communication, which is the most important and basic quality in a person's life. In sports, personality cultivation is regarded as a value concept. In sports, justice is an indispensable value concept. In a competition, the referee needs to cut the result fairly, and the athletes need to follow the rules and respect the opponent. In training, honesty is even more important. Sports is a process of constantly challenging themselves. Athletes constantly transcend the limits and create new records.

*Dongchang College of Liaocheng University, Liaocheng, 252000, China
†School of Physical Education and Health Science, Guangxi Minzu University, Nanning, 530006, China (Corresponding author, pqp2022@126.com)

This spirit of challenging the limit not only exists in competitive sports, but can also be applied in daily life. Challenge their own limits, need to have a kind of indomitable spirit and perseverance.

Athletes need to work hard in every training and competition, and be indomitable in the face of difficulties. Only in this way can we go beyond our limits and achieve faster, higher and stronger goals. The spirit of breaking the limit is not only the essence of sports, but also an important spirit in life. In study, work and life, only by constantly challenging their own limits, can we make continuous progress and achieve the maximization of self-value. Therefore, the spirit of breaking through the limit occupies an important position in the humanistic spirit of sports.

Another connotation of sports humanistic spirit is team cooperation. In sports, whether it is team competition or individual competition, athletes need to cooperate and support each other. Only on the basis of team cooperation, can we achieve the best results and performance.Teamwork is not only in the competition, but also in the daily training. In the training, the team members need to cooperate with each other, encourage each other, help each other, to overcome difficulties, overcome their own weaknesses, and constantly progress. Teamwork is also a value and a spirit. Only on the basis of team cooperation, can we realize the maximization of self-value, and can better complete various tasks in work and life. In the team cooperation, everyone's contribution is indispensable, only everyone can do the best, can achieve the best results and achievements. Team work is an indispensable part of sports humanistic spirit, which emphasizes team spirit and cooperation ability, and is a positive, unity and cooperation spirit.

Sports training is the key link to cultivate athletes 'physical quality and competitive level, and also an important way to cultivate athletes' personality cultivation. In sports training, through participating in different forms of sports activities, athletes can learn and exercise their own conduct and moral character, and constantly improve their moral level and psychological quality. Through sports training, athletes have learned to make unremitting efforts, overcome difficulties and overcome themselves, as well as life attitudes and values such as compliance with rules, respect for opponents, discipline and civilized competition. At the same time, sports training can also cultivate athletes' tenacity, courage, confidence and perseverance, so that they are in the face of setbacks and difficulties, perseverance, do not give up the belief in the pursuit of success.The factors of cultivating athletes' personality cultivation can make athletes have more sense of responsibility and responsibility in daily life, and face various challenges in life more happily, confidently and openly. Therefore, sports training is not only to improve the sports level of the athletes, but also to cultivate the personality cultivation of the athletes, shape the healthy personality of the athletes, stimulate the athletes' sense of social responsibility and patriotism, and improve the spiritual civilization quality of the athletes. These factors can help the athletes to better integrate into the society and contribute to the development of the society.

Sports training can not only improve the physical quality and competitive ability of athletes, but also improve the income level of athletes, which is also the spiritual value and significance of sports humanities. Sports training can provide a platform for athletes to show off their competitive level, and participating in various competitions can allow athletes to receive prize money and awards. Athletes can get more income through their own efforts and performance, which can be used to create more wealth value and improve the living conditions of athletes. At the same time, successful athletes can get more business opportunities, such as signing sponsors, endorsing promotion products, so as to earn more income. The actual effect of this aspect can improve the income level of the athletes and make a greater contribution to the athletes' family and society. In addition, this can also promote social development and promote economic growth.Successful athletes can guide more people to participate in sports activities through the personal image and influence of the athletes, and increase the development and potential of the sports industry. At the same time, the beautiful image and sportsmanship of athletes can also bring more positive energy and positive influence to the society. In sports training, improving the income level of athletes can not only help athletes improve their living conditions, but also bring more positive social impact and promote the development of sports industry and social progress.

Sports training is not only to improve the competitive level of athletes, but also to establish a healthy training environment, to reflect and penetrate the value and significance of sports humanistic spirit.The establishment of a healthy training environment can promote the physical and mental health of athletes, and improve their comprehensive quality and competitive ability. A healthy training environment should include perfect training facilities and equipment, scientific training plans and methods, and reasonable diet and rest arrange-

ments. At the same time, the mutual trust and respect between trainers and athletes is also an important part of the healthy training environment. Establishing a healthy training environment can cultivate athletes' good qualities and behavior habits, such as respecting rules, observing discipline, overcoming setbacks, teamwork, professionalism, etc. These qualities and habits can not only be improved in sports training, but also be applied in daily life. This quality and habit can not only enable athletes to achieve better results and performance in their competitive career, but also make them become better citizens and make more contributions to the society.In addition, a healthy training environment can also enhance the athletes 'confidence, form a healthy interpersonal relationship, and improve the athletes' mental state and cultural quality. These improvements can not only promote individual growth and development, but also make more contributions to social development and progress. Establishing a healthy training environment can not only improve the competitive level and physical quality of athletes, but also promote the improvement of comprehensive quality, and form a healthy personality. This is also one of the values and significance of the sports humanistic spirit.

Sports training is not only to improve the sports quality and competitive ability of athletes, but also to inherit and carry forward the sports culture, and to reflect and permeate the value and significance of the sports humanistic spirit. Inheriting sports culture can enable athletes and the public to better understand and feel the cultural heritage of sports, and better understand the significance and value of sports. Through inheriting and carrying forward sports culture, more people can understand the spirit and value of sports, form a healthy lifestyle and a positive attitude towards life, and play a more active role in social life. Inheriting sports culture can also promote the interaction and integration between athletes and spectators, enhance mutual understanding, and improve sports cohesion. At the same time, it can also better promote and develop various sports, spread sports culture knowledge, and promote the development of collective and individual.Therefore, inheriting sports culture is a necessary way to improve the sports training system and improve the national quality. Inheriting sports culture can also carry forward and inherit the national cultural tradition and enhance national identity and pride. Integrating national cultural elements in sports and combining traditional culture with modern science and technology can promote the innovation and inheritance of traditional culture, and promote the transformation and development of cultural tradition to modernity. Inheriting sports culture in sports training not only helps to shape the cultural temperament and spiritual pursuit of athletes, but also can promote the development of cultural inheritance and enhance the national spirit. This is also one of the values and significance of the sports humanistic spirit.

Exercise can not only exercise the body, but also contribute to the balanced development of physical and mental health. Coaches should encourage athletes to pay attention to their physical and mental health, and guide athletes to achieve their inner balance and health through sports. Competition is only one aspect of sports, and sports itself can bring athletes a lot of competition, such as teamwork, self-monitoring, discipline and so on. The coach should emphasize the value outside of these games, so that the athletes realize that as an athlete, whether in training or competition, should have these spiritual characteristics.In sports, success is not always fixed, and in many cases, the outcome of a competition can be influenced by various factors. Coaches can guide athletes to better analyze, evaluate and summarize various problems in sports training through self-thinking, which can help athletes expand their vision, enhance their critical thinking ability, and thus enrich their spiritual life. Humanistic spirit can enrich the spiritual life of athletes in sports training, so that athletes can get inner satisfaction while exercising[4]. In addition, the current authentication system also faces the following problems: fewer animation instances in operation, overflow; There are a lot of redundant frames in the video that can be easily attached to the pattern[5]; The network structure is unable to extract the essential features, which affects the improvement of the recognition rate. The measures taken by the author to solve the above problems include: integrating data augmentation algorithms in data preprocessing to reduce overfitting caused by small sample size; and Reduce the impact of nonvolatile data by filtering low-level video data; By integrating the color module into the remaining network, the extraction of discrimination is improved [6,7].

2. Methods. In the process of recognizing video images, it is necessary to process them first, and then classify them. If each frame is processed according to the input data, the calculation result of this algorithm can be greatly improved. On this basis, this paper designs a face recognition method based on R3D network, which extracts 16 frames from each video and weights them. Finally, a soft classifier is used to classify behaviors. On this basis, a new method based on data processing, feature extraction and behavior classification is proposed,as

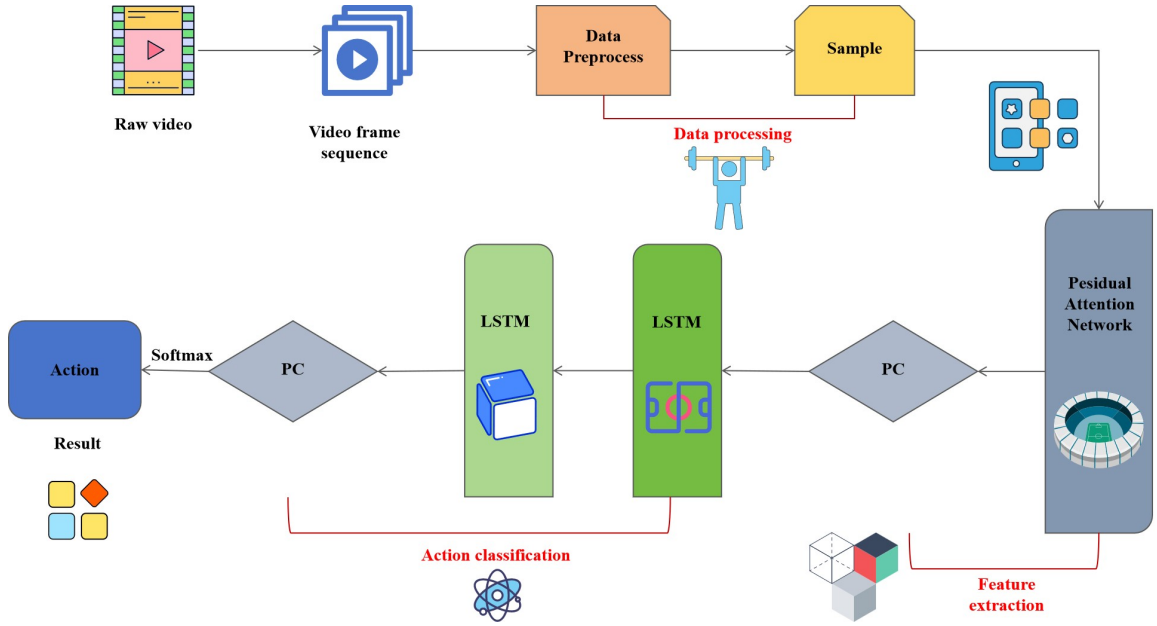

Fig. 2.1: Overall framework of the model

shown in Figure 2.1 [8].

2.1. Data processing.

(1) Data preprocessing. The traditional data pre-processing method is as follows: firstly, the fmeg module is used to sort the video into the system of video frame, and secondly, the first video frame is equal to the training requirement; and secondly, the first video frame is equal to the training requirement; Step three, focus on the cultivation and evaluation of video frames; Step four, convert the cropped video frame into a tensor form; Step five, tensor regular.

The algorithm has strong robustness, but it also faces two problems: first, the average confidence of the image in the image will cause the loss of the information boundary in the image; Secondly, the small number of action recognition samples of large samples leads to the phenomenon of overfitting during training.Therefore, in order to alleviate the above problems, the author proposes a data augmentation algorithm for videos (hereinafter referred to as Algorithm 1), in which, using a video frame sequence $\{f_1, f_2, ..., f_n\}$ to represent each action video V facilitates the indirect processing of video data using image processing methods [9]. If the video source has 50 frames and the output code range is (-6.6), after the data augmentation processing, the data can be expanded to 600 times. Meanwhile, the horizontal interpolation of video frames can also solve the problem of data edge loss by the average confidence. Therefore, the author incorporates data augmentation algorithms into data preprocessing, and the improved process of data preprocessing is as follows: firstly, data augmentation; and Step 2, scale; Step three, cut; Step four, alternating tensors; Step 5, regularly.

(2) Video frame sampling. The video frame structure mentioned in the human character recognition algorithm of R3D network is taken to measure the overall timing of the video frame system [10,11].

The specification is automatically generated with an L code of (0, R-16), where R is a video length parsed into a video frame; and Starting from frame L, select 16 frames of the sequence image as input to the model. Although this analytical method solves the computational cost problem caused by input in the network structure, it does not take into account the problem of data discrepancy in all the video frame sequences.If the initial output of the frame is in the low time frame data of the whole video, the input data obtained from the modeling process of R3D network human behavior recognition algorithm will be related to the model. Based on this, this paper proposes a human motion recognition method based on R3D network (hereinafter referred to as

"Algorithm 2"). In the algorithm 2, if the video frame rate is very low (n< 48), by using the index to ignore the effect of the frame parameter, the h number in the range of (0,n-16) is automatically generated, and 16 frames are selected from the h frame of the frame; In the case of high video frame rates, duplicate frames are removed in the start and end times and a random number h is generated in (n/3∼16,2 n/3∼16). Next, an image is selected from the 16th frame in the h frame of the text.

2.2. Feature extraction. Monitoring technology is to use neural networks to extract the information in the focus area, and to restrict other irrelevant information [12]. the Convolutional Attention Module (CBAM), as a lightweight module, has a rating of only 2.53×106, occupying very little expense. Therefore, in the special section, a subset of integrated circuit of CBAM is proposed.

(1) The basic structure of CBAM. CBAM includes two aspects: channel listening and spatial listening. Channel monitoring is an effective decision-making method, which assigns a higher weight to the channel with a large amount of information and uses it to study the source of the main message[13]. On this basis, this project proposes a method based on the maximum pool length and maximum pool length to compress the feature map F, and on this basis, the two feature fragments are introduced into the multi-layer perceptron (MLP), so as to reduce the coding and improve the performance. Finally, the MLP output signal is processed, and the channel density weight coefficient MC is obtained using the S-type function:

$$\begin{aligned} M_C =& \sigma(MLP(AvgPool(F))) + MLP(MaxPool(F)) \\ =& \sigma(W_1(W_0(F_{avg}^C)) + W_1(W_0(F_{max}^C))) \end{aligned} \tag{2.1}$$

CBAM multiplies the channel density weight factor MC by the input F, resulting in a new F'. Then, F 'is input into the airspace to obtain the spatial variation of the weight coefficient MS [14]. This paper presents a new algorithm. Finally, the F operation is performed on MS to obtain the final auditory feature F represented by formulas 2.2 and 2.3:

$$F' = M_C \bigotimes F \tag{2.2}$$

$$F" = M_S \bigotimes F' \tag{2.3}$$

(2) Improvement of CBAM. During training, each node in the network will dynamically adjust according to different inputs, which has a great impact on the subsequent characteristics. During collaborative learning, if the MLP weighs between the two groups of characteristics identical to the training, there will be a preceding problem. In order to solve this problem, improvements have been made to the channel monitoring system of CBAM. Firstly, concatenate and fuse the features after average pooling and maximum pooling, and then train the weights W_0' and W_1' through MLP, as shown in equation 2.4:

$$\begin{aligned} M_C =& \sigma(MLP([MaxPool(F); AvgPool(F)])) \\ =& \sigma(W_1'(W_0'([F_{max}^C; F_{avg}^C]))) \end{aligned} \tag{2.4}$$

In the formula, [MaxPool (F); AvgPool (F)] is the fused feature after concatenation. MLP consists of two FC layers, each with corresponding weights W_0' and W_1' [15]. After improving the channel attention module of CBAM, the parameter quantity of the weight W_0' obtained through training in the first FC layer of MLP is greater than the W0 before improvement. More, the performance of the model is stronger. Furthermore, although the improved W_1' has the same number of parameters as the pre improved W1, However,the development utilization of FC second layer in MLP can account for the maximum reservoirs and the average reservoirs,better the correlation between these two. For the sake of simplicity of description,an improved CBAM called C-CBAM,with only 2.99 units$\times 10^4$ buffer.

(3) Residual module. In the domain model, the network part uses the ResNet50 pattern and has sixteen sections.

To the right of the dotted box is a shorter connection that translates the input x directly into the output position. If the difference between x and F (x) is greater than 1. x changes in size; The difference in size. This

virtual block diagram represents the rest of the content, and it contains three levels, using 1 by 1. This project intends to use convolutional neural network to reduce the channel of input tensor, use 3x3 small convolution kernel to reduce the operational complexity, and then use 1xA convolution kernel and tensor channel to obtain F (x) output. Thus, the output value of the module as a whole is:

$$H(x) = F(x) + x \tag{2.5}$$

After F (x)=0, H (x)= x, this is the graph. Therefore, this project takes the residual probability F (x) as 0 to study the network part. In addition, it can be seen from formula (5) that the gradient loss and network problems caused by rules can be effectively avoided when the error reverse occurs in the residual network. Put 1 at x.

(4) Residual module integrated with G-CBAM. Put the rest of the pieces into GCBAM. Firstly, G-CBAM model is used to optimize the input features, and the optimal features are obtained. On this basis, according to the main information obtained, the depth feature of the original model is extracted. Finally, on this basis, the obtained results are further supplemented and integrated.

2.3. Action classification. Repetitive neural networks (RNNs) can handle timing problems well, but there are problems such as gradient loss when dealing with large-scale data[16]. To solve this problem, based on the specific structure of recurrent neural network, short-term memory model has the best performance in processing remote data. It inputs and outputs from the input gate, the forget gate, and the input gate.Among them, the input gate is located in the middle of the figure The layer, tanh layer, and a point by point multiplication "$\bigotimes$" determine how much input xt at the current time needs to be saved in the current unit state c_t; The recursive formula for updating LSTM is as follows:

$$f_t = \sigma(W_f h_{t-1} + U_f x_t + b_f) \tag{2.6}$$

$$i_t = \sigma(W_i h_{t-1} + U_i x_t + b_i) \tag{2.7}$$

$$\overline{c_t} = tanh(W_c h_{t-1} + U_c x_t + b_c) \tag{2.8}$$

$$c_t = f_t c_{t-1} + i_t \widetilde{c_t} \tag{2.9}$$

$$o_t = \sigma(W_o h_{t-1} + U_o x_t + b_o) \tag{2.10}$$

$$h_t = o_t \cdot tanh(c_t) \tag{2.11}$$

In the formula: W_f, W_i, W_c, W_o and U_f, U_i, U_c, U_o are the corresponding weight matrices, b_f, b_i, b_c, b_o is the corresponding bias, σ and tanh is the activation function.

3. Results and Analysis.

3.1. Experimental Environment. The author's test run environment is as follows: operating Ubuntu 16.04; Deep pytirch training 1.6.0%; Universal parallel computing architecture cuda10.2%; Deep neural network GPU fast cudnn7.6.5 library; GeForceRTX2080Ti graphics card with 11GB of memory; Picture card driving system nvidia450.80; 512GB Hard Disk [17].

3.2. Datasets. UCFYoutube.com has a total of 1,600 videos in 11 categories: shooting, swinging, swinging, cycling, horse riding, dog walking, diving, tennis, trampoline diving, volleyball. There are 25 different videos in each category, each with no less than four 320x240 pixel video clips. The KTH file contains 600 videos with a resolution of 160 px and 120 px. This case consists of 25 people who do 6 things in 4 different situations, including walking, jumping, speed, clapping, and swinging volleyball. HMDB51 materials consist of 6849 films, divided into 51 processing groups, with at least 101 films in each category and resolution of 320 pixels, × 240

Table 3.1: Experimental Parameters

parameter	numerical value
learming_rate	1x10-4
batch_size	30
epoch	100
worker	8
dropout	0.5
hidden_size	512
loss_function	crossentropy
optimizer	Adam

pixels. According to the category of actions, they can be divided into 5 types: facial acts, such as smiling and chewing; and A facial movement associated with surgery, such as smoking or eating; Physical exercise, such as shaking and walking; Interaction between bodies and objects, such as combing hair, dribbling balls, and drawing swords; Interactions between people, such as hugging and kissing. In this study, UCFYouTube and HMDB51 standards were used as 60 training topics, 20 were verified and 20 were tested. In terms of KTH data, because the sample number is relatively small, five different hybrid combinations are used in this method, 80% for training and 20% for testing.

3.3. Experimental Details. First, for UCFYouTube and HMDB51 data, their resolution is both 320 pixels×240 pixels, using it directly can increase the memory throughput due to excessive computation, need to be expanded, however, the KTH format is only 160x120 pixels, so the model can be imported directly. Secondly, in view of the high requirement of GPU computing power for video frequency recognition, this project intends to use transfer learning method to extract features from the model to improve training efficiency. That is, ResNet50, which retrains the web, has been ported to the ResNet standard adopted by the authors.Finally, in order to minimize the risk of network overfitting, each FC layer adopts Dropout technology to minimize the interference of nodes in FC layer with certain probability. The experimental results of the author are shown in Table 3.1.

3.4. Experimental process.

(1) The impact of attention modules on model performance. This project intends to make use of the accuracy and loss characteristics of YouTube (ResNetNetT), RLNetRLCBAM and RLNetG-CBAM modes (RLNetG-CBAM) (see Figure 3.1 and 3.2). All three models have large deviations in the initial learning stage, and tend to stabilize with the increase of the number of iterations. Compared with the traditional content-based network model, the combination of RLNet and content-based network model can effectively improve the classification effect of images, but the accuracy and loss rate of images will be greatly changed during training. The experimental results show that the combination of RLNet and G-CBAM has higher recognition rate and lower fall rate. The method has good stability because of its fast convergence speed, fast convergence speed and fast convergence speed. On this basis, a new method based on collaborative learning is proposed in this project. Through the correlation analysis of multiple elements, it can better deal with the correlation between multiple elements, reduce the impact of errors, reduce their importance, and enhance the stability and credibility of the model.

(2) Verification of the effectiveness of improvement measures. This project intends to use RLNet, RLNet1, RLnet2, 2 CBAM, 2G-CBAM and other models in UCFYouTube database to verify the above methods. The test results are shown in Table 3.2, where RLNet1 is the abbreviation of RLNetAlgol, and RLNet1 and 2 are the abbreviation of RLNetAlgolAlgo2 [18].

According to Table 3.2, various improvement measures have improved the recognition performance of the model by 1.57%, 1.17%, 1.89%, and 1.28%, respectively.

(3) Visualization of feature regions. Find out the implementation of the finite element method in the special section using Grade CAM technique. It can be clearly seen that the network extended with CBAM can not only locate the area where the main features reside, but also restrict other inefficient information.

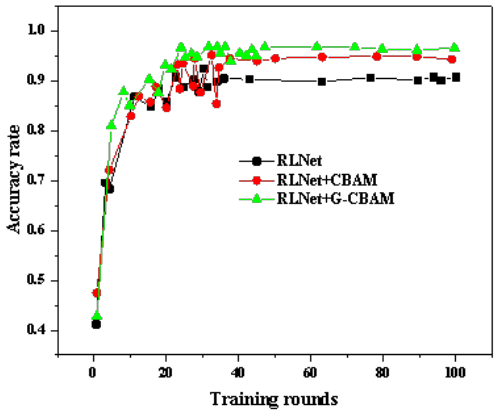

Fig. 3.1: Accuracy Curve

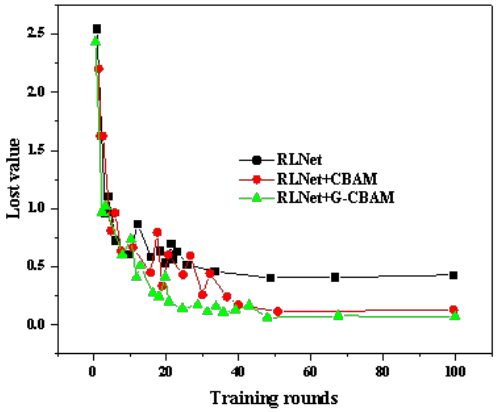

Fig. 3.2: Loss value curve

Table 3.2: Comparison of Various Improvement Measures

Ablation model	Accuracy/%
RLNet	90.86
$RLNet_1$	92.42
$RLNet_{1,2}$	93.58
$RLNet_{1,2}$+CBAM	95.46
$RLNet_{1,2}$+C-CBAM	96.73

Meanwhile, compared with CBAM, the improved G-CBAM has more complete and accurate localization of the main features, improving the network's ability to learn discrimination.

Table 3.3: Comparison with other methods on UCF YouTube

Method	Accuracy/%
Deep-Temporal LSTM	90.28
Proposed DB LSTM	92.85
Inceptionv3 + BiHSTM -Attention	94.39
$RLNet_{1,2}$ +G-CBAM(This article)	96.73

Table 3.4: Cross validation on the KTH dataset

Method	Cross validation results/%					
	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Average value
RLNet	96.33	85.51	96.85	86.68	88.52	90.78
$RLNet_1$	97.18	87.26	97.02	93.37	94.69	93.9
$RLNet_{1,2}$	98.51	89.18	100.00	94.63	95.18	95.5
$RLNet_{1,2}$ +CBAM	99.16	92.72	100.00	95.81	96.51	96.84
$RLNet_{1,2}$ +C-CBAM	100.00	94.57	100.00	97.46	98.28	98.07

Table 3.5: Comparison with other methods on UCF YouTube

Method	Accuracy/%
Deep-Temporal LSTM	93.91
CNNLSTM	94.21
Inceptionv3 + Bi-HSTM -Attention	95.68
$RLNet_{1,2}$ +G-CBAM(This article)	98.07

3.5. Experimental Results. In order to fully validate the method proposed in this article, experiments were conducted on three datasets: UCFYouTube, KTH, and HMDB51.

(1) Validation on UCFYouTube. After the training sample has been completed, the recognition rate of this sample on UCFYouTube data has reached 96.73%, which achieves better recognition than the existing recognition method, as shown in Table 3.3.

(2) Validation on KTH. The characteristics of KTH are that there are less information related to background, no interaction behavior, and the information is relatively simple. During the training process, the first 120 films in KTH are selected as the test samples, and the remainder is used as the training process, called the data, and then the first cross is validated. Similarly, a total of 5 cross-sectional trials were validated, and an average cost was taken as an experiment, as shown in Table 3.4 [19]. The The experimental results show that the recognition rate of RLNet1,2G-CBAM model of KTH reaches 98.06%.

Furthermore, as shown in Table 3.5, this method still has better recognition performance compared to other methods on the KTH dataset.

(3) Validation on HMDB51. 5151 chip data is mainly video, the data distribution is wide, the teaching content is complex. This project will use HMDB51 data collected by HMDB51 to test the recognition effect of RLNet1,2 GCBAM model in complex environment by comparing it with other algorithms. The results are shown in Table 3.6.

From Table 3.6, it can be seen that this method has some improvement compared to other motion recognition systems of HMDB51, but there is a significant difference in recognition accuracy compared to UCFYouTube and KTH [20]. It is shown that the method has a certain improvement in recognition accuracy. The main reason is that HMDB51 has higher video quality than other dual data rate, and has many disadvantages, such as camera motion, occlusion, background complexity, and changes in lighting quality, which leads to lower

Table 3.6: Comparison with other methods on HMDB51

Method	Accuracy/%
C3D	51.61
R3D	62.31
iDT+Video LSTM	63.01
$RLNet_{1,2}$ +G-CBAM(This article)	64.82

recognition rate.

4. Conclusion. The author proposes an in- depth learning strategy that integrates human resource management process. This method reduces the risk of model overfitting by integrating data augmentation algorithms into data preprocessing. By filtering low data rate video frames, the influence of redundant data is reduced. By integrating the light weight of G-CBAM into the residual network, an improved performance is achieved with a small number of parameters. Ultimately, recognition rates of 96.73%, 98.07%, and 64.82% were achieved on the UCFYouTube, KTH, and HMDB-51 datasets, respectively. Through the simulation of HMDB51 data, it is proved that the accuracy of this algorithm in complex cases is very low. So, the next step is to focus on how to enhance the credibility of the model in many negative ways.

REFERENCES

[1] Surasak, T., & Kitchat, K. (2022). Application of deep learning on student attendance checking in virtual classroom. 2022 4th International Conference on Electrical, Control and Instrumentation Engineering8841 (ICECIE), 1-4.
[2] Liu, T., Zheng, Q., & Tian, L. (2022). Application of distributed probability model in sports based on deep learning: deep belief network (dl-dbn) algorithm for human behaviour analysis. Computational intelligence and neuroscience, 2022(752), 7988844.
[3] Hui, Z., Jing, C., & Taining, W. (2022). Research on simulation analysis of physical training based on deep learning algorithm. Scientific Programming, 2022(7551), 1-11.
[4] Wang, J., & Zhao, R. (2022). Deep reinforcement learning and application in self-driving. Theories and Practices of Self-Driving Vehicles, 475(42),307-326.
[5] Donghwan, Y., Hyun-Lim, Y., Soonil, K., So-Ryoung, L., Kyungju, K., & Kwangsoo, K., et al. (2023). Automatic segmentation of atrial fibrillation and flutter in single-lead electrocardiograms by self-supervised learning and transformer architecture. Journal of the American Medical Informatics Association3245(1), 1.
[6] Lee, C. (2022). Deep learning-based detection of tax frauds: an application to property acquisition tax. Data technologies and applications85(3), 56.
[7] Zhou, Y. (2022). The application of curriculum ideology and politics in the training of judicial vocational education talents. Journal of Higher Education Research, 3(2), 155-159.
[8] Yang, Z. Q., Wei, M. F., Chen, L., & Xi, J. N. (2023). Research progress in the application of motor-cognitive dual-task training in rehabilitation of walking function in stroke patients. Journal of Neurorestoratology, 11(1), 100028-100028.
[9] Westover, J. (2022). Effective human resource training and development: examination and application of adult learning theory in the hr management context.845(52),821
[10] Li, S., & Bai, Y. (2022). Deep learning and improved hmm training algorithm and its analysis in facial expression recognition of sports athletes. Computational intelligence and neuroscience, 2022(721), 1027735.
[11] Cha, S., Jo, H. W., Kim, M., Song, C., Lee, H., & Park, E., et al. (2022). Application of deep learning algorithm for estimating stand volume in south korea. Journal of Applied Remote Sensing.85(554),792
[12] Zhao, X., Huang, W., Hao, W., & Liu, Y. (2023). A training strategy to improve the generalization capability of deep learning-based significant wave height prediction models in offshore china. Ocean engineering07(Sep.1), 283.
[13] Wang, Y., Chen, S., Hong, Y., Hu, B., Peng, J., & Shi, Z. (2023). A comparison of multiple deep learning methods for predicting soil organic carbon in southern xinjiang, china. Computers and Electronics in Agriculture, 212(75),087.
[14] Jush, F. K., Biele, M., Dueppenbecker, P. M., & Maier, A. (2022). Deep learning for ultrasound speed-of-sound reconstruction: impacts of training data diversity on stability and robustness.42(51),090
[15] Guo, J., Cao, W., Nie, B., & Qin, Q. (2023). Unsupervised learning composite network to reduce training cost of deep learning model for colorectal cancer diagnosis. IEEE journal of translational engineering in health and medicine, 11(572), 54-59.
[16] Rathore, M., & Singh, P. N. (2022). Application of deep learning to improve the accuracy of soil nutrient classification. 2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon), 1-5(563),587.
[17] K.-J., C., C.-H., D., T.-C., C., J.-J., X., & M.-T., L. (2023). Application of recurrence plots and vgg deep learning model to the study of condition monitoring of robotic grinding. International Journal of Precision Engineering and Manufacturing897(9), 24.

[18] Abdollahifard, S., Farrokhi, A., & Mowla, A. (2023). Application of deep learning models for detection of subdural hematoma: a systematic review and meta-analysis. Journal of neurointerventional surgery924(10), 15.
[19] Joel, M. Z., Umrao, S., Chang, E., Choi, R., Yang, D. X., & Duncan, J. S., et al. (2022). Using adversarial images to assess the robustness of deep learning models trained on diagnostic images in oncology. JCO clinical cancer informatics, 6(568), e2100170.
[20] Bartoli, A., Fournel, J., Maurin, A., Marchi, B., Habert, P., & Castelli, M., et al. (2022). Value and prognostic impact of a deep learning segmentation model of covid-19 lung lesions on low-dose chest ct. Research in Diagnostic and Interventional Imaging, 1(12), 100003 - 100003.

Edited by: Hailong Li
Special issue on: Deep Learning in Healthcare
Received: Jan 29, 2024
Accepted: Mar 8, 2024