Hindawi
Mobile Information Systems
Volume 2022, Article ID 4778245, 7 pages
https://doi.org/10.1155/2022/4778245

Research Article

Sequence Video and Artificial Intelligence Assisted Basketball Injury Risk Early Warning Method

Gui Yang and Xiao Xu

Beijing Sport University, Beijing 100084, China

Correspondence should be addressed to Xiao Xu; 1004320180066@bsu.edu.cn

Received 16 June 2022; Revised 25 July 2022; Accepted 29 July 2022; Published 5 September 2022

Academic Editor: Tao Dai

Basketball sport is a comprehensive nonperiodic collective sport, in which sports injuries of various parts are also prone to occur. Its higher exercise intensity not only effectively enhances the physical fitness of the athlete but also causes physical damage due to a series of actions such as frequent take-offs and landings, resulting in certain sports injuries. Therefore, the early warning of injury in basketball is very necessary, and it is very useful to protect the safety of players. In this paper, we investigate how to use the sequence video and artificial intelligence method for the basketball injury risk early warning in order to protect the player and better assist the player to improve their efficiency of training. First, we preprocess the video sequence. We convert the video sequence into image data and perform noise reduction and transformation operations on the image. Second, for the processed image data, we designed a convolutional neural network model to determine the damaged area. Third, we use the neural network model to take the image data with the detection area as the input, perform feature extraction on the data, and finally obtain the early warning value of basketball sports injury. The experimental results prove that the method proposed in this paper has good evaluation performance.

1. Introduction

Basketball is a comprehensive nonperiodic collective sport, which is determined by the diversity of its sports content structure and the variability and comprehensive characteristics of the competition process [1–3]. Engaging in basketball competitions and various basketball activities will help to cultivate the overall quality of the players, improve physical health, activate body and mind, and increase knowledge. It has a positive impact on exercising comprehensive talents and developing people's wisdom, cultivating excellent moral character and tenacious will. For example, the practical operation and actual application of basketball skills and tactics systems are completed by running, jumping, shooting, and other means under the requirements of a specific time, distance, venue, and facility conditions that are changing. In this process, both intelligence, physiology, and psychology must bear the comprehensive influence of various complex factors. Participating in basketball activities in an appropriate amount has a positive effect on promoting people's physiological functions, improving health, and comprehensively developing physical fitness and psychological cultivation.

However, in basketball, players have to carry out intense close attack, defense, and confrontation. Under changing and complex conditions, it is necessary to perform technical actions with high difficulty such as shooting, layup, dribbling breakthrough, and stealing, and the exercise intensity is large and the energy consumption is large. Therefore, sports injuries of various parts are also prone to occur [4–6]. Its higher exercise intensity not only effectively enhances the physical fitness of the athlete but also causes physical damage due to a series of actions such as frequent take-offs and landings, resulting in certain sports injuries.

According to the American Orthopaedic Sports Medicine (AOSSM) Sports Trauma and Overuse Prevention (STOP) Program [7, 8], the most common basketball injuries (both amateur and professional) are ankle sprains, finger contusions, knee injuries, thigh contusions, and facial injuries. Basketball injuries can cause unpredictable injuries to players.

The rapid development of informatization has brought technical convenience to basketball injury research. Through the data analysis of basketball injuries, it is possible to accurately understand the distribution of injuries among players of different genders in basketball games, as well as the distribution of injuries. This plays an important and fundamental role in taking targeted injury prevention methods and scientifically formulating recovery training programs. Data collection and analysis on the distribution of injuries in basketball shows that it is not difficult to find that physical injuries caused by basketball are a common phenomenon. Because the main body of basketball is mostly boys, and boys are more competitive. Therefore, among people who have experienced sports injuries, the proportion of boys is much higher than that of girls. Although the vast majority of sports injuries have a low number of injuries, they have a greater impact on the body. Therefore, when engaging in basketball, it is still necessary to pay attention and take preventive measures against sports injuries.

In the context of the rapid development of science and technology [9, 10], in order to effectively improve the professionalism and pertinence of injury prevention and recovery, it is necessary to comprehensively analyze the injury data information based on basketball sports, so as to understand the common injury types, identify the location of the injury, and do a good job of the cause of the injury analysis to minimize damage. According to the situation of sports injuries, combined with the analysis of basketball movements, it can be seen that the common injuries mainly occur in the ankle joint, knee joint, finger joint, and waist. The reasons are mainly as follows: first, before the exercise, the preparation activities were not implemented properly. Second, in the process of exercise, professional technical movements are not mastered in place. The third is the lack of physical fitness and function, and the overload of exercise intensity causes strain on parts of the body. Among the common types of injuries, most of them are contusions, abrasions, sprains, strains, and other minor physical injuries. Although they are less harmful to the body, they occur at a high rate. In addition, there are serious injuries to the body, such as fractures, which need special attention and take effective preventive measures to prevent serious injuries to the body.

Artificial intelligence has surpassed the human level in large-scale image recognition and face recognition [11–13]. Dedicated artificial intelligence systems for specific tasks (such as playing Go) can surpass human intelligence in single-item tests of local intelligence levels due to single tasks, clear requirements, clear application boundaries, rich domain knowledge, and relatively simple modeling. Artificial intelligence (AI) is a new technical science that studies and develops theories, methods, technologies, and application systems for simulating, extending, and expanding human intelligence. Artificial intelligence is a branch of computer science that attempts to understand the essence of intelligence and produce a new type of intelligent machine that responds in a similar way to human intelligence. Research in this area includes robotics, language recognition, image recognition, natural language processing, and expert systems. Since the birth of artificial intelligence, the theory and technology have become more and more mature, and the application field has also continued to expand. The technological products brought by artificial intelligence in the future will be the "containers" of human intelligence. Artificial intelligence can simulate the information process of human consciousness and thinking. Artificial intelligence is not human intelligence, but it can think like human beings and may surpass human intelligence.

The development of artificial intelligence provides more intelligent and accurate methods and technologies for sports injury risk early warning, especially in the application of basketball sports injury risk early warning tasks. A healthy physical state is the basic requirement for basketball players to participate in sports, and it is also the fundamental guarantee to effectively avoid injury. Under the background of informatization, strengthening the informatization management of the health status of basketball players can conduct a comprehensive and systematic health analysis of basketball players in a targeted manner and promote the continuous improvement of their physical health. After the basketball player completes the training task, the actual physical health status of the athlete can be recorded in detail, and the measurement and analysis of the health level and various indicators of the body can be done by means of information technology. Once a basketball player's body is abnormal or uncomfortable, a preliminary judgment can be made on the player's physical symptoms according to the analysis data of their health status, so as to prevent the increase of physical exercise burden and increase the probability of physical injury during exercise. This can effectively protect the physical health of athletes participating in basketball training activities and reduce sports injuries.

In this paper, we investigate how to use the sequence video and artificial intelligence method for the basketball injury risk early warning in order to protect the player and better assist the player to improve their efficiency of training. At the same time, we use the convolution neural network model for the video translation, image process, and feature extraction, so as to provide higher performance for the basketball injury risk early warning [14, 15].

The contribution of this paper can be summarized as follows:

(1) We propose the basketball injury risk early warning method by adopting the neural network model to extract more useful features

(2) We design a video transform algorithm and image denoising method based on a convolutional neural network model, which coverts the video data to images and filter the noise of images

(3) We experimentally demonstrate the effectiveness and performance of the method proposed in this paper on different conditions.

The rest of this paper is organized as follows: Section 2 introduces the proposed method for basketball injury risk early warning, including video data preprocessing, Injury area determination, and Injury detection and early warning based on a neural network model. Section 3 presents experimental studies and results to compare and demonstrate the performance of the proposed model for basketball injury risk early warning evaluation. Section 4 concludes the whole paper.

2. Proposed Method

In this paper, we aim to investigate how to use the neural network model and sequenced videos to detect and warn basketball injuries to better protect players and improve training efficiency.

First, we preprocess the video sequence. We convert the video sequence into image data and perform noise reduction and transformation operations on the image. Second, for the processed image data, we designed a convolutional neural network model to determine the damage area. Third, we use the neural network model to take the image data with the detection area as the input, perform feature extraction on the data, and finally obtain the early warning value of basketball sports injury.

2.1. Video Data Preprocessing. We use video data for early warning of basketball sports injuries. First, we process the video data into image data. Video data processing is not just about understanding every image in the video, but more about identifying the few best key themes that describe the video. The research content of video data processing mainly includes video conversion and image data processing. Before a video can be calculated, we design a video conversion algorithm to convert video into pictures.

Videos are made up of a series of still images, usually, 24–30 images per second or more, with each individual image called a frame. The continuous playback of these still images can lead to the false impression that the object is moving, and the faster the images are played, the smoother the motion will appear. Video to picture can be captured frame by frame, or continuous images or frames can be extracted by some software. The video conversion algorithm is shown in Algorithm 1.

After the video processing is completed, we can get the image data corresponding to the video. Because the video capture contains noise interference. Therefore, we denoise the image data. The basic image noise reduction algorithm usually removes the noise in the image by low-pass filtering the image. This method removes noisy pixels, but at the cost of blurring the image-all edges are blurred. More advanced noise reduction algorithms detect edges in images and low-pass filter them along lines perpendicular to the edges, a method that preserves the edges. However, low-pass filtering is not as effective because it has fewer pixels on average.

We use Gaussian low-pass filtering to denoise the image [16, 17]. The process of filtering an image by discrete Fourier transform is shown in Figure 1.

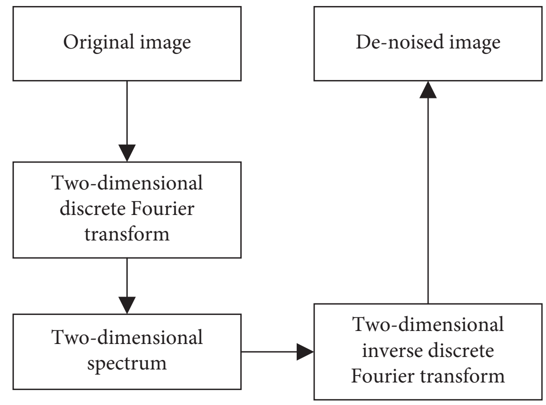

Figure 1: The process of filtering an image by discrete fourier transform.

The formula for the two-dimensional discrete Fourier transform is as follows:

$$F(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)e^{-j2\pi(ux/M+vy/N)}, \quad (1)$$

where x and y are the horizontal and vertical coordinates of the original image, respectively, M and N are the height and width of the original image, and f (x, y) is the gray value at the coordinates (x, y) on the original image. Euler's formula is shown as follows:

$$e^{-j2\pi(ux/M+vy/N)} = \cos\left(2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)\right) - j\sin\left(2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)\right). \quad (2)$$

Therefore, each pixel on the two-dimensional spectrum obtained by the two-dimensional discrete Fourier transform [18] of the original image is a complex number. When processed in the Mat data format in OpenCV [19], each pixel value has two channels. The formula for the two-dimensional inverse discrete Fourier transform is as follows:

$$f(x,y) = \frac{1}{MN}\sum_{u=0}^{M-1}\sum_{v=0}^{N-1} F(u,v)e^{j2\pi(ux/M+vy/N)}, \quad (3)$$

where the symbol representation is the same as the above-mentioned two-dimensional discrete Fourier transform.

2.2. Injury Area Determination. After video processing and image denoising, we design an injury region determination method based on the CNN model.

The deep learning models currently used in image recognition and analysis research mainly include Convolutional Neural Networks (CNNs) [20], Deep Belief Networks (DBNs) [21], Stacked Auto-Encoders (SAE) [22], and so on. The convolutional neural network is superior to the latter two methods in detection accuracy and speed. The convolutional neural network was proposed by biologists Huber and Wiesel in the early research on cat visual cortex, it is a variant of multilayer perceptron (MLP). LeCun of New York University proposed the convolutional neural network model LeNet [23] in 1998 and applied it to the detection of handwritten characters. In 2012, Professor Hinton led his student Krizhevsky to The convolutional neural network model AlexNet [24] was applied to the image classification task, and won the first place in the ImageNet [25] Large-scale Visual Recognition Challenge (ILSVRC), which reduced the Top-5 error rate to 15.3%. Currently, Microsoft's ResNet [26] and Google's InceptionV4 [27] reduce the Top-5 error rate to within 4%, surpassing human performance on this specific task. Girshick combined the region prediction + convolutional neural network method in 2014. (R-CNN) [28] was applied to the object detection task and achieved good performance.

CNN have the following advantages in object detection tasks.

(1) Multiple convolution kernels. Adding multiple convolution kernels can learn a variety of features to better describe the image.

(2) Local perception. Generally speaking, people's cognition of the outside world is from local to global, and the spatial relationship of the image is also closer to the adjacent pixel regions, and the correlation of the farther pixels is weaker. Therefore, each neuron first senses the local area and then synthesizes the local information at a higher level to obtain global information. The number of training parameters can also be reduced by local sensing.

(3) Parameter sharing. In the local connection, the parameters of the filter corresponding to each neuron are the same, which greatly reduces the number of training parameters.

(4) Pooling. After using multicore convolution, the extracted feature dimension is huge. In order to reduce the dimension, the maximum or average method is used to reduce the dimension. The pooled features can be maintained when scaling or changing.

(5) Sparsity restriction. The neuron data in the hidden layer is large, so adding a sparsity restriction to this makes most neurons in a suppressed state and only a few neurons are activated, which helps us to extract a small number of neurons. Extract the more essential features of the image. This greatly reduces the number of parameters of the network and speeds up the training speed.

The above advantages enable CNN to extract image features by itself when processing images, including color, texture, shape, and topological structure of the image, which has good robustness and computational efficiency.

In this section, we design a CNN model to mark the region feature of the player in the images. The main process contains two steps. First, we preprocess the images, including crop, data augmentation, resize, and transform. Second, the preprocessed images are input into the CNN model to generate the images with marked regions. The processing flow is shown in Figure 2.

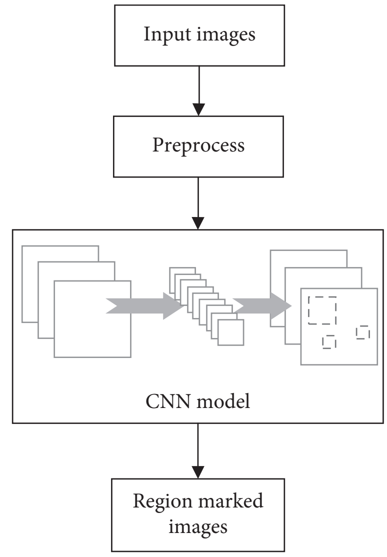

Figure 2: The processing flow of injury area determination based on CNN model.

2.3. Injury Detection and Early Warning. After obtaining the above images with marked regions, we propose an injury detection and early warning method based on a neural network model.

By constructing an adaptive neural network model to determine whether the image sports damage area is sports damage, the neural network belongs to the structure of multiple input and single output. The acquired sports injury signal feature U is used as the input vector of the neural network; the output result is b, which represents the corresponding sports injury type +1 is the value of the training sample, indicating that it is a sports injury; −1 is also the value of the training sample, indicating that it is not a sports injury. The neuron activation functions are all bisexual sigmoid functions, and the formula is as follows:

$$h(U) = \frac{1 - e^{-U}}{1 + e^{-U}}. \tag{4}$$

Suppose q is the number of hidden layer nodes of the output neural network model, hM is the activation function of the hidden layer nodes, θ_a $(a = 1, 2, \ldots, q)$ is the threshold; 1 is the number of output layer nodes, $h0$ is the hidden layer The activation function of the node, β is the threshold. Z_{xa} is the connection weight between the x-th hidden node and the ath input node, and Z_x is the connection weight between the x-th hidden node and the output node. Then, we can calculate the value of b as follows:

$$b = h_0 \left\{ \sum_{x=1}^{q} Z_x h_M \left[\sum_{a=1}^{10} Z_{xa} U_a + \theta_x \right] + \beta \right\}. \tag{5}$$

The sports injury is classified by the output b value. If $b \geq 0$, it belongs to sports injury, and if $b < 0$, it does not belong to sports injury.

The steps of building an adaptive neural network model are as follows:

```
Input: video data
Output: image data
(1) initialize the parameters;
(2) read the video data steam;
(3) get the frame rate;
(4) for each frame i = 1, 2, . . ., do
(5) get the name of image;
(6) get the save path;
(7) calculate image stream data;
(8) write the image into the path;
(9) end for
(10) Return the image data.
```

ALGORITHM 1: The video conversion algorithm.

(1) Construct a training sample set and a test sample set. When the initial external structural parameters of the adaptive neural network model are certain, the training sample set is learned through the BP (Back Propagation) algorithm of the neural network, and the internal connection weight parameters of the learning model are obtained.
(2) Through the test sample set, test the learning model, calculate the discrimination bias, and obtain the fitness function of the genetic algorithm.
(3) Use the learning mechanism of the genetic algorithm to adjust the external structural parameters of the adaptive neural network model, and based on the adjusted external structural parameters, obtain the internal connection weight parameters of the new adaptive neural network model through the neural network BP algorithm.
(4) Calculate the discrimination bias and fitness function in the same way and then continue to adjust the external parameters until it meets the stopping requirements of the genetic algorithm.
(5) Output the adaptive neural network model with the best generalization ability.

3. Experiments and Results

In this section, we experimentally verify the basketball injury early warning method based on the CNN model proposed in this paper.

In the experiment, we selected 50 basketball sports injury personnel as the experimental objects, and randomly divided the 50 sports injury personnel into 5 groups to test the feasibility and accuracy of the automatic detection of sports injury of this method. The neural network is trained by the training sample set, the neural network is tested by the test sample set, and the best adaptive neural network structure parameters are obtained according to the test results.

The performance metrics adopted in our experiments include Accuracy (Acc), Sensitivity (Sen), Specificity (Spe), and Positive Predictivity Value (PPV), which are common metrics used in the classification task. Their definitions are given as follows:

$$\begin{aligned}\text{Acc} &= \frac{\text{TP}+\text{TN}}{\text{TP}+\text{TN}+\text{FP}+\text{FN}},\\ \text{Sen} &= \frac{\text{TP}}{\text{TP}+\text{FN}},\\ \text{Spe} &= \frac{\text{TN}}{\text{TN}+\text{FP}},\\ \text{PPV} &= \frac{\text{TP}}{\text{TP}+\text{FP}},\end{aligned} \tag{6}$$

where TP is the number of true positive detections, TN is the number of true negative detections, FN represents the number of false negative detections, and FP represents the number of false positive detections. Each indicator represents the measurement of the classification error between the predicted value of electricity and the actual value of electricity. The more accurate the predicted value is, the bigger each indicator is, and the better the performance of the model is.

In order to verify the influence of the number of layers of the designed deep model on the performance of the model, we tested the models with different numbers of layers. Table 1 shows the performance of the network models with different numbers of layers. According to Table 1, with the increase of the hidden layer of the model, the training accuracy of the network first increases and then decreases. When the number of hidden layers of the model is 7, the maximum number of training steps is 17, and the highest accuracy is obtained. This shows that it is not that the number of hidden layers of the model is better, because the more hidden layers of the model, the easier it is to cause overfitting of the model, but it will reduce the performance of the model.

Furthermore, we compare the impact of two image preprocessing methods on model performance, including noise reduction processing and injury region detection. That is, we compare model performance with and without denoising and injury region detection. We list these conditions as follows:

(1) Without the denoising operation
(2) With the denoising operation

Table 1: The performance of the network models with different numbers of layers.

Layers	Training steps	Acc/(%)
3	20	75.4
5	20	78.7
7	17	90.4
9	15	86.6
11	15	84.2

Table 2: The comparison results of model performance under different conditions.

Conditions	Acc	Sen	Spe	PPV
(a)	82.5	83.9	82.3	80.6
(b)	85.5	82.4	82.7	84.5
(c)	84.6	83.1	84.2	83.8
(d)	88.7	86.3	85.3	87.9
(e)	81.4	80.1	78.3	80.5
(f)	90.4	88.5	89.7	90.1

(3) Without the damage region detection

(4) With damage region detection

(5) Without the denoising operation and the damage region detection

(6) With the denoising operation and the damage region detection

Table 2 shows the comparison results of model performance under different conditions. From Table 2, it can be seen that both denoising operation and injury region detection help to improve the model performance. At the same time, we can also see that the influence of noise interference on the model is not large, while the influence of region detection is relatively large.

The above-given experimental results show that the basketball injury early warning method based on the neural network model proposed in this paper has good performance.

4. Conclusions

In this paper, we propose a basketball injury early warning method based on a neural network model to protect the player and better assist the player to improve their efficiency of training. First, we preprocess the video sequence. We convert the video sequence into image data and perform noise reduction and transformation operations on the image. Second, for the processed image data, we designed a convolutional neural network model to determine the damaged area. Third, we use the neural network model to take the image data with the detection area as the input, perform feature extraction on the data, and finally obtain the early warning value of basketball sports injury.

However, as a novel method, this paper has also some limitations. On one hand, the time complexity of the algorithm has not been analyzed. On other hand, the scale of the dataset is not sufficient. In the future, we will improve the method around the mentioned two aspects. In addition, we also will consider to build one practical platform for the early injury risk warning.

Data Availability

The data used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest in this paper.

References

[1] F. Allard, S. Graham, and M. E. Paarsalu, “Perception in sport: basketball,” *Journal of Sport & Exercise Psychology*, vol. 2, no. 1, pp. 14–21, 1980.

[2] J. Sampaio, B. Gonçalves, and N. Mateus, “Basketball,” *Modelling and Simulation in Sport and Exercise*, pp. 108–126, Routledge, Oxfordshire, UK, 2018.

[3] B. Li and X. Xu, “Application of artificial intelligence in basketball sport,” *Journal of Education, Health and Sport*, vol. 11, no. 7, pp. 54–67, 2021.

[4] J. S. Newman and A. H. Newberg, “Basketball injuries,” *Radiologic Clinics of North America*, vol. 48, no. 6, pp. 1095–1111, 2010.

[5] Ş Dane, S. Can, R. Gürsoy, and N Ezirmik, “Sport injuries: relations to sex, sport, injured body region,” *Perceptual & Motor Skills*, vol. 98, no. 2, pp. 519–524, 2004.

[6] T. H. Trojian, A. Cracco, M. Hall, M. Mascaro, G. Aerni, and R. Ragle, “Basketball injuries: caring for a basketball team,” *Current Sports Medicine Reports*, vol. 12, no. 5, pp. 321–328, 2013.

[7] B. S. Delay, R. J. Smolinski, W. M. Wind, and D. S. Bowman, “Current practices and opinions in ACL reconstruction and rehabilitation: results of a survey of the American Orthopaedic Society for Sports Medicine,” *The American Journal of Knee Surgery*, vol. 14, no. 2, pp. 85–91, 2001.

[8] D. E. Hill and J. R. Andrews, “Stopping sports injuries in young athletes,” *Clinics in Sports Medicine*, vol. 30, no. 4, pp. 841–849, 2011.

[9] P. Costamagna and S. Srinivasan, “Quantum jumps in the PEMFC science and technology from the 1960s to the year 2000: Part II. Engineering, technology development and application aspects,” *Journal of Power Sources*, vol. 102, no. 1-2, pp. 253–269, 2001.

[10] E. J. Lavernia and T. S. Srivatsan, “The rapid solidification processing of materials: science, principles, technology, advances, and applications,” *Journal of Materials Science*, vol. 45, no. 2, pp. 287–325, 2010.

[11] A. N. Ramesh, C. Kambhampati, J. R. T. Monson, and P. Drew, “Artificial intelligence in medicine,” *Annals of the Royal College of Surgeons of England*, vol. 86, no. 5, pp. 334–338, 2004.

[12] R. Mitchell, J. Michalski, and T. Carbonell, *An Artificial Intelligence Approach*, Springer, Berlin, Germany, 2013.

[13] J. He, S. L. Baxter, J. Xu, J. Xu, X. Zhou, and K. Zhang, “The practical implementation of artificial intelligence technologies in medicine,” *Nature Medicine*, vol. 25, no. 1, pp. 30–36, 2019.

[14] B. B. Traore, B. Kamsu-Foguem, and F. Tangara, "Deep convolution neural network for image recognition," *Ecological Informatics*, vol. 48, pp. 257–268, 2018.

[15] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Proceedings of the 2017 International Conference on Engineering and Technology (ICET)*, pp. 1–6, IEEE, Antalya, Turkey, August 2017.

[16] J. Lv, X. Wang, and M. Huang, "ACO-inspired ICN Routing mechanism with Mobility support," *Applied Soft Computing*, vol. 58, pp. 427–440, 2017.

[17] F. Qu, D. Ren, X. Liu, J Zhenyu, and Y Lin, "A face image illumination quality evaluation method based on Gaussian low-pass filter,"vol. 1, pp. 176–180, in *Proceedings of the 2012 IEEE 2nd International Conference on Cloud Computing and Intelligence Systems*, vol. 1, pp. 176–180, IEEE, Hangzhou, China, November 2012.

[18] H. J. Nussbaumer, "The fast Fourier transform," *Fast Fourier Transform and Convolution Algorithms*, pp. 80–111, Springer, Berlin, China, 1981.

[19] G. Bradski, "The openCV library," *Dr. Dobb's Journal: Software Tools for the Professional Programmer*, vol. 25, no. 11, pp. 120–123, 2000.

[20] J. Gu, Z. Wang, J. Kuen et al., "Recent advances in convolutional neural networks," *Pattern Recognition*, vol. 77, pp. 354–377, 2018.

[21] Y. Hua, J. Guo, and H. Zhao, "Deep belief networks and deep learning," in *Proceedings of the 2015 International Conference on Intelligent Computing and Internet of Things*, pp. 1–4, IEEE, Harbin, China, January 2015.

[22] J. Masci, U. Meier, D. Cireşan, and S. Jürgen, "Stacked convolutional auto-encoders for hierarchical feature extraction," *International Conference on Artificial Neural Networks*, pp. 52–59, Springer, Berlin, Heidelberg, 2011.

[23] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[24] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, 2012.

[25] J. Deng, "A large-scale hierarchical image database," in *Proceedings of the IEEE Computer Vision and Pattern Recognition*, Miami, FL, USA, June 2009.

[26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.

[27] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alexander, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Proceedings of the Thirty-first AAAI Conference on Artificial Intelligence*, San Francisco, CA, USA, February 2017.

[28] R. Girshick, J. Donahue, T. Darrell, and M. Jitendra, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587, Columbus, OH, USA, June 2014.