

# Application of Convolution Network Model Based on Deep Learning in Sports Image Information Detection

Xiaoqiao Zhang<sup>1</sup>

<sup>1</sup>The Department of Information, Beijing City University, Beijing, China

**Abstract:** In recent years, convolution neural network has achieved great success in single image super-resolution detection. Compared with the traditional method, this method achieves better reconstruction detection effect. However, the network structure of the existing reconstruction model is shallow, and the convolution kernel has a small acceptance, so it is difficult to learn a wide range of motion image features, which affects the quality of motion image information detection. Aiming at the problems and shortcomings of the existing sports image information detection based on convolution neural network, this paper proposes the application of convolution network model based on deep learning in sports image information detection. In this paper, we get the average SSIM value from the data of set5, set14, bsd100 and urban100 by using the X4 model of different algorithms. The average SSIM value of set5 is 0.865, which shows that the quality of sports image reconstruction and the reconstruction efficiency of the model can be improved by using the local image features of different scales, which provides technical support for sports image information detection. The research in this paper has important practical significance for the further development of the two and the reform of the convolution network model in sports image information detection.

## 1 INTRODUCTION

In recent years, with the development of artificial intelligence, deep learning has shown amazing effect in processing massive data. Its powerful ability to learn and process data exceeds that of human beings in some fields [1-3]. For example, in the field of image recognition competition, the accuracy of image recognition based on deep learning algorithm has far exceeded the traditional algorithm. In addition to these achievements, new breakthroughs in deep learning include natural language processing, speech recognition, image semantic segmentation, image recognition, target detection, etc. [4-5].

In the application of deep neural network, there is a type of transmission network. Its typical application is to turn an ordinary photo into the style of a famous painting. In the process of texture synthesis using convolutional neural network, it is found that the statistical features of feature mapping in convolutional neural network can reflect the style of image, and feature mapping itself is the deep feature representation of network input image, reflecting the content features of image [6-8]. Then, we can adjust the randomly initialized image to the image similar to the famous painting in style, but the content is still ordinary photo. Although image stylization has achieved good results, there are still some problems to be solved: one is the puzzle problem. Even with the most advanced transformation network scheme, it usually takes 15 minutes to train a model [9-10]. Secondly, there are

some problems, such as the style representation of the results after migration is not obvious enough. Obviously, there is still room for improvement.

Analysis shows that the application of deep learning convolution network model in image information detection is still insufficient. In this paper, the application of convolution network model based on deep learning in sports image information detection is established. In the research, according to the actual situation of sports image information detection, the advantages of convolution network model are analyzed. The effective combination of the two can improve the detection effect of sports image information.

## 2 DEEP LEARNING AND BATCH NORMALIZATION LAYER IN CONVOLUTIONAL NEURAL NETWORKS

### 2.1 Overview of Deep Learning

Deep learning is closely related to the development of neural network, which is a mathematical model to simulate the synaptic connection of human brain. The earliest neural networks had only one layer of perceptron, which could only be used in binary classification, and could not solve the basic XOR problem. In the process of neural network training in 1980s, the problem of gradient disappearance often appeared which hindered the further development of neural network. Until 2006, unsupervised training was proposed to initialize the weights of neural

e-mail: zhqxq@bcu.edu.cn

networks and fine tune them with supervision to solve the problem of gradient disappearing. Since then, the level of neural network has become deeper and deeper, and the model has become more and more complex. On the other hand, due to the progress of computer hardware, it is no longer difficult to train a sufficiently complex neural network. Neural network is widely used in various computer vision problems, speech recognition, natural language processing and other fields. It has achieved the results that traditional machine learning algorithm cannot achieve, and has achieved better performance than human in some problems.

## 2.2 Batch Normalization in Convolutional Neural Networks

Convolution neural network is a deep neural network with convolution structure. Its development is closely related to the development of deep learning. The convolution neural network can effectively reduce the number of network parameters and alleviate the over fitting problem of the model through parameter sharing, which provides the possibility for training large-scale deep network with high-dimensional data such as images.

When training the neural network, we usually normalize the input data and the distribution becomes variance 1. When a shallow network is used, the output near the output layer cannot change significantly with the training of the model. This numerical instability increases the training complexity and the risk of over fitting. Therefore, for each layer of the network, before the output data is sent to the active layer, all the data should be forced to be under the unified data distribution, so as to reduce the instability of the data distribution and accelerate the convergence problem of the neural network. Specifically, it is assumed that the input of the network batch normalization layer is batch data:

$$B = \{x_1, \dots, m\} \quad (1)$$

output data:

$$\{y_i = BN_{\gamma, \beta}(x_i)\} \quad (2)$$

Calculate the mean value of batch data:

$$\mu_B \leftarrow \frac{1}{m} \sum_{i=1}^m x_i \quad (3)$$

## 3 CNN MODEL OPTIMIZATION

### 3.1 Data Preprocessing

Data in this paper include sports data set and large auxiliary image data set. The moving image data set is the target data set. Using Imagenet data set as auxiliary training data set, CNN model is trained. In order to detect the moving image information, it is necessary to preprocess the image.

Image is composed of image sequence, so image processing can be transformed into video key frame processing. Firstly, the SBD shot boundary detection algorithm is used to segment the image into multiple shots,

and then the key frames are extracted from the shot for detection. In order to improve the detection speed, only one key frame is extracted to represent the shot content, which is used for image semantic detection. After the above processing, the image data set is not large enough to train into a large-scale CNN network. Therefore, this paper introduces the Imagenet data set as the auxiliary data set to initialize the CNN model, and then uses the motion data to fine tune the CNN model after training.

### 3.2 Construction of CNN Model

#### (1) Neural layer structure of CNN

Convolutional neural networks usually contain multiple convolution layers, lower sampling layers and normalization layers. Finally, the full connection layer feature mapping is connected to a vector input to the final classifier to obtain the probability output. This section details the structure of each floor. The input of convolution layer is a number of two-dimensional planes. Each convolution kernel is connected to all input channels. Convolution is carried out in three-dimensional space to obtain the response output of position. Finally, convolution checks the whole input space and convolutes to get the feature map. Usually, multiple convolution kernels are set in each convolution layer, and each convolution kernel extracts different features, so each feature map represents the feature plane extracted by corresponding convolution kernel.

#### (2) Neuron activation function

Purpose of neuron activation function is to introduce nonlinear operation into neural network, which can simulate complex nonlinear function, improve the model complexity and classification ability of neural network, and enable neural network to deal with nonlinear classification problems. The convergence speed of the algorithm is much lower than that of the nonlinear training model. In addition, this kind of relu's unsaturated activation function does not need comparative normalization and similar data preprocessing, because the response value of larger input value can be amplified after the unsaturated activation function. Therefore, it is only necessary to subtract the average value of each dimension from each dimension of the input data so that the average value of each dimension is 0.

## 4 DISCUSSION

### 4.1 Validity of MFFRR Reconstruction Model in Sports Image Information Detection

In the motion image super-resolution algorithm based on convolution neural network, div2k provided by NTIRE 2017 super-resolution challenge is selected as the training data set. The data set is the benchmark data set of image super-resolution (2k resolution), which consists of 900 training images, 150 verification images and 150 test images. Because the reconstruction method deepens the depth of the network model, in order to avoid over fitting phenomenon, this paper expands the training set, rotates all training images horizontally, then rotates 900270 and

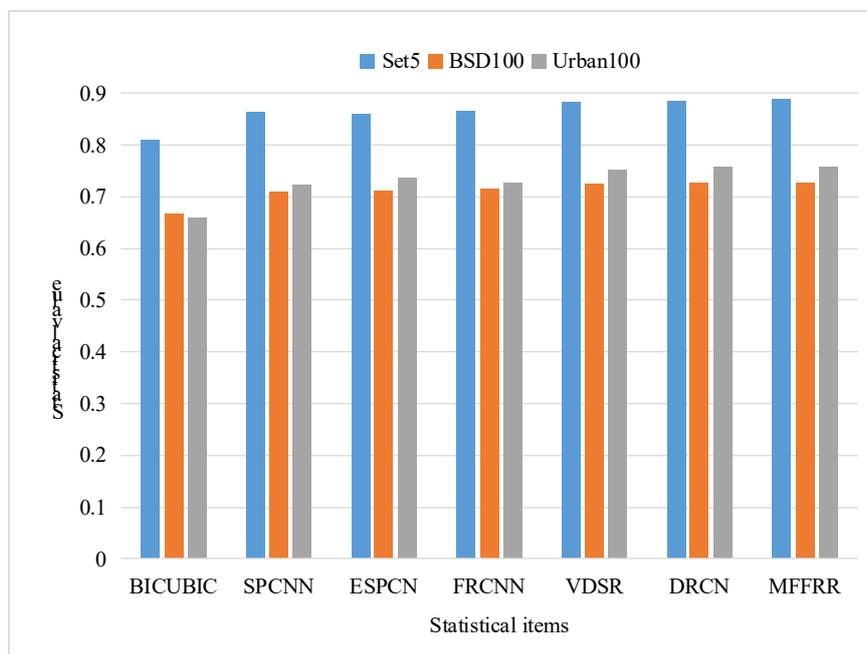
2700 respectively, and the down sampling is 0.5 times, 0.6 times, 0.7 times and 0.8 times of the original value respectively. Finally, the image number port in the training set is expanded by  $6 \times 6 = 36$  times, and the total number of images is  $900 \times 36 = 32400$ . The idea of image enhancement is to make the network model robust to image motion. In the experimental test, div2k test data set is not used because it is not public. Four benchmark datasets, set5, set14, BSD100 and urban100, are used to evaluate the quality of image super-resolution reconstruction. The set5 dataset contains five classic images, all of which are smaller and have lower resolution. The length and width of the image range from 250 to 450 pixels.

Data in Table 1 and Figure 1 are from X4 models with different algorithms in set5, set14, bsd100 and urban100

data sets. Table 1 and Figure 1 show that the SSIM index value of MFFRR model is better than that of most contrast reconstruction models based on convolution neural network. In addition, the restored image has superior performance in visual quality, and the edge contour and image detail information are more complete. Experiments show that the use of multi-scale feature extraction can obtain more abundant shallow features, and the use of multi-path recursive residuals can further deepen the depth of the network, and can learn richer image deep features. It shows that the use of different scales of shallow and deep local image features can improve the quality of sports image reconstruction and the reconstruction efficiency of the model the test provides technical support.

**TABLE 1** AVERAGE SSIM VALUES OF X4 MODELS WITH DIFFERENT ALGORITHMS ON SET5, SET14, BSD100 AND URBAN100 DATASETS.

data set	SCALE	BICUBIC	SPCNN	ESPCN	FRCNN	VDSR	DRCN	MFFRR
Set5	x4	0.81	0.863	0.861	0.866	0.884	0.886	0.888
Set14	x4	0.703	0.751	0.823	0.754	0.767	0.770	0.771
BSD100	x4	0.668	0.710	0.712	0.715	0.725	0.727	0.728
Urban100	x4	0.658	0.722	0.736	0.727	0.752	0.756	0.758



**Figure 1** average SSIM values of X4 models with different algorithms on set5, set14, bsd100 and urban100 datasets

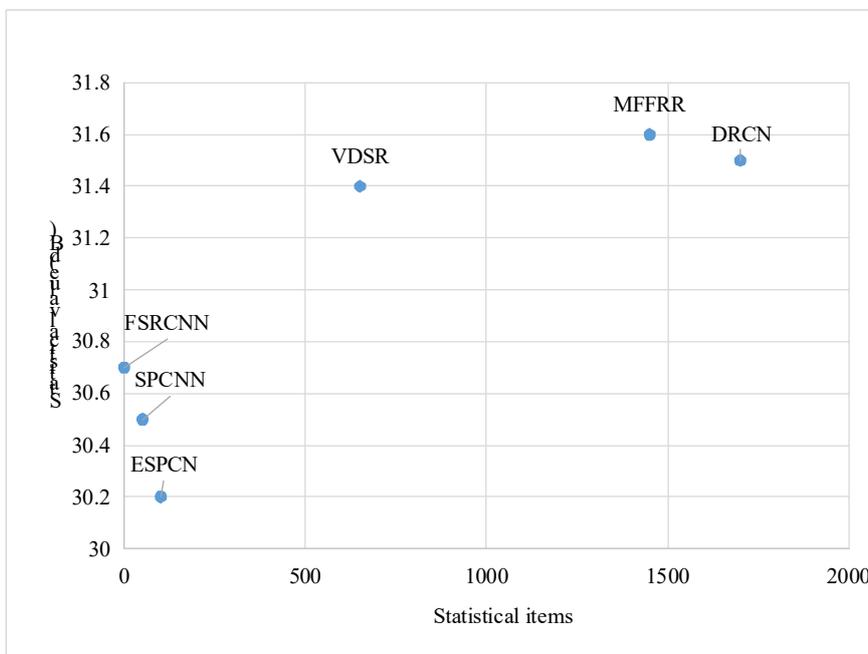
#### 4.2 Efficiency Comparison

In order to further verify the performance of MFFRR reconstruction model proposed in this paper, the number of training parameters and image reconstruction time of some classical methods are evaluated and compared on sets dataset. Figure 2 Comparison of accuracy and number

of parameters of set5 dataset under different methods when the scale factor is X4. When the magnification factor is x4, the training parameters and reconstruction performance of MFFRR model are compared with other models. The x-axis represents the number of parameters to be trained for each model, and the y-axis represents the average PSNR value of all images reconstructed by each model in the set data set. It can be seen from the analysis

in Figure 2 that the parameters needed to be trained in the multi-scale feature fusion recurrent residual network model are compared with those of SRCNN, FSRCNN and ESPCNN. Because the depth of network model in this paper is increased compared with other methods, there are more VDSR models. The network has 25 layers. Accordingly, the PSNR value of the reconstructed image

is improved compared with other methods under appropriate parameters. The results show that the MFFRR model proposed in this paper has good image quality. It is proved that the MRN can make full use of local multi-scale features and use hierarchical features to obtain accurate SR images and recover more image details.



**Figure 2** Example of a figure caption. (*figure caption*)

### 4.3 Research Prospect

Starting from the detection of sports image information, this paper deeply analyzes the root of the construction and optimization technology of depth model. According to the characteristics of depth model, it analyzes the transfer learning technology based on deep network. Deep learning in the map has achieved high resolution and depth in classification, target recognition and recognition. Xi Jinping has shown strong learning ability on large scale data. Deep learning model usually contains millions of learning contents. The parameter scale of learning parameters is much larger than that of traditional shallow model, which also gives deep model stronger learning ability, but it also increases the complexity of model training.

In view of the sports image resolution reconstruction technology involves many theoretical techniques, reconstruction models and practical application scenarios, each field has different requirements for image quality. This paper only studies and analyzes the reconstruction model based on convolutional neural network. Compared with other SISR methods, this method improves the reconstruction quality obviously, but there are still some shortcomings.

(1) Although the quality of motion image reconstruction is improved compared with other comparison models, this model is based on deep network, which has the problems of high complexity and slow

image reconstruction speed, and cannot meet the application scenarios with high real-time requirements. How to reduce the complexity of reconstruction model and improve the efficiency of image reconstruction without reducing the impact of image reconstruction quality is a problem that needs further study.

(2) In real scene, the degradation process of moving image is different from the image degradation model used in the experiment. The image quality in practical application scene is affected by many factors. The degradation process is very complex, and the quality of low score image is different. Therefore, it is necessary to further understand the process of image degradation.

### 5 CONCLUSIONS

This paper takes the application of convolution network model in sports image information detection as the main line of research. After research, this paper considers that convolution network model is an important part of sports image information detection. Through data testing, the convolution network model of deep learning can improve the quality of sports image reconstruction, enhance the ability of sports image information detection, and can improve the detection speed. From the test and analysis results of this paper, in order to make full use of the convolution network model in sports image information detection, we must combine the convolution network model with the actual situation and image features of

sports image information detection, and pay attention to scientific introduction. This study has achieved ideal results and made a certain contribution to the application of convolution network model of deep learning in sports image information detection.

## REFERENCES

1. Hao, X., Zhang, G., & Ma, S. (2016). Deep learning. *International Journal of Semantic Computing*, 10(03), 417-439.
2. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., & Ghafoorian, M., et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42(9), 60-88.
3. Zhao, Z., Chen, W., Wu, X., Chen, P. C. Y., & Liu, J. (2017). Lstm network: a deep learning approach for short-term traffic forecast. *Iet Intelligent Transport Systems*, 11(2), 68-75.
4. Shrawan, R., Shloak, G., & Basant, A. (2018). Devanagri character recognition model using deep convolution neural network. *Journal of Statistics & Management Systems*, 21(4), 593-599.
5. Jeong, J., Lee, D., Jung, H., & Yang, H. (2020). Automatic convolution neural network model compression framework for resource-constrained embedded systems. *Journal of KIISE*, 47(2), 136-146.
6. Mari, K., & Ganesh, V. (2020). Multi-labelled emotion with intensity-based sentiment classification model in tweets using convolution neural networks. *International Journal of Advanced Trends in Computer ence and Engineering*, 9(2), 1650-1656.
7. Santos, M. A. D., Vveinhardt, J., Calabuig, F., & Francisco Javier Montoro Ríos. (2017). Involvement and image transfer in sports sponsorship. *Engineering Economics*, 27(1), 78-89.
8. Bachleda, C., Fakhar, A., & Elouazzani, Z. (2016). Quantifying the effect of sponsor awareness and image on the sports involvement–purchase intention relationship. *Sport Management Review*, 19(3), 293-305.
9. Zhu, K., & Ying, L. (2016). Information source detection in the sir model: a sample path-based approach. *IEEE/ACM Transactions on Networking*, 24(1), 408-421.
10. Gupta, N., Pillai, G. V., & Ari, S. (2018). Change detection in landsat images based on local neighbourhood information. *Image Processing, IET*, 12(11), 2051-2058.