Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 6
ISET 2024

Systematic Analysis of Image Restoration Methods Based on Deep Learning

Zhengqi Hu*

College of optoelectronics, Xiamen University of Technology, XiaMen, 361024, China

* Corresponding Author: 1809030217@stu.hrbust.edu.cn

Abstract. In today's digital age, people often use images to convey information. However, during the process of sending images, they can become damaged for various reasons, causing the images to lose their original meaning. In such cases, image repair is necessary to restore the damaged areas to their original, undamaged appearance. Traditional image repair methods often struggle to produce effective results for highly damaged images. However, image repair using deep learning techniques can address these challenges effectively. This paper will detail the impact of image repair on Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory networks (LSTM), Convolutional Recurrent Neural Networks (CRNN), and other models, as well as the advantages and disadvantages of each approach. From the data, it is evident that combining multiple models for image repair is the most effective method. This approach reduces the consumption of hardware resources and time required for image repair, thereby improving the overall efficiency of the process.

Keywords: CRNN; Deep learning; LSTM; Image Restoration.

1. Introduction

Images have always been a good way for people to understand information, to communicate with each other, and to preserve useful information. However, due to various reasons, the image is damaged, and the damaged image will hinder the transmission of information, so not all images can fully express the meaning it wants to express, at this time the role of the image will disappear, and this kind of thing is very common in life, at this time, the image will be used to repair. Image repair can restore the damaged or lost area of the image to its original appearance as much as possible. However, the effect of traditional image restoration is not satisfactory. In many cases, the information and effect expressed by the repaired result are quite different from the meaning of the original image [1]. In recent years, with the entry of the digital age, digital image has become an important way of modern information expression. In the process of digital image transmission, the image will be lost or other problems, at this time, the use of traditional restoration methods is not enough to solve, not only that, for example, in the restoration of cultural relics, the use of traditional restoration methods is not good. Hardware and software have continuously improved in recent years, in many areas, deep learning has made significant progress, and image repair is one of them. The use of deep learning can better improve the repair effect of images, so that the information after image repair is consistent with the information expressed in the original image, and the image repair based on deep learning can not only repair traditional images such as cultural relics and artworks, but also repair digital images. Therefore, image repair based on deep learning is particularly important in this era, which is now the most mainstream repair method.[2]

In Chapter 2, this paper will introduce three methods of image restoration based on deep learning, and then introduce the method of combining two of them. The third chapter will focus on the application of these methods. Finally, the conclusion is drawn in chapter four.

2. Classification and foundation of deep learning

2.1. CNN's theoretical analysis

Convolutional Neural Network is CNN for short. A convolution layer, pooling layer, nonlinear activation layer, and fully connected layer are all part of CNN's neural network model. CNN can extract and abstract image features through multiple convolution layers and pooling layers, use convolution operations to capture local features in the image, and reduce the data dimension through pooling operations. Parameter sharing and space invariance are characteristics of CNN, it can effectively deal with the changes of translation, scaling and rotation in the image. The role of CNN in image restoration includes image denoising, image patching and image super resolution [3]. If a CNN model is trained, it can learn the internal structure and features of the image, so that it can fill in damaged or missing pixels, and can obtain high-quality repaired images. These make CNN an effective tool for processing image tasks. Compared with traditional image classification methods, the CNN-based classification method is a learning process that involves the end-to-end process. The original image should be inputted, the training and prediction process should be conducted in the network, and the result should be outputted. The limitations of traditional classification methods are overcome because this method does not require manual extraction of specific image features. This is why CNN has the greatest advantage in image classification. The CNN model is shown in the figure1.

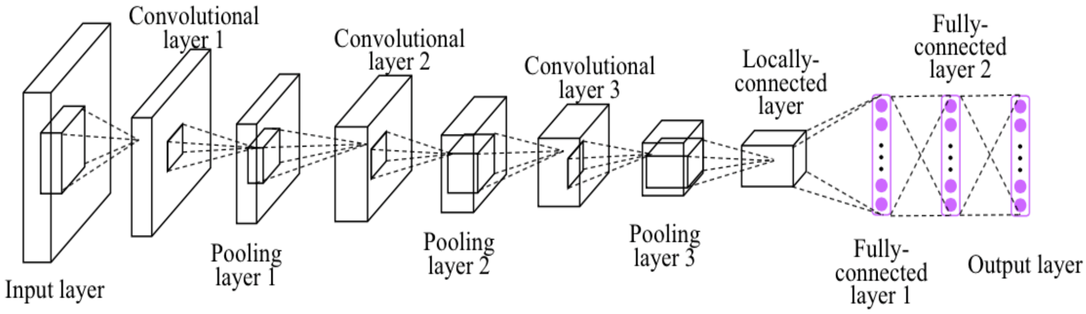

Figure 1. CNN

2.2. RNN's theoretical analysis

Recurrent Neural Networks is RNN for short. Sequence data is used as input in a recurrent neural network, makes all nodes are connected by a chain during recursion in the evolution direction of the sequence. An input layer, a hidden layer, and an output layer are all part of it.

The Hidden Layer features an arrow that indicates the cyclic update of data. This arrow can pass the information processed at the current time to the next time. This is why RNN can realize the function of time memory. The internal memory of an RNN can be utilized to process any arbitrary sequence of inputs. Previous information is remembered by the network and used in the current output calculation, as stated by the memory function., which is the connection between the nodes between the hidden layers has disappeared, but they are still connected, and the output of the hidden layer is a combination of the input layer and the output from a previous time. The RNN model is shown in the figure2.

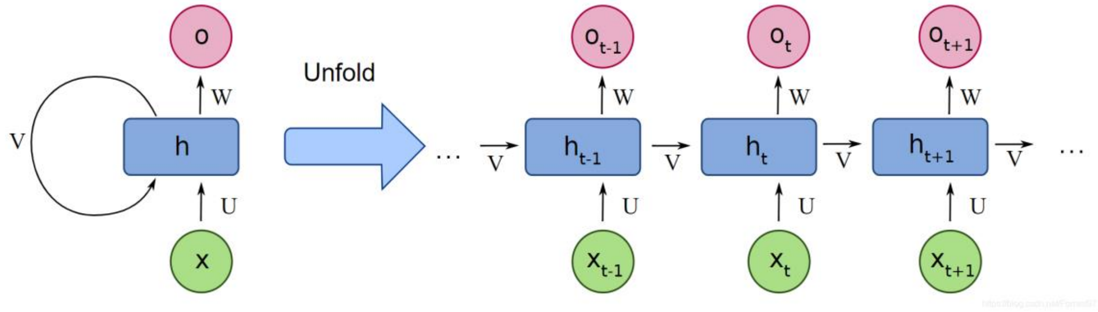

Figure 2. RNN

However, RNNS have a fatal problem, which is that they do not handle long-term dependency problems well. When the prediction point is far from the dependent data, a long-term dependence issue arises., It makes acquiring the necessary knowledge challenging. This is due to the fact that gradient explosion and dispersion are more likely to happen as the time interval increases. Gradient dispersion will affect the updating of network parameters, making it impossible to learn new information, while gradient explosion will make the updating of network parameters too fast to learn useful information

2.3. LSTM's theoretical analysis

Long Short-Term Memory Network is LSTM for short. LSTM is a deep learning model that was developed with the specific purpose of resolving the long-term dependence issue of general RNNS, which is often used to process sequence data [5]. The four components of an LSTM are an input gate, an output gate, a forgetting gate, and a cell state. The mechanisms in LSTM are designed to enhance its ability to handle long-term dependencies in sequences. The role of the forgetting door: The model can forget old information when new information is input, using the forgetting door to complete the process. Update cell status picture: Update the old cell status picture to the current cell status picture. The output gate generates a value based on the cell state, which is then processed through a filtering process. The LSTM model is shown in the figure3.

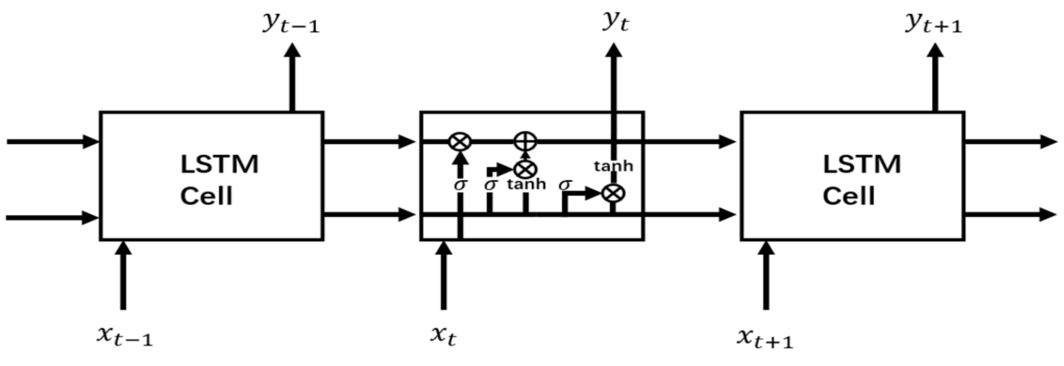

Figure 3. LSTM

LSTM adds a cell state to its structure, influencing information changes based on current input, gate structure, previous cell state, and hidden layer state. Compared with RNN, LSTM can not only deal with short-term dependency problems, but also deal with long-term dependency problems.

2.4. CRNN's theoretical analysis

Convolutional Recurrent Neural Network is CRNN for short. The main use of the convolutional recurrent neural network structure is to recognize text sequences of different lengths from start to finish. Transforming text recognition into a sequence learning problem based on timing without first cutting individual text, In short, recognizing sequences through images.

From bottom to top, there are three parts that comprise the entire CRNN network structure: Input images are used by CNN to extract features and generate feature maps using depth CNN. The feature sequence can be predicted and each feature vector in the sequence can be learned by RNN using bidirectional RNN, also produce the distribution of prediction labels (correct value); A final tag sequence is created via CTC loss, which transforms a number of tag distributions extracted from the loop layer. The CRNN model is shown in the figure4.

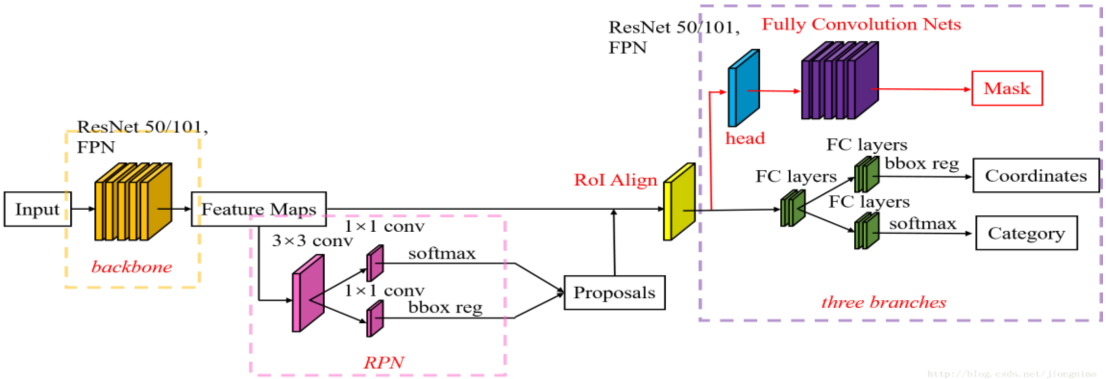

Figure 4. CRNN

In theory, CRNN has the following advantages :1. It is end-to-end and improves the recognition speed 2. It can process sequences of any length and solve the problem that the aspect ratio of Chinese text images varies greatly in different scenes. No predefined dictionary is required to improve the accuracy and robustness of text recognition. 4. Smaller models with fewer parameters, it means no full connection layer required.

3. Technical analysis of image restoration methods based on deep learning

3.1. Image repair based on CNN

There are many advantages for CNN-based image repair methods. For example, key features can be extracted from input images without the need for manual feature extraction, which greatly simplifies the feature extraction process. In addition, CNN can automatically learn image features and has strong generalization ability [6]. Therefore, it can learn image features from different types of image data and then apply these features to various image restoration tasks. While CNN has advantages, it also has major disadvantages. For example, when performing convolution operations, only local information is considered. Therefore, under certain conditions, this method cannot take into account the influence of global information, such as segmentation tasks or global feature extraction, which has certain limitations.

To solve the problem that CNN cannot process global information, Yu introduced an image repair technology that utilizes CNN and Transformer. The advantages of the two methods are integrated and applied to image restoration to improve the quality of image restoration. In the process of restoration, Yu introduced a sharpening module to improve image restoration quality by refining local details due to blurred results due to image content complexity [7]. Transformer technology is used to repair the image network structure frame, and the technology used in machines differs from Transformer technology. There are two coding parts, namely the first coding module and the second coding module. The network structure comprises a Transformer module, a decoding module, and a sharpening module. The first coding module is composed of a fully connected layer, whose main function is to extract the feature vector of each small block. The second coding module employs a convolutional neural network to reduce image size and extract local information. The Transformer module is responsible for extracting higher-level semantic features from input feature vectors and projecting them into the codebook. The decoding module in the model is to further decode the output of Transformer. The sharpening module uses the reconstructed image from the previous decoding module to enhance local details and eliminate fuzzy areas in the reconstructed image.

The datasets used include CelebA and Paris Street View. RFR is a kind of progressive network with coarse restoration and then fine restoration, GAN is a kind of image restoration method based on generative adversal network, and LGnet is a kind of image restoration algorithm based on U codec. CNN+Transfomer compared with data from other models is shown in table 1.

Table 1. CNN+Transfomer compared with data from other models

MODEL	PSNR	SSIM
RFR	20.11	0.813
GAN	23.89	0.904
LGNET	24.19	0.921
CNN+Transfomer	26.74	0.935

In order to better depict the details of image restoration, quantitative index evaluation was carried out on these methods. The indexes of PSNR and SSIM are extremely similar since they both take into account the variation in pixel level between the original and restored images. The repair method proposed by Yu is the highest in the value of PSNR and SSIM, which indicates that the repair method proposed by Yu can better repair the details of the image.

3.2. Image repair based on RNN

Generative image restoration algorithms based on RNN usually decompose the image into a pixel sequence, and use RNN to traverse the pixel sequence of the entire image, learn the feature distribution of the global sample, and thus generate pixel values of the missing regions one by one. The generative image restoration algorithm based on RNN faces the problem of high computational complexity and long time when processing large-scale images. In the later stage of traversing pixels, the correlation between pixels gradually weakens, which makes the algorithm perform poorly for complex missing patterns.

Zhang designed the module SARNN capable of discovering and discovering interest graphs generated from input element graphs [8]. SARNN consists of a two-wheeled network of recurrent neural RELUs and RNNS. The primary purpose of the RNN's first round is to create a feature map that compiles the input image's location points' up and down information. The second round of the RNN is used to further collect non-local up-down information to generate a globally aware feature map.

SARNN first has two layers of RNN network. In order to accumulate information throughout the image, the initial layer of an RNN network effectively propagates information, which enhances the coherence of semantic information and can also recognize the context between remote feature values to establish semantic continuity. To generate global perception feature values, RNN will gather local context information. To create a feature map that is globally aware, Layer 2 RNNS collects additional non-local context data. Observing that direction is crucial for identifying significant clues in shaded and unshaded regions. Shaded regions were detected by ZHANG using this two-wheeled, four-way RNN architecture.

The drop was projected in four primary directions by Zhang through the use of an RNN model [8]. A network with ReLU and IRNN. The input image's position points are used to create a feature map for contextual information during the first round of IRNN; To create a feature map that is globally aware, IRNN gathers non-local context information in the second round. To selectively highlight projected shadow features, ZHANG included an additional branch to gather spatial context information. It also visualizes the attention maps generated by SARNN. Zhang's method is compared with the other five methods on the RICE1 dataset is shown in table 2.

Table 2. Zhang's method is compared with the other five methods on the RICE1 dataset

MODEL	PSNR	SSIM
CGAN	26.54	0.90
CyleGAN	25.88	0.89
Spa-GAN	30.23	0.95
SC-FEGAN	-	-
GMCCNN	-	-
SARNN	30.85	0.96

And Zhang's method is compared with the other five methods on the RICE2 dataset is shown in table 3.

Table 3. Zhang's method is compared with the other five methods on the RICE2 dataset

MODEL	PSNR	SSIM
CGAN	25.38	0.81
CyleGAN	23.91	0.79
Spa-GAN	28.36	0.90
SC-FEGAN	20.22	0.73
GMCCNN	19.21	0.74
SARNN	28.56	0.91

It can be seen from the figure that Zhang's method is outstanding in both the value of PSNR and the value of SSIM, so this method is very effective [8].

3.3. Image repair based on LSTM

Because of the design defect of RNN, it cannot effectively capture the information of a long time ago, and it is prone to gradient dispersion and gradient explosion problems with the increase of time interval. Gradient dispersion will affect the updating of network parameters, resulting in the network unable to learn new information, while gradient explosion will make network parameters updated too fast, unable to learn useful information, affecting network performance. By implementing a gating mechanism to regulate the accumulation speed of information, LSTM can effectively solve the gradient dispersion problem of RNN. Through this mechanism, LSTM can not only capture short-term dependencies of sample sequences, but also learn long-term dependencies.

Zhang proposed a progressive image restoration algorithm based on generative adversarial neural networks. The algorithm divides the whole repair process into several sub-stages, each of which only needs to focus on repairing a part of the area, and connects these sub-stages through LSTM [9]. In each sub-stage, to produce forged images, a generation network with a short encoder-decoder structure is employed, the input images are judged to be real data by using the global and local discriminant networks. The algorithm includes data preprocessing module, model training module and image restoration module. In the data preprocessing part, the data in the data set is processed to meet the network requirements. Firstly, the images in the data set are randomly clipped and training data is made. Then, through masks of different sizes, training sets in the form of real data, damaged data and film are made for different stages for training subsequent neural networks. In the part of model training, firstly design and build the neural network model corresponding to the algorithm, and then use the prepared training set to train the model. The model consists of generative network, global discriminant network, local discriminant network and LSTM network. To generate forged images, the generative network is employed based on the damaged input images. The input image's realness or forged image generated by the generator is judged by the discriminant network. The global discrimination network's input is comprised of both genuine and fabricated images, and the input of the local discriminant network consists of the area where the real image and forged image are located. LSTM is used to connect the bottleneck layers of the different stages of the generation network and pass the information repaired in the previous stage to the next stage. In order for the image repair module to complete the repair task, it is necessary to use the trained generating network to repair the damaged image input. The repair effect table of three repair methods is shown in table 4.

Table 4. The repair effect table of three repair methods

MODEL	PSNR	SSIM
CE	24.420	0.8685
GLI	26.567	0.8831
LSTM	27.384	0.8996

Among them, CE and GLI are classical methods in the field of semantic image restoration. It can be seen from the results that the average PSNR and average SSIM values of the repair method proposed by Zhang are the highest among all methods, so Zhang's algorithm is effective [9].

3.4. Image repair based on CRNN

Because CNN model and RNN have their own unique advantages and disadvantages, CRNN model combines the advantages of RNN model describing image statements with the advantages of traditional CNN model feature acquisition, and proposes a new RCNN fusion model

Huo proposes a model that combines the advantages of description image statements of the RNN model with the feature acquisition advantages of the traditional CNN model, and proposes a new RCNN fusion model to build a deep-level model, and input the shallow and deep information extracted by multi-layer convolutions into the RNN [10]. The timing information and semantic information of the image are mined by using its structural advantages, and the U-net network structure is used to restore and reconstruct the obtained information. The module can be divided into four modules.

The first module is the preprocessing module of data set. The aim of preprocessing the image data set is to minimize the impact of other irrelevant factors on the final recognition outcome. Furthermore, the preprocessing method can enhance the model's robustness by making the data set more diverse and randomized. The second module is the image feature acquisition module. We need to take some methods to reduce the dimension of image data, that is, we say feature extraction of the image, extract important information in the image, and then use the extracted feature information to carry out more operations on the image, which can significantly reduce the space required for storage and the complexity of the calculation process. The third module is semantic information acquisition module. An image is first divided into many blocks, and then feature extraction is carried out on these blocks to obtain a series of features. Each small block represents an object of interest, and these objects are identified through the recognition layer to extract higher level semantics. Using the knowledge related to the domain object to extract the local features of the object, such a retrieval mechanism will be more effective. The fourth module is the recovery and reconstruction module. The U-net network structure is used for restoration and reconstruction, and the information of the image is extracted and combined in the process of coding, so that the information obtained by U-net is more comprehensive. During the decoding process, the image is up sampled to the original size. The repair effect table of three repair methods is shown in table 5.

Table 5. The repair effect table of three repair methods

MODEL	PSNR	SSIM
Deep reinforcement learning and GAN	27.58	0.965
CNN+GAN	27.92	0.959
CRNN	28.09	0.961

The table indicates that when the cyclic convolutional neural network model is employed to repair the distorted image, the PSNR of the reconstructed image is the highest among the methods and the SSIM effect is also reasonable, the algorithm's effectiveness can be demonstrated by this.

4. Conclusion

The significance of image restoration is to recover the information of the original image from the damaged, missing or noisy image. With the rapid progress of deep learning technology, there are more and more models for image repair, and the repair effect of each model is not the same. Therefore, this paper introduces the repair effect of each model and the method combining several models. The data above shows that, since the image restoration methods of a single model have their own limitations and advantages, the effect of image restoration methods based on a single model is not as good as that of combining multiple models. Therefore, if the advantages of each model can be combined, the effect of the generated method will be greatly improved.

Building upon this foundation, several challenges are evident. Firstly, the development of a model capable of automatically detecting and repairing damaged areas of images is a primary objective. While many repair methods yield excellent results, their practical implementation is hindered by the substantial computing resources and costs required. Therefore, it is crucial to effectively reduce the computing time and resources necessary for image repair. Additionally, future endeavors should include conducting experiments or simulations on the current algorithms, if conditions permit, to further verify the discrepancies between real-life test results and theoretical outcomes.

References

[1] Lugmayr A, Danelljan M, Romero A, et al. Repaint: Inpainting using denoising diffusion probabilistic models. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 11461-11471.

[2] Zhang H, Xu H, Tian X, et al. Image fusion meets deep learning: A survey and perspective. Information Fusion, 2021, 76: 323-336.

[3] Su J, Xu B, Yin H. A survey of deep learning approaches to image restoration. Neurocomputing, 2022, 487: 46-65.

[4] Das S, Tariq A, Santos T, et al. Recurrent neural networks (RNNs): architectures, training tricks, and introduction to influential research. Machine Learning for Brain Disorders, 2023: 117-138.

[5] Bouktif S, Fiaz A, Ouni A, et al. Multi-sequence LSTM-RNN deep learning and metaheuristics for electric load forecasting. Energies, 2020, 13(2): 391.

[6] Liu Y, Pu H, Sun D W. Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices. Trends in Food Science & Technology, 2021, 113: 193-204.

[7] Yu F. Research on image restoration based on Transformer Technology. Fuyang normal university, 2023.

[8] Zhang E. Image completion algorithm based on attention mechanism research. Nanjing university of posts and telecommunications, 2021.

[9] Zhiqiang Z. Based on the depth study of image restoration method research. Donghua university, 2020.

[10] XingYu H. Circular convolution neural network based on the depth of distorted image restoration. Changchun university of science and technology, 2023.