A Conceptual Representation of Image Restoration using Deep Neural Network Structure in Digital Image Processing

Student ID: 2019380130

Student Name: Bisenbay Guldariya

Course: Digital Image Processing

Course Code: U10M12021

June 16th 2022

CONTENTS

1. Introduction:

Decades ago, people could save important moments in their lives in paper based photos to be framed or stored in the albums. The great disadvantage of these photos was that the oxidation could easily damage the qualities of the photos and they would easily get destroyed. However, as the technology industry was developing at high speed in different areas, it gave a great solution to virtually store all photos in a computer. On the other hand, it was not enough to bring the damaged images into their original state. For that reason, the sensational discoveries in Artificial Intelligence and Digital Image Processing gave new possibilities to enhance the image quality of centuries ago. Artificial Intelligence, as well known as AI is the science of instilling intelligence in machines in order to be able to do tasks that traditionally require human beings. AI based systems are evolving rapidly in terms of application, adaptation, processing, speed and capabilities [1]. Because of the high potential that AI holds, people are aiming to increase the usability of AI into a broad range. Therefore, many fields have conjugated to drive AI development and the major branch of it is image processing. Image processing method concentrates on applying some operations on an image in order to perform data extraction from the image or to get an improved image. It is a technology that is long-established in the development of statistical vision systems. The classical approach to image processing consists of four processing stages: preprocessing, segmentation, feature extraction and classification [2]:

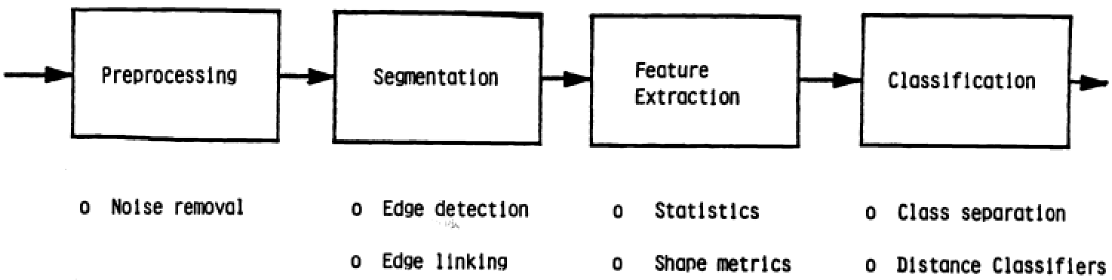


Figure 1. The Classical Approach to Image Processing

Traditional image restoration algorithm uses diffusion-based methods that focuses on missing scene information and perceptual information. This method spreads the local structure to the location part or constructs one pixel of the missing part every time, while maintaining surrounding pixels' consistency. Therefore, these traditional algorithms focus on defect filling, but incompitent in solving blur and fading parts of the photos.

Another better solution is deep learning a branch of machine learning. It is an algorithm that uses artificial neural networks as the architecture to characterize and learn data. In recent years, deep learning has been widely used in computer vision tasks due to its powerful representation of image data. Some researchers have tried to apply deep learning into image tasks to improve the high quality content structure and structure of image restoration algorithms. Texture capabilities, especially deep convolutional neural networks and generative adversarial networks, perform well in feature extraction and image generation and have good image restoration effects on old photos. Deep neural network calculations and other neural network methods mainly rely on the illusion of a pre-trained neural network to fill in the missing parts of the image. Repairing technology based on deep neural network saves time and effort to perform the restoration [3, 4]. As one of the branches of image processing technology, this technology can be used to repair damaged photo paintings and remove redundant texts. Considering all the stated advantages of deep neural networks, more people are using deep neural network to restorate old photos.

The main idea of this work is to have a better understanding about concepts of image processing, methods to restorate the damaged and blurred images, experiment design, process and results. Then by comparing test final outcomes.

2. Concepts of Image Restoration and Processing of Damaged Images Based on Deep Neural Network

2.1. Deep Neural Network Concept on Image Restoration

Deep learning models are usually composed of neural networks.The commonly used deep learning models are Convolutional Neural Networks (CNN) and Generative Neural Networks (GNN). Compared with traditional image restoration modeling, the deep learning-based method takes advantage of big data [5–7]. It learns the deep features of the image from training to obtain the representation ability and significantly improves the quality of image restoration [8].

2.2. Convolutional Neural Network

Convolutional Neural Network is a neural network that was specifically designed to use the spatial structure relationship to reduce the number parameters that are necessary to learn and improve backpropagation algorithms.

2.2.1. Convolutional layer

The convolutional layer consists of many neurons, and the shared weights among the three network connection methods are used between every two convolutional layers. In image processing, the convolution operation is generally composed of several convolution kernels, which are needed to extract different features in the image, such as horizontal, vertical, and diagonal edges, and the deeper the convolutional layer, the higher the level of features extracted. CNN extracts features by automatically learning the weights of the convolution kernels during training, and the features extracted by different convolution kernels will be combined into the input of the next layer, so more advanced features can be obtained [9–11].

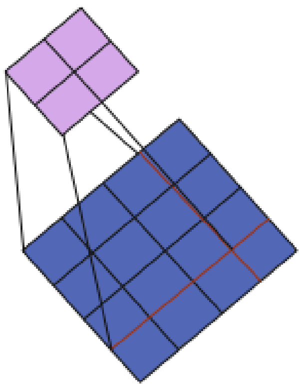


Figure 2. Covolution Calclation Process

2.2.2. Pooling Layer

The pooling layer is one of the commonly used components in the current convolutional neural network. It imitates the human visual system to reduce the dimensionality of the data and uses more advanced features to represent the image. Pooling has the effect of reducing information redundancy, improving the scale invariance of the model and preventing overfitting.

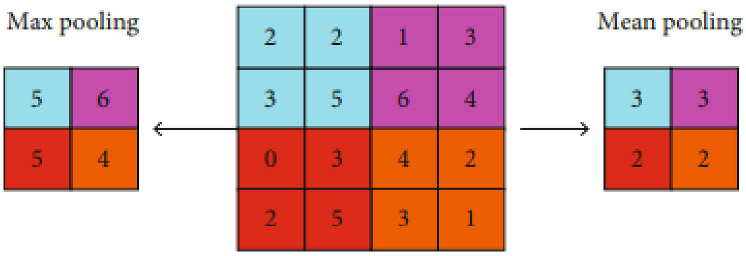


Figure 3. Pooling Operation

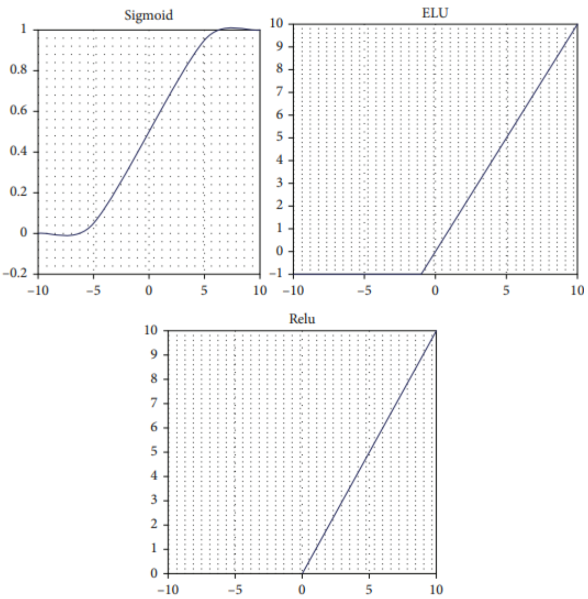
2.2.3 Nonlinear Activation Layer

In order to solve the problem of inseparable nonlinearity, the convolutional neural network usually introduces an activation function layer, so that the network has the learning ability of nonlinear mapping. Common activation functions include Sigmoid function, ELU function, and ReLU function [12–14]. The formulas and according graphs are provided below:

$$\mathrm{Sigmoid}(x) = \frac{1}{1+\exp(-x)},$$

$$\mathrm{ELU}(x) = \max\left[\alpha(\exp(x)-1), x\right]\alpha > 0,$$

$$\mathrm{ReLU}(x) = \max(0, x).$$



2.3. Generative Confrontation Network

Confrontation network is a generative model which was developed in recent years, and it is used in many applications, such as image generation and text generation. The model has two main characteristics. 1) The model does not require any prior assumptions; 2) The model relies on the forward propagation of its generator to generate real-like samples, which is simple. GAN is divided into two parts: generation network and discriminant network. The steps are as follows: first, the generation network generates a sample that is as similar to the original sample as possible according to the input information, then judges the generated sample and the original sample through the discriminant network, then continues to repeat this process, and finally reaches a balance [15]. In the figure 4, Z is an input information, Generator means a generation network, Xfake is a sample generated by generation network and Xreal represents a real sample. Between Xfake and Xreal distinguishes discriminator which represents discrimination network.

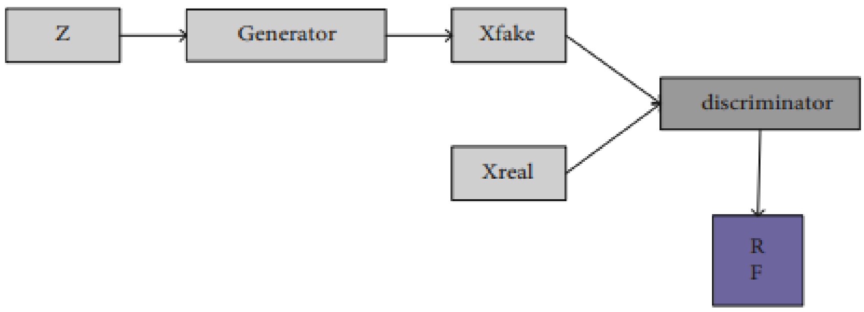


Figure 4. Schematic Diagram of Steps to Generate a Confrontation Network

3. Experiment on Image Restoration and Processing of Damaged Images Based on Deep Neural Network

3.1. Experiment Concepts, Design and Environment

The experiment that was held consisted of 2 parts: 1) Experiment on image blur restoration; 2) Experiment of image damage repair. In this experiment 2 image data sets were used and these data sets contained human faces and natural landscapes information. The data sets were named

CelebA-HQ and Places2. Because of the convenience and experience, the experiment tools were PyTorch 1.2.0 and the Ubuntu 16.04 operating system.

Package	Version
OpenCV-Python	4.5.18
SciPy	1.5.4
Torch	1.2.0
Wheel	0.33.1
TensorboardX	2.1

Table 1. Python partial Dependency Package

Parameter	Weight value
KL loss weight coefficient	20
Weight coefficient of appearance matching loss	20
Counter loss weight coefficient	1

Table 2. Loss Function Weight Value Design

The evaluation indicators that are used as the measurement of reconstruction quality in the image field and very suitable for the evaluation of image restoration of old photos in this experiment are peak-signal-to-noise ratio PSNR and structural similarity SSIM whose value range is [0, 1]. The logic of the measurements is the larger the values, the repair effect will be better.

3.2. Peak-signal-to-noise ratio

This method evaluates the generated image by measuring the degree of distortion or noise level of the image, and the evaluation result is expressed by the ratio between credible information and noise. The higher the PSNR value, the smaller the distortion of the image conversion process, and the more realistic the generated image. PSNR is the most used objective evaluation method when evaluating generated images. The calculation formula of PSNR:

$$\begin{cases} \text{MSE} = \dfrac{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left[I_0(i,j) - I(i,j)\right]^2}{M \times N}, \\ \text{PSNR} = 10 \log\left(\dfrac{G_f^2}{\text{MSE}}\right). \end{cases}$$

Where, MSE = square error

(i, j) = position of the real image

I_0 (i, j) = the pixel value of the real position of the real image at that position

I (i, j) = value at (i, j) of the restored image

M x N = area size of the restored image

3.3. Structural Similarity

This method evaluates the image by comparing the similarity of two given images. The larger the SSIM value gets, the higher the similarity between the restored image and the real image, and closer to the maximum SSIM value - 1. SSIM refers to the brightness, contrast, and structure of the image when modeling the degree of distortion of the image, in order to give a comprehensive evaluation result that reflects the properties of the object structure in the scene modeling. Among them, the brightness and contrast are represented by the average and deviation of the images is standard, while the structural similarity between the images is represented by the covariance of the images. SSIM pays more attention to the similarity of two images, and it reflects the attributes of the object structure when modeling the scene. The calculation formula:

$$\text{SSIM}(x,y) = \frac{\left(2\mu_x\mu_y + C_1\right)\left(2\sigma_{xy} + C_2\right)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x^2 + \sigma_y^2 + C_2\right)}.$$

Where, x,y = input image

μ = average value of x

σ_x^2 = variance of x

μ_y = average value of y

σ_y^2 = variance of y

σ_{xy} = covariance of x, y

3.4. Comparison Algorithm Selection

The selected comparison algorithms for image restoration in this work are SRCNN (based on convolutional neural network), SRGAN(based on generative adversarial network), and Criminisi which is the original image image restoration algorithm based on texture synthesis.

The experiment has compared these algorithms with the proposed algorithm in the work, that later were found as more convenient.

3.5. Blur Image Restoration Experiment

For this section of the experiment, 1000 pictures were chosen from CelebA - HQ and Places 2 datasets. All algorithms were examined on these selected images and were explored with PSNR and SSIM for image blur restoration.

3.5.1. Analysis of PSNR Results in Blur Image Restoration Experiment

The figure 5 illustrates the results obtained by applying each algorithm on images. Comparing the PSNR of each algorithm on CelebA – HQ dataset is always slightly higher than the PSNR value of each algorithm on the Places 2 dataset which means that the repairing methods are shown to have better effect on faces rather than on the landscape scenes.

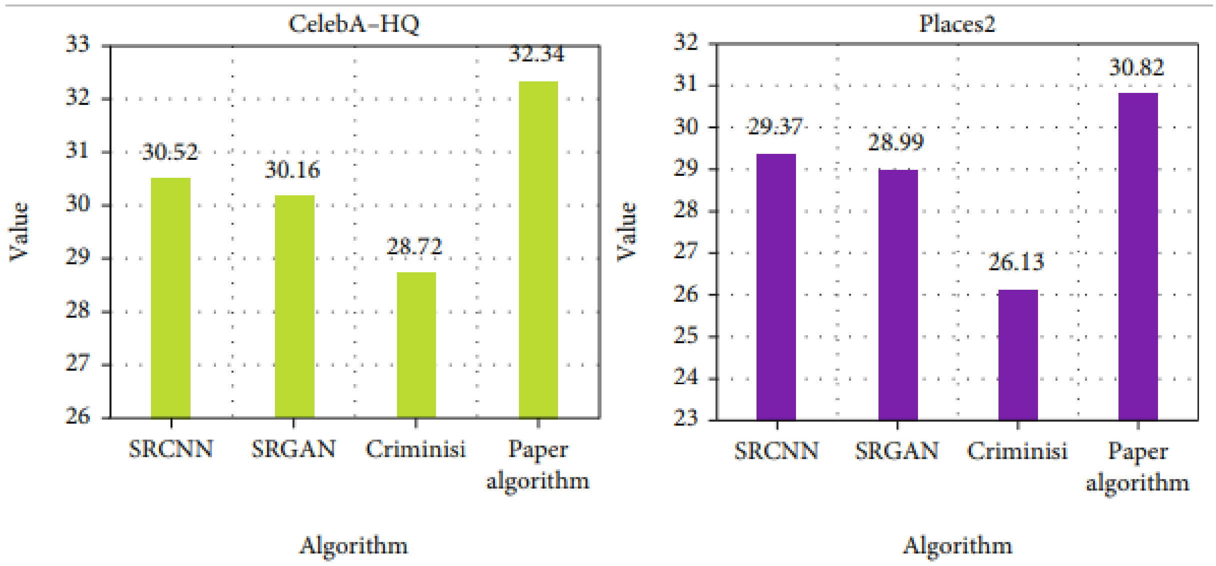


Figure 5. PSNR results of Blur Image Restoration Experiment

3.5.2. Analysis of SSIM Results in Blur Image Restoration Experiment

The figure 6 illustrates the results obtained by applying each algorithm on images. The SSIM value of paper algorithm on the CelebA-HQ dataset is 0.885. The SSIM value on the Places2 dataset is 0.879. Like PSNR, the algorithm with the smallest SIMM value is Criminisi. The SSIM values of this algorithm in the CelebA-HQ dataset and Places2 dataset are 0.849 and

0.831. Therefore, in the SSIM measurement, the value of paper algorithm is always greater than other algorithms. To sum up, paper algorithm is better than other algorithms for the structural similarity of the repaired image.

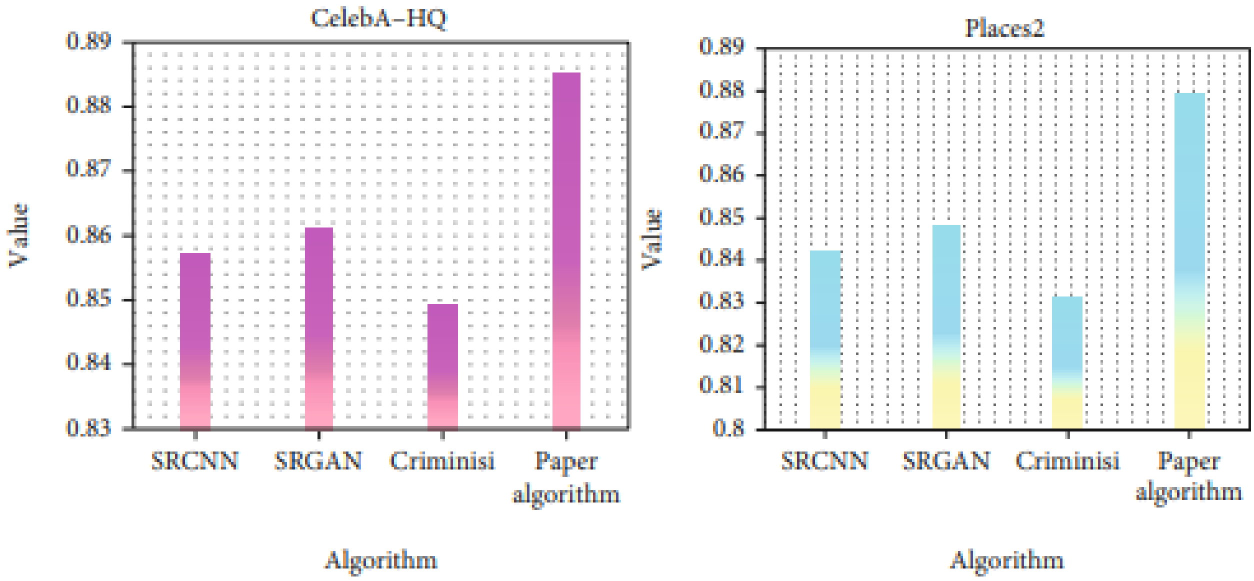


Figure 6. SSIM results of Blur Image Restoration Exp.

3.6. Damage Degree Image Restoration Experiment

For this section of the experiment, 100 pictures were chosen from CelebA - HQ and Places 2 datasets. However, they were processed with photoshop tools in order to simulate damage on the images and the results were divided into 5 groups each with 20 images. The 5 groups' damaged area of the image counts for 5%, 10%, 15%, 20%, and 25% of the entire image.

3.6.1. Analysis of PSNR Results in Damage Degree Image Restoration Experiment

The figure 7 illustrates the results obtained by applying each algorithm on images. Comparing the PSNR of each algorithm on the images, it can be noticed that the greater the degree of the damage on the picture, the more difficult it was to restore the photo. Even though the algorithm that was developed in this experiment is suggested to be the best option with average value 19.11, the other algorithms PSNR values are slightly close too. Hence, it is hard to determine which algorithm had the lowest PRSN, but according to the outcomes, the developed paper algorithm turned out to be the best.

3.6.2. Analysis of SSIM Results in Damage Image Restoration Experiment

The figure 8 illustrates the results obtained by applying each algorithm on images. Calculation of the average SSIM value of the paper algorithm is 0.776 under different damage degrees. The outcomes indicate that the SSIM value of the experiment's algorithm is the smallest by directed downwards and greater in a level. That is why the algorithm is shown to have the best repair effect.

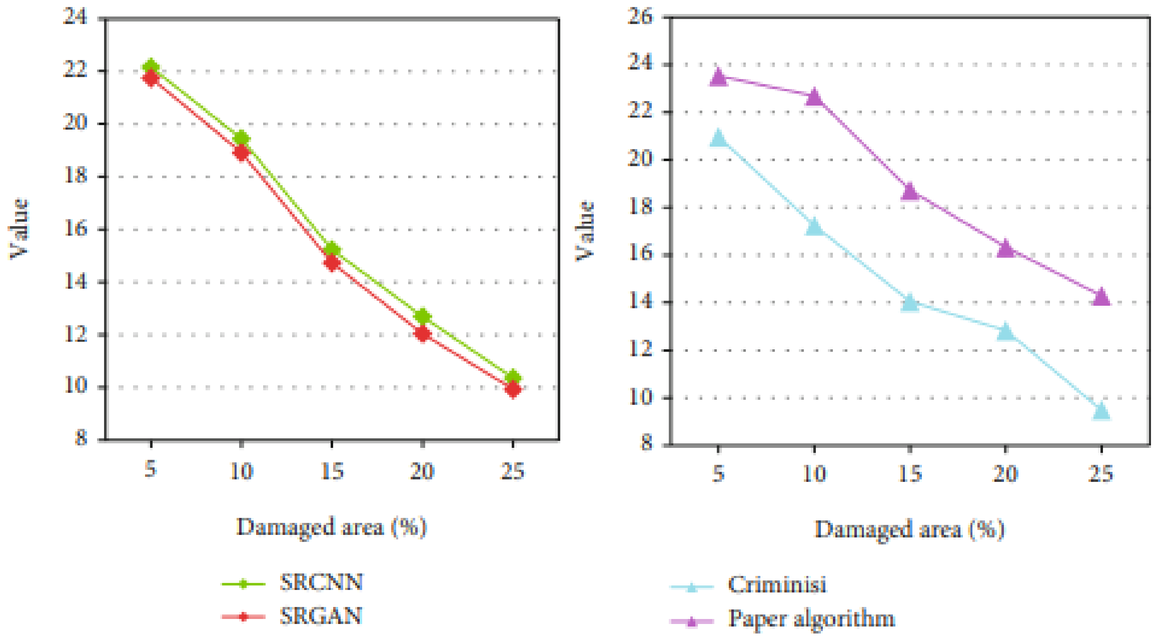


Figure 7. PSNR results of Damage Degree Image Restoration Experiment

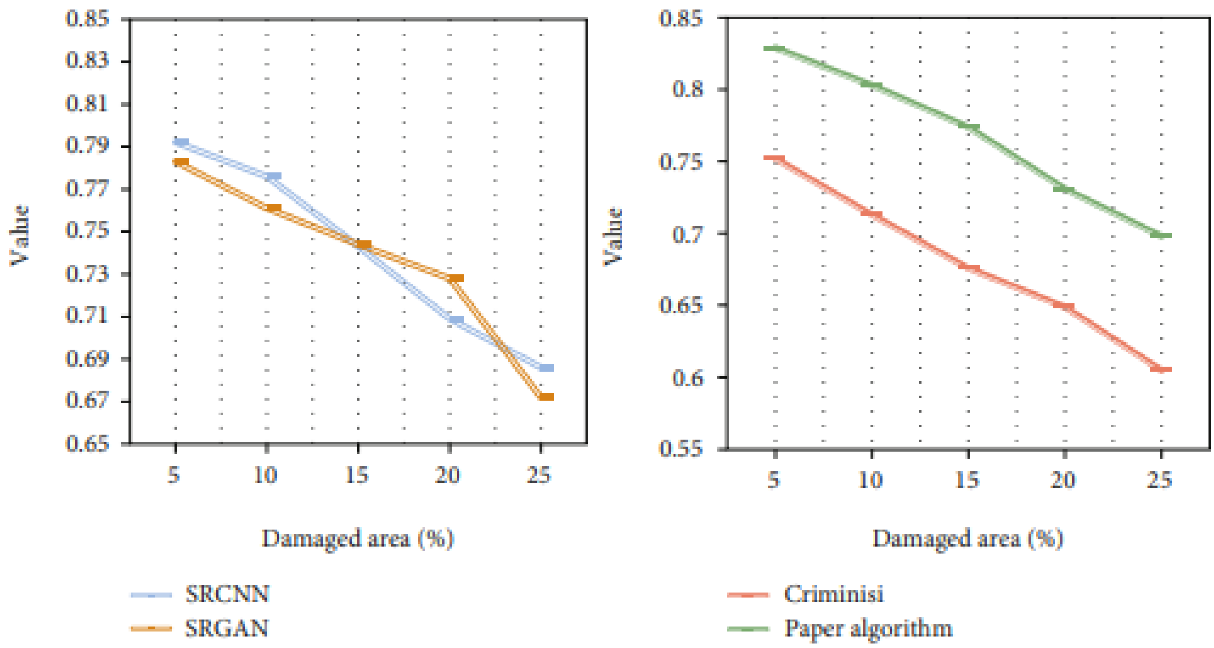


Figure 8. SSIM results of Damage Degree Image Restoration Experiment

4. Related Work and Review

Image restoration is a great technology that is still intensively being researched by many scholars. Research work published by H. Duan and X. Wang proposed a neural dynamics image restoration method which was named Gang method. It uses Echo State Network (EST) and obtained final results were compared with the state-of-the-art methods. Experiments showed that the proposed neutral dynamics' method performs better image recovery on fuzzy or on noisy instances [16]. Another research was introduced where a nonlocal (NL) extension to TV regularization uses pixel-level content adaptive distribution to model the sparsity of image gradients. Researches also used the NL similarity of natural images to achieve a more accurate estimation of the gradient. Even though the experimental results show that the proposed method can produce better objective and subjective image quality, the experimental method is more complicated and has less operability [17]. Y. T. Peng and P. C. Cosman proposed an underwater scene depth estimation method based on image blur and light absorption, which can be used to restore and enhance underwater images in image formation models. This method accurately estimates the depth of underwater scenes. The experimental results of restoration of real and synthetic underwater images indicates that the proposed method is more precise to other underwater image restoration methods based on IFM, but the disadvantage of this experiment is that the dimensions of the experimental test are less [18]. Xu et al. introduced a two-level domain decomposition method to directly solve the total variational minimization problem. The method is composed of overlapping domain decomposition technology and coarse grid correction, and an iterative algorithm forms a small sized and better condition system. Experiment results show that this method is fast and robust, especially effective on large-size images, but this method has some flaws in detail design [19]. Dong et al. introduced a "general model" which includes most of the existing wavelet frame-based models as special cases. Performed an asymptotic analysis of the general model when the image resolution reached infinity, thereby establishing a connection between the general model in the discrete environment and the new variable model in the continuum environment. However, the disadvantage of this study is that the established connection did not have convincing results [20].

5. Conclusion

Pictures taken many centuries ago are often blurred and damaged. Proposing a suitable and effective restoration method, helps to renew these images and influences to design, and apply new tools in this area. The development of computer vision technology in digital image processing and in artificial intelligence, influenced image restoration to become one of the important research topics in the field of computer vision. The experimetns and researach results have shown that emage enhancement, its image resolution compared to image restoration have completely different concept. Image super-resolution aims to recover natural and realistic textures for a high-resolution image from its degraded low-resolution counterpart [21]. The purpose of this work is to introduce basic understanding of the selected topic, its algorithms and results that were obtained in the experiment in order to test the differences and their functionality. Even with the final results, the process of the work showed that algorithms still needed improvement in repairing the damages.

6. References

[1] A. Ghosh, D. Chakraborty and A. Law, "Artificial Intelligence in Internet of Things," IET Research Journals, vol. 3, issue 4, pp. 1-11, 2015.

[2] John F. Gilmore, "Artificial Intelligence in Image Processingn," SPIE Digital Image Processing, vol. 528, pp. 192-201, 1985.

[3] S. Ding, S. Qu, Y. Xi, and S. Wan, "Stimulus-driven and concept-driven analysis for image caption generation," Neurocomputing, vol. 398, pp. 520–530, 2020.

[4] O. I. Khalaf, C. A. T. Romero, A. A. J. Pazhani, and G. Vinuja, “VLSI implementation of a high-performance nonlinear image scaling algorithm,” Journal of Healthcare Engineering, vol. 2021, Article ID 6297856, 10 pages, 2021.

[5] R. Surendran, O. I. Khalaf, and C. Andres, “Deep learning based intelligent industrial fault diagnosis model,” CMC-Computers, Materials & Continua, vol. 70, no. 3, pp. 6323–6338, 2022.

[6] S. Wan and S. Goudos, “Faster R-CNN for multi-class fruit detection using a robotic vision system,” Computer Networks, vol. 168, p. 107036, 2020.

[7] H. Zhu, H. Wei, B. Li, X. Yuan, and N. Kehtarnavaz, “Realtime moving object detection in high-resolution video sensing,” Sensors, vol. 20, no. 12, p. 3591, 20

[8] Z. Lei and W. Zuo, “Image restoration: from sparse and low-rank priors to deep priors [lecture notes],”IEEE Signal Process-ing Magazine, vol. 34, no. 5, pp. 172–179, 2017.

[9] Z. Zha, X. Zhang, Y. Wu et al., “Non-convex weighted ℓ_p nuclear norm based ADMM framework for image restoration,” Neurocomputing, vol. 311, pp. 209–224, 2018.

[10] H. Wei and N. Kehtarnavaz, Determining number of speakers from single microphone speech signals by multi-label convolutional neural network, IEEE IECON, 2018.

[11] Y. Zhao, H. Li, S. Wan et al., “Knowledge-aided convolutional neural network for small organ segmentation,” IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 4, pp. 1363– 1373, 2019.

[12] H. Lu, S. Li, Q. Liu, and M. Zhang, “MF-LRTC: multi-filters guided low-rank tensor coding for image restoration,” Neurocomputing, vol. 303, pp. 88–102, 2018.

[13] S. Wan, L. Qi, X. Xu, C. Tong, and Z. Gu, “Deep learning models for real-time human activity recognition with smartphones,” Mobile Networks and Applications, vol. 25, no. 2, pp. 743–755, 2020.

[14] M. G. Kim, H. Ko, and S. B. Pan, “A study on user recognition using 2d ecg based on ensemble of deep convolutional neural networks,” Journal of Ambient Intelligence & Humanized Computing, vol. 11, no. 5, pp. 1859–1867, 2020.

[15] F. Karami, K. Sadik, and L. Ziad, “A variable exponent nonlocal p (x) -Laplacian equation for image restoration,” Computers & Mathematics with Applications, vol. 75, no. 2, pp. 534–546, 2018.

[16] H. Duan and X. Wang, “Echo state networks with orthogonal pigeon-inspired optimization for image restoration,” IEEE Transactions on Neural Networks & Learning Systems, vol. 27, no. 11, pp. 2413–2425, 2016.

[17] H. Liu, R. Xiong, X. Zhang, Y. Zhang, S. Ma, and W. Gao, “Non-local gradient sparsity regularization for image restoration,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 27, no. 9, pp. 1909–1921, 2017.

[18] Y. T. Peng and P. C. Cosman, “Underwater image restoration based on image blurriness and light absorption,” IEEE Trans- actions on Image Processing, vol. 26, no. 4, pp. 1579–1594, 2017.

[19] M. G. Kim, H. Ko, and S. B. Pan, “A study on user recognition using 2d ecg based on ensemble of deep convolutional neural networks,” Journal of Ambient Intelligence & Humanized Computing, vol. 11, no. 5, pp. 1859–1867, 2020.

[20] B. Dong, Z. Shen, and P. Xie, "Image restoration: a general wavelet frame based model and its asymptotic analysis," SIAM Journal on Mathematical Analysis, vol. 49, no. 1, pp. 421–445, 2017.

[21] Michal Irani and Shmuel Peleg. Improving resolution by image registration. CVGIP, 53(3):231–239, 1991.