JOURNAL OF SOUTHERN COMMUNICATION ENGINEERING
ISLAMIC AZAD UNIVERSITY BUSHEHR BRANCH

E-ISSN: 2980-9231
https://jce.bushehr.iau.ir

https://doi.org/10.30495/jce.2025.1993480.1329

Vol. 14/ No. 54/Winter 2025

Research Article

Improving Super-Resolution in Face Images by Modeling Image Degradation Using Pairs of High-Quality and Low-Quality Images

Ahmad Dolatkhah, Assistant Professor [1*]

[1]Department of information and Communication, Amin Police University, Tehran, Iran
ict.ahmad93@gmail.com

Correspondence
Ahmad Dolatkhah, assistant professor, department of information and Communication, Amin Police University, Tehran, Iran, ict.ahmad93@gmail.com

Received: 17 February 2024
Revised: 14 April 2024
Accepted: 17 April 2024

Abstract

Improving image quality for identification and authentication in security and surveillance systems is of particular importance, and today, using artificial intelligence, the quality of images can be significantly improved. In this regard, the present paper, focusing on the details of face images, has improved the image failure detection model in the adversarial generator network, which led to a suitable performance in the meta-dissolving of face images. Most of the CNN networks that have been presented in recent years require a large set of images with appropriate annotations for proper performance, and they usually perform poorly in the case of degradation that have not been trained, which is addressed in this research to improve this challenge. In this work, pairs of high-quality and low-quality images are used to train the image degradation detection model; This information is then transferred to real images. The naturalness of the output images is one of the most important challenges in this field. The obtained results show that the criterion of perceptual similarity of the obtained image is equal to 38.4%, which is comparable to recent researches. As a result, using the proposed model, more natural images were produced.

Keywords: Quality of face image, Adversarial generative network, Super-resolution, Deep learning.

Highlights

- Improving super-resolution in face images by modeling image degradation using pairs of high-quality and low-quality images
- Improving image quality for identification and authentication in security and surveillance systems.
- Using SynNet and DegNet networks, the image damage detection model was improved and the image details were preserved.

Citation: A. Dolatkhah, "Improving Super-Resolution in Face Images by Modeling Image Degradation Using Pairs of High-Quality and Low-Quality Images," *Journal of Southern Communication Engineering*, vol. 14, no. 54, pp. 69–82, 2025, doi:10.30495/jce.2025.1993480.1329, [in Persian].

1. Introduction

Currently, hardware limitations and high costs associated with hardware improvements hinder the enhancement of image quality in imaging systems. Additionally, increasing quality always requires more storage space, which can itself become problematic[1].
High-quality and low-quality images can be distinguished using the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) metrics. Images of excellent quality should have high PSNR values, indicating a very high signal-to-noise ratio. Such images should exhibit fewer errors and less pixel difference compared to the reference (high-quality) image[2]. Quality images should have an SSIM close to 1, which indicates a high structural similarity with the reference quality images, meaning that these images share a significant amount of structural features and information with the high-quality image. Overall, these metrics serve as important tools for assessing image quality and can be utilized in the process of enhancing and improving images.
To enhance low-quality images, various methods exist. Some approaches utilize algorithms and image processing filters, such as contrast enhancement and color correction, to improve image quality. In cases where we are not dealing with complex noise, noise reduction filters can be employed to reduce noise levels and enhance image quality. Finally, when faced with complex artifacts in an image, image processing techniques such as super-resolution should be used, as these methods can add new information to the image using complex algorithms and improve its resolution. Therefore, to enhance images that suffer from complex artifacts, deep learning concepts and neural networks should be employed [3].
The use of neural networks to improve image quality has multiple reasons. Neural networks are well-equipped to recognize complex patterns and features in images. This capability allows them to identify and enhance errors and deficiencies present in images. Furthermore, given the vast amount of available image data, neural networks can be trained to improve image quality using rich and extensive training datasets. This can lead to better performance and higher quality in the final images. On the other hand, with advancements in technology, deep neural networks such as Generative Adversarial Networks (GANs) offer more capabilities for enhancing image quality. These networks are specifically used for generating high-quality images and improving the quality of real images. Deep learning is one of the topics that has been widely discussed in recent years, attracting the attention of many researchers in the field of artificial intelligence and even other areas. In fact, the year 1993 can be considered the year when the first successful convolutional neural network (CNN) and the LeNet network were introduced, and in 1997, recurrent neural networks (RNNs) were invented and introduced. However, in recent years, due to the expansion and advancements in the areas that will be mentioned, we have witnessed a flourishing discussion on deep learning[4-5].

2. Innovation and contributions

In this article, the main objective is to expand the ESRGAN network to reconstruct real images with low resolution by combining training pairs with a more practical degradation process. Real complex degradations are usually the result of intricate combinations of various degradation processes, such as camera imaging systems, image editing, and internet transmission. In other words, degradations are modeled with multiple iterative processes, each process serving as a classic degradation model. Empirically, a second-degree degradation modeling process has been used to achieve an appropriate balance between simplicity and efficiency [6].
The classic degradation model only includes a fixed process, while the image may contain complex degradations. Such a complex degradation process cannot be modeled with the initial classic model. Therefore, this research presents an optimized degradation model. This model consists of n iterative degradation processes, where each degradation process utilizes the classic degradation model with the same method but with different parameters. Recently, data-driven methods have been proposed to design a practical degradation model by constructing complex combinations, sampling, noise, and compression. Although these methods exhibit a wider variety of degradations and demonstrate significant generalization in managing LQ images in the real world, they still cannot cover some complex real degradations [7].
In this article, the main objective is to expand the ESRGAN network to reconstruct real images with low resolution by combining training pairs with a more practical degradation process. Real complex degradations are usually caused by intricate combinations of various degradation processes, such as camera imaging systems, image editing, and internet transmission. In other words, degradations are modeled with multiple iterative processes, each process serving as a classical degradation model. Empirically, a second-order degradation modeling process has been used to achieve an appropriate balance between simplicity and efficiency. This article employs a U-Net discriminator with spectral normalization (SN). The U-Net recognizes the true values for each pixel and can provide fine feedback to the generator. At the same time, the U-Net structure and complex degradations also increase training instability. In this design, spectral normalization has been utilized to stabilize training dynamics.

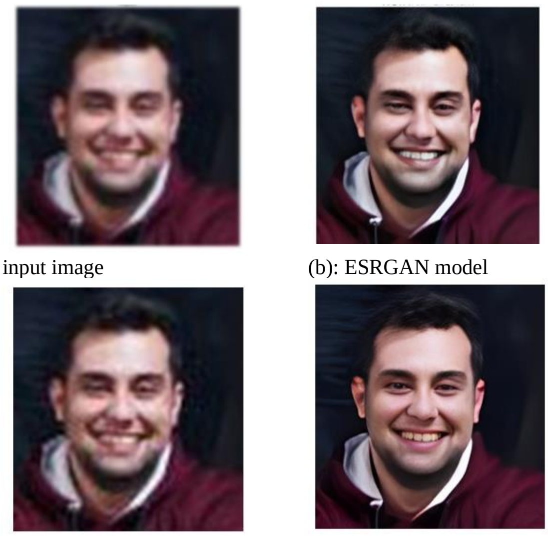
3. Materials and Methods

In this project, an initial effort has been made to create degradation in images, which includes learning the actual degradation from pairs of high-quality and low-quality images using DegNet, and transferring this learning to HQ images to generate their realistic LQ samples with SynNet. In the first case, instead of using a single LQ image to predict its degradation parameters, real LQ images and their synthetic HQ counterparts are used as input to create the degradation. To separate the image content and the type of degradation, a framework has been designed that predicts degradation accurately through multiple connected layers to produce convolution weights. This information is then transferred to natural images. Therefore, this project has two main subnetworks: DegNet for learning the degradation representation and SynNet for synthesizing low-quality images from high-quality input and the parameter Ω obtained from the previous stage. After receiving Ω, inspired by Style-GANs that control the style of the generated image with a vector in the W space, a similar structure has been chosen in this work to map the degradation representation Ω to the W space through several fully connected layers. Then, instead of introducing extensive noise in StyleGAN, the content of the SynNet image is provided by the features of the input HQ images. Finally, using the transformation D and the image content C, the degraded image is reconstructed with a modulated convolution operation (MC Block), where the degradation styles act as convolution weights to control the degradation process of the given input image. Figure 5 provides an overview of this model.

4. Results and Discussion

In this plan, the first step is to create distortions in the images. For this purpose, pairs of low-quality (LQ) and high-quality (HQ) images are used to learn the distortions. Instead of using a single LQ image to predict its distortion parameters, real LQ images and their corresponding artificial HQ images are used as input to generate the distortions. To separate the content of the image from the type of distortion, a neural network has been designed that accurately predicts the distortion through several connected layers to produce convolution weights. It should be noted that the proposed method may have limited effectiveness in non-facial scenarios, as the modeling has been performed for facial images.

In Figure 1, the outputs of the ESRGAN model, BSRGAN, and the proposed model are shown. The comparison metrics examined in this article include Peak Signal-to-Noise Ratio (PSNR), expressed in decibels, Structural Similarity (SSIM), Learned Perceptual Image Patch Similarity (LPIPS), and face verification accuracy in the dataset, expressed as a percentage.



(a): input image (b): ESRGAN model

(c): BSRGAN model (d): proposed method

Figure 1. (a) input and output comparison of (b) ESRGAN [8], (c) BSRGAN [9] and (d) proposed model

A quantitative comparison of these results is presented in Table 1. As can be seen, the proposed model has achieved a higher PSNR while maintaining the SSIM compared to recent studies. More importantly, the LPIPS metric has decreased, indicating that the proposed model offers greater perceptual similarity.

5. Conclusion

In this work, complex image defects of faces have been modeled, and pairs of high-quality and low-quality images were used as input for the image defect detection model. By doing this and utilizing the SynNet and DegNet networks, the image defect detection model in the generative adversarial network was improved, and image details were preserved. As a result, we observed satisfactory performance in terms of resolution. The results indicate that with the introduction of this model, the image parameters improved to PSNR=25.42 dB, SSIM=71.5%, and LPIPS=38.4%, showing a significant advancement compared to recent studies.

6. Acknowledgement

This research has received no external funding.

7. References

[1] P. Kaur and H. S. Pannu, “Comparative analysis of continuous and discrete orthogonal moments for face recognition,” *Proc. Int. Conf. Electron. Commun. Aerosp. Technol. ICECA*, vol. 2017-Janua, pp. 449–453, 2017, doi: 10.1109/ICECA.2017.8203724.

[2] N. Aloysius and M. Geetha, “A review on deep convolutional neural networks,” *Proc. IEEE Int. Conf. Commun. Signal Process. ICCSP*, vol. 2018-January, pp. 588–592, 2017, doi: 10.1109/ICCSP.2017.8286426.

[3] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” *Artif. Intell. Rev.*, vol. 53, no. 8, pp. 5455–5516, 2020, doi: 10.1007/s10462-020-09825-6.

[4] J. Kim, J. K. Lee, and K. M. Lee, “Accurate image super-resolution using very deep convolutional networks,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-December, pp. 1646–1654, 2016, doi: 10.1109/CVPR.2016.182.

[5] S. Guo, Z. Yan, K. Zhang, W. Zuo, and L. Zhang, “Toward convolutional blind denoising of real photographs,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2019-June, pp. 1712–1722, 2019, doi: 10.1109/CVPR.2019.00181.

[6] X. Ji, Y. Cao, Y. Tai, C. Wang, J. Li, and F. Huang, “Real-world super-resolution via kernel estimation and noise injection,” *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2020-June, pp. 1914–1923, 2020, doi: 10.1109/CVPRW50498.2020.00241.

[7] Z. Luo, Y. Huang, S. Li, L. Wang, and T. Tan, “Unfolding the alternating optimization for blind super resolution,” *Adv. Neural Inf. Process. Syst.*, vol. 2020-December, 2020.

[8] K. Zhang, J. Liang, L. Van Gool, and R. Timofte, “Designing a Practical Degradation Model for Deep Blind Image Super-Resolution,” *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 4771–4780, 2021, doi: 10.1109/ICCV48922.2021.00475.

[9] X. Li, C. Chen, S. Zhou, X. Lin, W. Zuo, and L. Zhang, “Blind Face Restoration via Deep Multi-scale Component Dictionaries,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12354 LNCS, pp. 399–415, 2020, doi: 10.1007/978-3-030-58545-7_23.

[10] Z. Liu et al., “Fine-Grained Face Swapping Via Regional GAN Inversion,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis*. Pattern Recognit., vol. 2023-June, pp. 8578–8587, 2023, doi: 10.1109/CVPR52729.2023.00829.

[11] A. Sargsyan, S. Navasardyan, X. Xu, and H. Shi, “MI-GAN: A Simple Baseline for Image Inpainting on Mobile Devices,” *Proc. IEEE Int. Conf. Comput. Vis*., pp. 7301–7311, 2023, doi: 10.1109/ICCV51070.2023.00674.

Appendix

Table 1: Comparison of the proposed model with recent studies

model	PSNR (dB)	SSIM (%)	LPIPS (%)	Face recognition accuracy in the dataset Chat cx(%)	reference
BSRGAN	24.48	71	39.1	98.7	[8]
ESRGAN	24.50	71.5	38.7	98.8	[9]
RGI	18.95	81	38.6	98.8	[10]
MI-GAN	24.65	72.5	39	98.7	[11]
Proposed model	25.42	71.5	38.4	98.8	This work

Declaration of Competing Interest: Authors do not have conflict of interest. The content of the paper is approved by the authors.

Publisher's Note: All content expressed in this article is solely that of the authors, and does not necessarily reflect the views of their affiliated organizations or the publisher, editors, and reviewers. Any content or product that may be reviewed and evaluated in this article is not guaranteed or endorsed by the publisher.

Author Contributions: All authors reviewed the manuscript.

Open Access: Journal of Southern Communication Engineering is an open access journal. All papers are immediately available to read and reuse upon publication.