Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 4 (2023)

Advancing Image Deblurring Performance with Combined Autoencoder and Customized Hidden Layers

Meysam Gorji[1], Elham Hafezieh, Ali Tavakoli
Mathematics department, University of Mazandaran, Babolsar, Iran

Abstract: This article introduces a novel approach to image deblurring by combining a Fourier autoencoder model. The proposed model effectively removes blur artifacts and restores image details by capturing frequency information using the Fourier Transform. In addition, the article presents a method to enhance deblurring by identifying optimal directions using an autoencoder model, trained on a dataset of blurry and sharp images to learn latent features for removing blur and restoring clarity. The encoded representations are used by the decoder to reconstruct a sharper version of the input image. A combination of two autoencoder models is employed, with a Convolutional Neural Network (CNN) handling the initial deblurring process and a fully connected model optimizing the deblurring parameters. This integrated approach leverages the strengths of CNNs in feature extraction and the flexibility of fully connected networks to produce higher quality, clearer images.

Keywords: Image deblurring, Fourier Transform, Autoencoder.

1. Introduction

Image deblurring is the process of restoring clear and sharp details from blurred images, and recent advancements have been made in this field [1]. Researchers have developed sophisticated algorithms and deep learning models for image deblurring, utilizing mathematical techniques, statistical models, and convolutional neural networks [2]. The goal of image deblurring is not only to remove blur but also to restore important visual information lost during the blurring process. Fourier transform is a widely used technique in image deblurring, allowing analysis of frequency components and addressing motion or defocus blur [3, 4]. However, Fourier transform-based methods have limitations, assuming linearity and shift-invariance that may not represent real-world scenarios accurately [5, 6]. Deep learning models, such as autoencoder deblurring, have been developed to handle complex scenarios [7, 8]. Autoencoder deblurring uses autoencoders with an encoder and decoder architecture to transform blurry images into latent representations and reconstruct deblurred images [9]. Training involves a large dataset of blurry images paired with sharp images to learn from examples and extract relevant features for generating deblurred outputs [10, 11]. The objective is to minimize the difference between predicted and ground truth sharp images through optimization, improving image quality and fine detail recovery. Autoencoder deblurring is versatile and can handle different blur types and complexities. It leverages the power of deep learning and data-driven techniques to generate high-quality deblurred images. Image deblurring is crucial in various fields such as photography, medical imaging, surveillance, and computer vision. Ongoing research and development in image deblurring techniques have the potential to further enhance image quality and clarity, benefiting industries and domains that rely on clear and detailed visual information [12].

Objectives

The primary objective of this article is to improve image deblurring by combining a Fourier transform with an autoencoder model. By training the autoencoder on a large dataset of blurry and sharp images, it learns to extract latent features that effectively remove blur and restore clarity. In this proposed approach, a combination of two autoencoder models is employed, with a Convolutional Neural Network (CNN) used for initial deblurring and a fully connected model used to optimize the deblurring parameters. The integrated approach of leveraging CNNs for feature extraction and fully connected networks for parameter optimization enhances the overall deblurring performance and produces higher quality, clearer images.

2. Methods

2.1 Motion blurred images

Motion blur in an image is caused by the movement of either the camera or the subject during the exposure time of the photograph. When the camera or the subject moves while the shutter is open, the result is a blurred image with streaks or smudges in the direction of the motion. In general, the appearance of motion blur can be described by utilizing the concepts of discrete Fourier transform and its inverse. In this section, we will provide an explanation of how a motion blurred image is perceived.

Let M and N be two integer numbers. The 2-D discrete Fourier transform of a function $f(x,y)$ denoted by $F(u,v)$ is defined as follows:

$$F(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\ e^{-j2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)}. \tag{1}$$

in which $u = 0,1,\dots,M-1, v = 0,1,\dots,N-1.$

Also, the inverse discrete Fourier transform of a function $F(u.v)$ is defined as:

$$f(x,y) = \frac{1}{MN}\sum_{u=0}^{M-1}\sum_{v=0}^{N-1} F(u,v)\ e^{j2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)}. \tag{2}$$

where $x = 0,1,\dots,M-1,\ y = 0,1,\dots,N-1.$

The effect of motion blur on an image can be simulated by convolving the original image $f(x,y)$ with $h(x,y)$, and adding noise $n(x,y)$ [13, 14]. This convolution process generates a blurred image, which results from the motion blur effect:

$$g(x,y) = h(x,y) * f(x,y) + n(x,y). \tag{3}$$

The function $h(x,y)$ is a filter that applies a convolution operation to the original image. It impacts the surrounding pixels of each individual pixel and results in a blurring effect on the pixel. The Fourier transform of (3) is as:

$$G(u,\mathrm{v}) = H(u,v)F(u,v) + N(u,v). \tag{4}$$

Suppose the image $f(x,y)$ is blurred using a uniform linear motion with motion values $\mathrm{x}_0(\mathrm{t})$ and $\mathrm{y}_0(\mathrm{t})$, representing motion in the x and y directions respectively. If T is the exposure duration, $H(u,v)$ can be obtained using the following equation:

$$H(u,v) = \int_0^T e^{-j2\pi(ux_0(t)+vy_0(t))}\ dt. \tag{5}$$

If the motion variables $x_0(t)$ and $y_0(t)$ are known, it is possible to directly obtain $H(u,v)$ from Equation (5). By substituting the given values of $x_0(t)$ and $y_0(t)$ into the equation, we can calculate the values of $H(u,v)$ for each u and v. Suppose a and b are unknown parameters. a and b represent the motion-blur parameters, which define the direction of motion blur in the image. We assume that the image is blurred by a uniform linear motion, occurring between the sensor and the scene during the imaging process. To model this, we can consider the functions $x_0(t) = \frac{at}{T}$ and $y_0(t) = \frac{bt}{T}$, where t represents the time of exposure and T is the total duration of the exposure. Based on this assumption, we can express Equation (5) as follows:

$$H(u,v) = \frac{T}{\pi(ua+vb)}\, sin\,[\pi(ua+vb)]\, e^{-j\pi(ua+vb)}. \tag{6}$$

Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 4 (2023)

Therefore, based on the blurred image generated by directions a and b, we address two main objectives in this work. First, we want to enhance image deblurring and then, identifying the optimal directions a and b in (6) that led to creation of blurriness. The aim is to develop methods that effectively reduce blurriness and restore the sharpness of the image. Additionally, the article focuses on identifying the most suitable directions a and b to achieve the desired deblurring results.

2.2 Proposed method

To achieve better image deblurring, this article proposes the utilization of an autoencoder model. The autoencoder is a deep learning architecture that consists of an encoder network and a decoder network. By training the autoencoder on a large dataset of blurry and corresponding sharp images, it learns to extract the latent features that can effectively remove blur and restore clarity.

The encoded representations capture important information about the image, enabling the decoder to reconstruct a sharper version of the input. Through this approach, the article aims to achieve superior deblurring performance compared to traditional methods.

In this approach, a combination of two autoencoder models is employed as visualized in Fig. 1. In the first model, a Convolutional Neural Network (CNN) architecture is applied to handle the initial deblurring process. The CNN is capable of learning and extracting complex spatial features from the input image, allowing it to effectively reduce the blur.

The CNN is trained to extract meaningful features from the input blurry image. It learns to identify and capture important patterns, edges, and textures that are relevant for understanding the content of the image. Once the CNN encoder has extracted these high-level features from the image, we move to the decoding stage. Here, instead of directly generating the deblurred image from the encoded features, we make use of Fourier transform, which is defined in (4). In the decoding stage, we use Fourier transform to reconstruct and correct these features, resulting in a higher-quality image.

In the fully connected model, we want to customize a hidden layer according to (6), to optimize and find the best directions a and b for image deblurring. By defining this custom hidden layer and feeding the encoded representations from the first model into the fully connected network, we can introduce additional trainable parameters to the model that specifically focus on determining the optimal values of a and b. This allows the network to learn and adapt to the specific characteristics of the input image, improving the deblurring performance. We can modify the fully connected model architecture to include this custom hidden layer, enabling the network to optimize and refine the motion-blur parameters for better image deblurring results.

In summary, the use of a CNN in the initial encoding stage and Fourier transform in the decoding stage allows for the extraction, reconstruction, and improvement of important features from the blurry image, ultimately enhancing its overall quality. Also, in the fully connected model the best directions a and b will be obtained.

Our model aims to achieve better PSNR and least Mean Square Error (MSE) which is defined as:

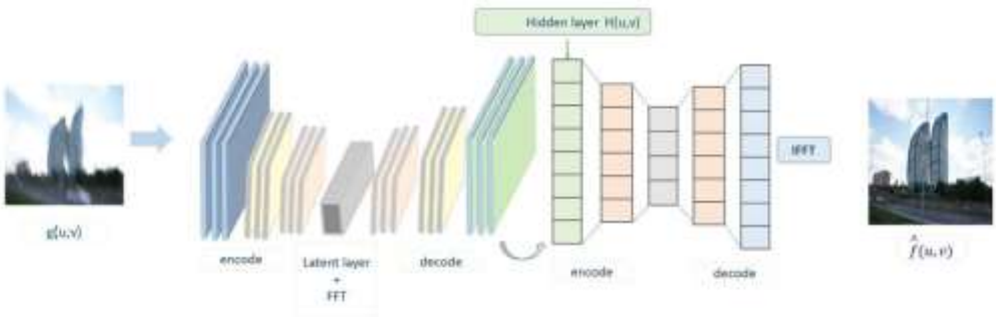

Fig. 1. The Hybrid Autoencoder Architecture: Merging CNN and Fully Connected Layers.

3 Dataset

The GoPro dataset is utilized for the research in this paper due to its widespread adoption in deblurring benchmarks. This involves using both the unaltered original data and modified sharp data. The dataset comprises 2103 high-resolution images sized at 720 × 1280, along with their corresponding images featuring simulated motion blur. The training set contains the sharp-blur pairs, while the testing set consists of 1111 images.

4 Results

Our motion deblurring method was implemented in python 3.10.4, and the experiments were conducted on a system with a Core i7 Intel(R) processor, 16 GB RAM, 3.2GHz, and GTX 3050 Ti GPU. The outcomes of our deep neural network are displayed in Table 1. To evaluate the performance of the proposed model, we compared two standard metrics, PSNR and SSIM, with various state-of-the-art models. In terms of efficiency, our model demonstrates favorable performance in comparison. As indicated in Table 1, the results reveal that our suggested method surpasses the previously examined techniques [15-18], according to the specified metrics. The highest SSIM and PSNR values for each measure in Table 1 are highlighted in bold. The data in Table 1 clearly demonstrates that our approach achieved the most notable SSIM and PSNR values. Despite being trained using the SSIM loss function, it is intriguing to note that the method actually performs better in terms of the PSNR metric. Figure 2 shows a comparison of the model's deblurring results with those of the compared methods on a randomly selected test sample. Regarding visual perception, the proposed method introduces two primary distinctions: Firstly, our recommended approach showcases enhanced capability in mitigating motion blur, especially in demanding situations like low-light conditions and camera shake in outdoor settings, as depicted in Figure 2; Secondly, the results obtained serve as proof that our framework effectively reduces object distortion and produces sharper edges compared to earlier models. Additionally, the test image that underwent restoration using the suggested method displayed a minor alteration in color compared to the original image. This implies that the method might benefit from additional training time or possibly require statistics from the test samples distinct from those in the training set. The SSIM and PSNR values point out that Deblur-GAN+ [15], SRN [16], and MBANet [17] are not as effective as the proposed and HWDCNN [18] models.

Table 1. The outcomes derived from evaluating the metrics of the motion deblurring models.

Methods	PNSR(dB)	SSIM
Deblur-GAN+ [8]	32.35	0.9582
SRN [9]	34.73	0.9414
MBANet [15]	35.33	0.9686
HWDCNN [1]	36.67	0.9696
Proposed	**37.45**	**0.9721**

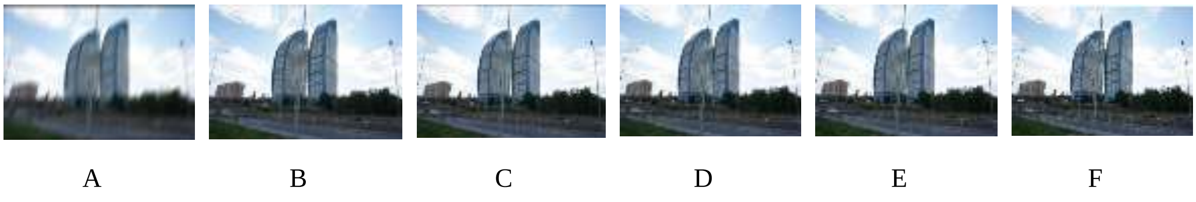

Fig. 2. Visual comparison of different deblurring models. A) blurred image, B) Deblur-GAN+ [15] C) SRN [16], D) MBANET [17], E) HWDCNN [18], F) Proposed method.

5 Conclusion

The purpose of this paper is to tackle the problem of image deblurring caused by motion blur. In this study, we propose a new approach that combines a Convolutional Neural Network (CNN) and a Fully Connected Neural Network to create an autoencoder model. Additionally, we introduce a novel hidden layer that determines the optimal directions responsible for the blurriness in the image. The results obtained from our method demonstrate superior performance compared to existing techniques when tested on the same dataset used for training.

Refrences

[1] Xu, J., Feng, A., Hao, Y., Zhang, X., & Han, Y. (2016). Image deblurring and denoising by an improved variational model. *AEU-International Journal of Electronics and Communications*, 70(9), 1128-1133.

[2] Lee, Y. (2023). Image deblurring method using denoising diffusion probabilistic model.

[3] Mao, X., Liu, Y., Liu, F., Li, Q., Shen, W., & Wang, Y. (2023, June). Intriguing findings of frequency selection for image deblurring. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 37, No. 2, pp. 1905-1913.

[4] Yae, S., & Ikehara, M. (2023). Inverted Residual Fourier Transformation for lightweight single image deblurring. *IEEE Access*, 11, 29175-29182.

[5] Fatima A, Younas I, Ali MW. An Overview on Recent Advances in Biosensor Technology and its Future Application. Arch Pharm Pract. 2022;13(1):5-10.

[6] Kopzhassarova U, Bexultanova L, Olkova I, Sarzhanova G, Belgibayeva G. Assessing the capabilities of CLIL technology in development students' foreign language professional skills. J Adv Pharm Educ Res. 2021;11(4):116-20.

[7] Lucy, L. B. (1974). An iterative technique for the rectification of observed distributions. *Astronomical Journal*, Vol. 79, p. 745 (1974), 79, 745.

[8] Richardson, W. H. (1972). Bayesian-based iterative method of image restoration. *JoSA*, *62*(1), 55-59.

[9] Nah, S., Hyun Kim, T., & Mu Lee, K. (2017). Deep multi-scale convolutional neural network for dynamic scene deblurring. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3883-3891.

[10] Taheri Mirghaed M, Karamaali M, Bahadori M, Abbasi M. Identification and Prioritization Technologies and Types of Threats in Future Warfare Using Future Studies Approach. Entomol Appl Sci Lett. 2022;9(1):7-19.

[11] ELGENDY TYAAA. Proposed Model for Selection of The Internal Auditor Using Analytical Network Process: Case Study. J Organ Behav Res. 2022;7(1):138-55.

[12] Heaton, J. (2018). Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning: The MIT Press, 2016, 800 pp, ISBN: 0262035618. Genetic programming and evolvable machines, 19(1-2), 305-307.

[13] Zhang, J., Zhang, C., Wang, J., Xiong, Q., Zhang, Y., & Zhang, W. (2021). Attention driven self-similarity capture for motion deblurring. In *2021 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1-6. IEEE.

[14] Zheng, S., Wu, Y., Jiang, S., Lu, C., & Gupta, G. (2021). Deblur-yolo: Real-time object detection with efficient blind motion deblurring. In *2021 International Joint Conference on Neural Networks (IJCNN)*, pp. 1-8. IEEE.

[15] Shao, W. Z., Liu, Y. Y., Ye, L. Y., Wang, L. Q., Ge, Q., Bao, B. K., & Li, H. B. (2020). DeblurGAN+: Revisiting blind motion deblurring using conditional adversarial networks. *Signal Processing*, 168, 107338.

[16] Tao, X., Gao, H., Shen, X., Wang, J., & Jia, J. (2018). Scale-recurrent network for deep image deblurring. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 8174-8182).

[17] Zhou, Q., Ding, M., & Zhang, X. (2020). Image Deblurring Using Multi-Stream Bottom-Top-Bottom Attention Network and Global Information-Based Fusion and Reconstruction Network. *Sensors*, 20(13), 3724.

[18] Grosu, G. F., & Terebes, R. (2022). Reinforced Hybrid Wiener Deconvolutional-Convolutional Autoencoders Based Image Deblurring. In *2022 International Symposium on Electronics and Telecommunications (ISETC)* (pp. 1-4). IEEE.