DEEP LEARNING APPLICATION FOR IMAGE ENHANCEMENT

A. Elaraby[1,a], I. Elansary[2], A. Nechaevskiy[3]

[1]*Department of Computer Science, Faculty of Computers and Information, South Valley University, Qena, 83523, Egypt*

[2]*Modern Academy for Computer Science and Management Technology, Cairo, Egypt*

[3]*Meshcheryakov Laboratory of Information Technologies, Joint Institute for Nuclear Research, 141980, Joliot-Curie 6, Dubna, Russia*

Email: [a] ahmed.elaraby@svu.edu.eg

Recently, deep learning has obtained a central position toward our daily life automation and delivered considerable improvements as compared to traditional algorithms of machine learning. Enhancing of image quality is a fundamental image processing task and. A high-quality image is always expected in several tasks of vision, and degradations like noise, blur, and low-resolution, are required to be removed. The deep techniques approaches can significantly and substantially boost performance compared with classical ones. One of the main research areas where deep learning can make a major impact is imaging. This work presents a survey of deep learning on image enhancement and describes its potential for future research.

Keywords: Machine Learning, Deep Learning, Image Enhancement.

Ahmed Elaraby, Ismail Elansary, Andrey Nechaevskiy

1. Introduction

Image quality enhancing is a fundamental problem in image processing that has received great attention over several decades. A high-quality image is always expected in different tasks of vision, and degradations like low-resolution, blur, and noise, are required to be removed. While the classical techniques for this task have achieved great progress, the recent top performer, deep techniques approaches, can significantly and substantially boost performance compared with classical ones. The advantages of deep learning techniques which enable it to realize such success are its high representational capacity and the strong nonlinearity of the approaches [1].

Image enhancement is adjusting process of digital images so that the results are more appropriate for display or additional image analysis [2-5]. Image enhancement methods divide to two main categories: "spatial domain" approaches and "frequency domain" approaches [1]. The term spatial domain in image processing indicates to an image plane itself, and methods in this category are depending to direct processing of image pixels. Spatial domain enhancement methods also divide to two categories: the spatial domain filtering category and transformation category. The former is depending on neighborhoods and the latter is depending on individual pixels. The spatial domain transformation approaches in common use are basic on histogram processing and gray level transformation. The frequency domain processing approaches are depending on modifying an image Fourier transforms. Let $f(a,b)$ be the original image and its Fourier transform $F(x,y)$, while $h(a,b)$ be a filter and $H(x,y)$ its Fourier transform, and then$f(a,b)$ is transformed to $g(a,b)$ after convolving with $h(a,b)$ and $G(x,y)$ is Fourier transform of $g(a,b)$.

The procedure can be written in the spatial domain as:

$$g(a,b) = f(a,b) * h(a,b) \qquad (1)$$

The frequency domain as:

$$G(x,y) = F(x,y) \cdot H(x,y) \qquad (2)$$

Where "$*$"is convolution operator and "$\cdot$"means multiplication operator. It is clear from Eqs. (1,2) that if $h(a,b)$ is selected correctly, the image $f(a,b)$ will then be influentially enhanced.

The rest of this paper will explore the development of advanced deep approaches for image enhancement by researching several fundamental issues with various motivations.

2. Deep Learning Applications on Image Enhancement

In image denoising, the basic concept of learning in image denoising is to use training data to refine a module of the designed model. Dictionary learning is an example of this; image patches can be interpreted using a coefficients sequence on a collection of bases, resulting in a redundant dictionary. The dictionary is supposed to express the general structures of natural images in such a way that clean image patches are well estimated on it for the denoising task. As a result, as in KSVD [6], LSSC [7], and CSR [8], the dictionary is an essential item that is learned from a collection of high-quality training data or the to-be-processed degraded image. Deep learning is another example, in which the goal is to learn a discriminative restoration function. The first attempts to use (CNNs) and stacked auto-encoders to denoising natural images [9] are promising, demonstrating that these deep models can work like or better than traditional wavelet or Markov random field-based denoising approaches.

Agostinelli et al. [10] investigated for building multi-column stacked auto encoders to deal with different forms of noise. DnCNN [11] used the residual learning technique to train deep CNN models and achieved state-of-the-art results. These studies show that deep architecture, strong learning, and high representational ability can significantly improve image denoising efficiency. Example of image denoising result using deep learning is shown in figure 1.

Figure 1. Image denoising applying neural network-based models [9]

In image deblurring, the most critical issue is ill-posedness. The observed fuzzy images do not stably and uniquely evaluate as sharp images in the non-blind case, owing to the blur operator's ill-conditioned existence [11]. Researchers have been concentrating on creating new strategies and prototypes, as well as improving the efficiency of optimization techniques, to deal with ill-posedness. For the task of deblurring, learning-based approaches are proposed. The motives in this case can be divided into two categories. In the first group, the proposed methods [11-13] aim to learn a subspace in which the sharp picture can be found. A subspace can be created by extracting local patterns from multiple sharp images, where the sharp images and target images share similar information, allowing for an accurate representation of details in the target image. In the second group, the aim is to find a restoring mechanism that can transform a fuzzy target image into a sharp target image [13]. Multiple sharp images with their fuzzy equivalents are widely used to train the restoring function parameters in this scenario. The "Regression Tree Field" (RTF) was used by "Schmidt" et al. [14] to model a nonlinear regressor that defines the parameters of local deblurring. Example of image deblurring result using deep learning is shown in figure 2.

Figure 2. Estimation of image deblurring utilizing explicit deconvolution [15]

In Image Super-resolution, the aim of super-resolution is to create a high-resolution image from one or more low-resolution images, recapturing high-frequency information that were lost during the imaging process. A belief network is a type of learning technique that can be expressed in terms of a Markov network. The images are analyzed with patch representation using a Markov network [16]. An observation function connects the low-resolution patches and their corresponding high-resolution patches, defining how well one fits the other. A transition function connects the neighbor patches in the reconstructed image. The model uses the belief propagation algorithm to restore the high-resolution image after the parameters of the functions have been well trained.

High-resolution images form a manifold with similar local geometry to that provided by low-resolution images, according to manifold learning-based approaches [17]. In order to do restoration, the relationship between points on the low-resolution manifold can be directly applied to the high-resolution manifold. However, the assumption of identical manifolds is overly restrictive, and it cannot be fulfilled in many situations. To address this issue, it is proposed that two explicit mapping functions be learned to find a common manifold for low and high-resolution images [18]. The idea behind the sparse coding-based approaches [19-20] is that a signal can be interpreted by a sparse code on an over-complete dictionary. The linear relationship between the sparse codes of low and high-resolution images can be recovered in the super-resolution task, resulting in encouraging restoration results. A high-resolution image can be obtained from low-resolution patches by dividing images into patches. The mapping function types like simple functions, support vector regression, and anchored neighborhood regression [21] are indeed shallow types with limited representational capacity. It has been shown that stacking shallow types into a deep one can significantly improve the performance of super-resolution, which is related to deep learning. Kim et al. conducted research into the use of very deep convolutional networks for SISR [22], who suggested residual learning to train a 20-layer CNN model and achieved high efficiency. They also presented a profoundly recursive convolutional network [23] that uses a limited number of model parameters to allow for long-range pixel dependencies. In the SR mission, use self-examples on different scales to fine-tune a pre-trained convolutional auto-encoder [24]. Shi et al. and Dong et al. [25-26], constructed networks in which the majority of the computation occurs in the LR space and the up-scaling process occurs only in the last layer of the networks to speed up the SR computation. Example of image super-resolution result using deep learning is shown in figure 3.

Figure 3. Obtaining a high-resolution image [26]

Besides the above CNN-related works, the studies on other feed-forward neural networks were conducted, such auto-encoders [27-28] and sparse coding-based networks [29-30]. An LR image patch can be generated from various HR image patches which reside in a low-dimensional natural image ramified. The MSE-based solution is a pixel-wise average of the possible HR patches on the ramified, thus exhibiting blurry or over-smoothing effects, and lacking high-frequency details. To address this task, in [31] presented a perceptual loss which measures the MSE between the VGG feature maps of the SR result and the ground truth. In [32], by incorporating this loss and an adversarial loss, Ledig et al. developed a generative adversarial network which can recover photo-realistic textures. Using similar ideas, inference models can be applied to the statistics of CNN feature maps, such that the statistics of SR solutions and that of natural images are as close as possible [32].

3. Conclusion

In this paper we highlighted several aspects of image enhancement. Providing a compendium of the advances made in the new and exciting sub-field of deep learning for image enhancement. The paper discusses image enhancement categories in spatial domain and frequency domain. Deep learning technology and background is discussed. Furthermore, applications of deep learning on image enhancement are analysis in most important tasks like image super-resolution, deblurring and denoising.

References

[1] R. C. Gonzalez and R. E. Woods, Digital Image Processing, (3rd Edition). Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006.

[2] A. Elaraby; D.Moratal: A Generalized Entropy-Based Two-Phase Threshold Algorithm for Noisy Medical Image Edge Detection, Scientia Iranica, Vol. 24; 6, 2017.

[3] A. Elaraby; et al: New Algorithm For Edge Detection in Medical Images Based on Minimum Cross Entropy Thresholding, IJCSI, Vol. 11; 2, 2014.

[4] A. Elaraby; et al: New Algorithm For Edge Detection Based on Exponential Entropy, (ISCYR), 29-30 April 2014, Assiut University, Egypt.

[5] A. Elaraby, A. Nechaevskiy: An effective segmentation approach for liver computed tomography scans using fuzzy exponential entropy // Computer Research and Modeling, 2021, 13(1), pp. 195–202.

[6] W. Dong, X. Li, L. Zhang, and G. Shi, "Sparsity-based image denoising via dictionary learning and structural clustering," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2011, pp. 457–464.

[7] V. Jain and S. Seung, "Natural image denoising with convolutional networks," in NIPS, 2009, pp. 769–776.

[8] J. Xie, L. Xu, and E. Chen, "Image denoising and inpainting with deep neural networks," in NIPS, 2012, pp. 341–349.

[9] H. C. Burger, C. J. Schuler, and S. Harmeling, "Image denoising: Can plain neural networks compete with bm3d?" in CVPR, 2012, pp. 2392–2399.

[10] H. C. Burger, C. Schuler, and S. Harmeling, "Learning how to combine internal and external denoising methods," in GCPR, vol. 8142, 2013, p. 121.

[11] F. Agostinelli, M. R. Anderson, and H. Lee, "Adaptive multi-column deep neural networks with application to robust image denoising," in NIPS, 2013, pp. 1493–1501.

[12] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," IEEE Trans. Image Process, 2017.

[13] N. Joshi, W. Matusik, E. H. Adelson, and D. J. Kriegman, "Personal photo enhancement using example images," ACM Transactions on Graph, vol. 29, no. 2, p. 12, 2010.

[14] J. Ni, P. Turaga, V. M. Patel, and R. Chellappa, "Example-driven manifold priors for image deconvolution," IEEE Trans. Image Process, vol. 20, no. 11, pp. 3086–3096, 2011. 23, 24.

[15] C. J. Schuler, H. C. Burger, S. Harmeling, and B. Sch¨olkopf, "A machine learning approach for non-blind image deconvolution," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 1067–1074.

[16] W. T. Freeman and E. C. Pasztor, "Learning to estimate scenes from images," Proceedings of Advances in Neural Information Processing Systems, pp. 775–781, 1999.

[17] M. Bevilacqua, A. Roumy, C. Guillemot, and M.-L. A. Morel, "Neighbor embedding based single-image super-resolution using semi-nonnegative matrix factorization," in Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2012, pp. 1289–1292.

[18] B. Li, H. Chang, S. Shan, and X. Chen, "Low-resolution face recognition via coupled locality preserving mappings," IEEE Signal Process. Lett., vol. 17, no. 1, pp. 20–23, 2010.

[19] X. Gao, K. Zhang, D. Tao, and X. Li, "Image super-resolution with sparse neighbor embedding," IEEE Trans. Image Process, vol. 21, no. 7, pp. 3194–3205, 2012.

[20] L. He, H. Qi, and R. Zaretzki, "Beta process joint dictionary learning for coupled feature spaces with application to single image super-resolution," in CVPR, 2013, pp. 345–352.

[21] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in ECCV, 2014, pp. 184–199.

[22] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in CVPR, 2016.

[23] J. Kim, J. K. Lee, and K. M. Lee "Deeply-recursive convolutional network for image super-resolution," in CVPR, 2016.

[24] Z. Wang, Y. Yang, Z. Wang, S. Chang, W. Han, J. Yang, and T. Huang, "Self-tuned deep super resolution," in CVPR Workshops, 2015, pp. 1–8.

[25] W. Shi, J. Caballero, F. Husz´ar, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in CVPR, 2016, pp. 1874–1883.

[26] C. Dong, C. C. Loy, and X. Tang, "Accelerating the super-resolution convolutional neural network," in ECCV. Springer, 2016, pp. 391–407. 21, 85, 94

[27] Z. Cui, H. Chang, S. Shan, B. Zhong, and X. Chen, "Deep network cascade for image super-resolution," in ECCV, 2014, pp. 49–64.

[28] R. Wang and D. Tao, "Non-local auto-encoder with collaborative stabilization for image restoration," IEEE Trans. Image Process, vol. 25, no. 5, pp. 2117–2129, 2016.

[29] K. Zeng, J. Yu, R. Wang, C. Li, and D. Tao, "Coupled deep auto-encoder for single image super-resolution," IEEE Trans. Cybern., no. 99, pp. 1–11, 2016.

[30] Z. Wang, D. Liu, J. Yang, W. Han, and T. Huang, "Deep networks for image super-resolution with sparse prior," in ICCV, 2015, pp. 370–378.

[31] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," ECCV, 2016.

[32] J. Bruna, P. Sprechmann, and Y. LeCun, "Super-resolution with deep convolutional sufficient statistics," ICLR, 2016.