JOURNAL OF CRITICAL REVIEWS

ISSN- 2394-5125 VOL 8, ISSUE 02, 2021

Quality Enhancement of Hybrid Bayesian Regularization Neural Network for Image Restoration

Sukanth Behera[1], Saradiya Kishore Parija[1]

[1]Assistant Professor, Department of ECE, Gandhi Institute for Technology (GIFT), Bhubaneswar, India

Abstract

Neural networks have shown very promising results for various image restoration (IR) tasks. However, the design of network architectures remains a major challenging for achieving further improvements. While most existing DNN-based methods solve the IR problems by directly mapping low quality images to desirable high-quality images, the observation models characterizing the image degradation processes have been largely ignored. In this paper, we first propose a denoising-based Hybrid Bayesian Regularization IR algorithm, whose iterative steps can be computed efficiently. Then, the iterative process is unfolded into a Hybrid Bayesian Regularization neural network(HBRNN), which is composed of multiple denoisers modules interleaved with back-projection (BP) modules that ensure the observation consistencies. As such, the proposed network not only exploits the powerful denoising ability over DNNs, but also leverages the prior of the observation model. Experimental results prove that our model greatly outperforms numerous state-of-the-art restoration methods in terms of both peak signal-to-noise ratio (PSNR) and structure similarity index metrics (SSIM).

Index Terms— HBRNN, image denoising, image super resolution, JPEG image deblocking, high-quality.

I. INTRODUCTION

Image compression is a task to convert images into small footprint. Lossy image compression aims a higher compression ratio while allows some distortion of decompressed images. Given specific compression ratio, lossy image compression algorithms are designed to minimize reconstruction distortion in terms of peak signal-to-noise ratio (PSNR), structural similarity (SSIM) or other metrics. Traditionally, lossy image compression algorithms (e.g. JPEG and BPG) are based on block-wise discrete cosine transform (DCT), quantization and entropy coding. The DCT and quantization steps introduce losses and distortions (e.g. blocking, blurring and ringing). The DCT is based on the spacial continuity of image signals, however, ignores the prior distribution of photographs. Recently, learned image compression algorithms with deep networks are developed and achieve better image quality than traditional approach. Many of the deep networks based algorithms utilize deep model to transform images instead of DCT. Comparing to DCT transformed representations, the deep networks encoded features are not orthogonal between dimensions, and not discriminative for low and high frequencies, which make it more difficult to ignore high frequency redundancies to further minimize features entropy. Mentzer, et al. [1] proposed to use context models (PixelCNN [2]) as entropy coder and achieved better results than others. In this work, we investigate a simple pipeline cascading BPG image compression and deep networks based image restoration. The deep networks for image restoration are supposed to learn the prior distribution of images, so that can enhance BPG-compressed images both quantitatively in term of peak signal-to-noise ratio (PSNR), and qualitatively for human perception. Deep neural networks are widely used for low-level image restoration problems recently. Deep networks based restoration for compressed images is firstly introduced by Dong et al. [4]. The work is inspired by SuperResolution Convolutional Neural Network (SRCNN)

ISSN- 2394-5125 VOL 8, ISSUE 02, 2021

[5], in which convolutional networks show potential in low level vision tasks. Since then, several follow-up work [3, 7] further improve the power of deep networks to remove artifacts. Recently, the Enhanced Deep residual networks for Super-Resolution (EDSR) [10] achieves significant improvement for image super-resolution. In EDSR, the deep networks consist of multiple blocks with linear residuals. The residual blocks have 2 convolutional layers connected with ReLU activation. We further improve the EDSR with wide activation SR networks (WDSR) (Fig. 1) from three aspects: wide activation, weight normalization in training and simplified global residual pathway. The WDSR are more effective in term of PSNR for image super-resolution. In this work for image restoration, we adopt the WDSR by removing up-sampling pixel-shuffle layers in the final stage.

In this paper, we propose a Hybrid Bayesian Regularization neural network to take advantages of both the optimization- and discriminative learning-based IR methods. First, we propose a denoising-based IR method, whose iterative process can be efficiently carried out. Then, we unfold the iterative process into a feed-forward neural network, whose layers mimic the process flow of the proposed denoising-based IR algorithm. Moreover, an effective Bayesiandenoiser that can exploit the multi-scale redundancies is proposed and plugged into the deep network.

2. METHODOLOGY

In proposed system, an adaptive residual network (ARN) for high quality image restoration is proposed. ARN builds very accurate mappings between the degraded images and their corresponding clean ones. ARN give great performance in image restoration. ARN can handle many image restoration applications, such as Gaussian image denoising, single image super resolution, image deblocking, and so on at the same time. Compared to existing networks, which could only solve one or two problems, our ARN are applicable in different problems by just changing the corresponding training datasets. ARN is made up of one 3*3 convolutional layer, six ARUs, and one 3*3 convolutional layer sequentially.

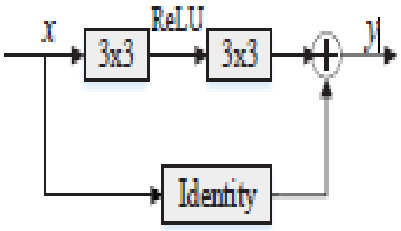

Fig.1 Residual Unit

It will be more difficult to train such a deep network because of vanishing gradients. the residual building block is made of two 3*3 convolutional layers, the corresponding ReLU, and one identity skip connection. Obviously, such shortcuts connections introduce neither extra parameter nor computation complexity. Shortcut connects the original input with the following stacked layers, rather than linking the outputs of the previous layers to the current one. This matters a lot for reserving the useful information of input. Thirdly, we introduce and to smartly balance the importance between the original input and the output of previous layer.

The first 3*3 convolutional layer is used to extract the features from the input low-quality images. Then, the extracted features will be sent to the six ARUs. After that, the last 3*3 convolutional layer works as reconstructing high-quality images Then, the extracted features will be sent to the six ARUs. After that, the last 3*3 convolutional layer works as reconstructing high-quality images. Each ARU

ISSN- 2394-5125 VOL 8, ISSUE 02, 2021

has two 3*3 convolutional layers, and for example, (64,128) in the 3rd ARU represents there are 64 channels in the first 3*3 convolutional layer and 128 channels in the second 3*3 convolutional layer. Facts proved that 3*3 convolutional layers have sufficient ability to extract good features as long as the network is deep enough. Furthermore, they will produce fewer parameters than larger convolutional kernels, which will reduce computation complexity greatly.

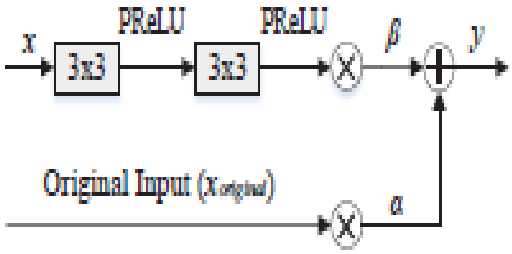

Fig.2 Adaptive Residual Unit (Aru)

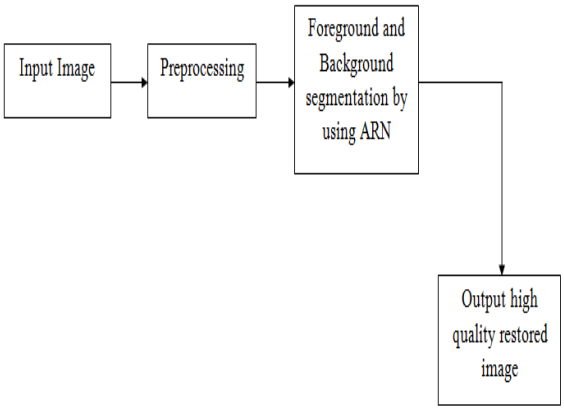

Fig.3 ARN Block Diagram

Here, adopt MSE, the most common method to measure the differences between two images, as the loss function in our model. Learning the end-to-end function F from low-quality images to its high-quality counterparts needs to estimate the weights represented by the convolutional kernels as well as the scaling parameters of the shortcuts. This is achieved by minimizing the MSE between the outputs of the network and the original image.ReLU has been a common activation function used in deep learning after traditional sigmoid-like units. For ReLU, if the input is less than zero, the result is zero, which is very helpful to generate a sparse representation. For neural networks, the values of the activation layers can be viewed as a sparse representation of the input. Hence, ReLU can accelerate the process of network convergence and achieve better solutions. This is why an increasing number of researchers prefer to using ReLU other than conventional sigmoids. The goal of image denoising is to recover the latent clean image from its corrupted version.

HBRNN

For solving the denoising-based IR methods, based on which a HBRNN. Considering the denoising-based IR problem, we adopt the half-quadratic splitting method, by which the equally constrained optimization problem can be converted into a nonconstrained optimization problem.

ISSN- 2394-5125 VOL 8, ISSUE 02, 2021

The x-subproblem is a quadratic optimization problem that can be solved in closed-form, as x (t+1) = W−1 b, where W is a matrix related to the degradation matrix A. Generally, W is very large, so it is impossible to compute its inverse matrix. Instead, the iterative classic conjugate gradient (CG) algorithm can be used to compute x (t+1), which requires many iterations for computing x (t+1). In this paper, instead of solving for an exact solution of the x-subproblem, we propose to compute x (t+1) with a single step of gradient descent for an inexact solution.

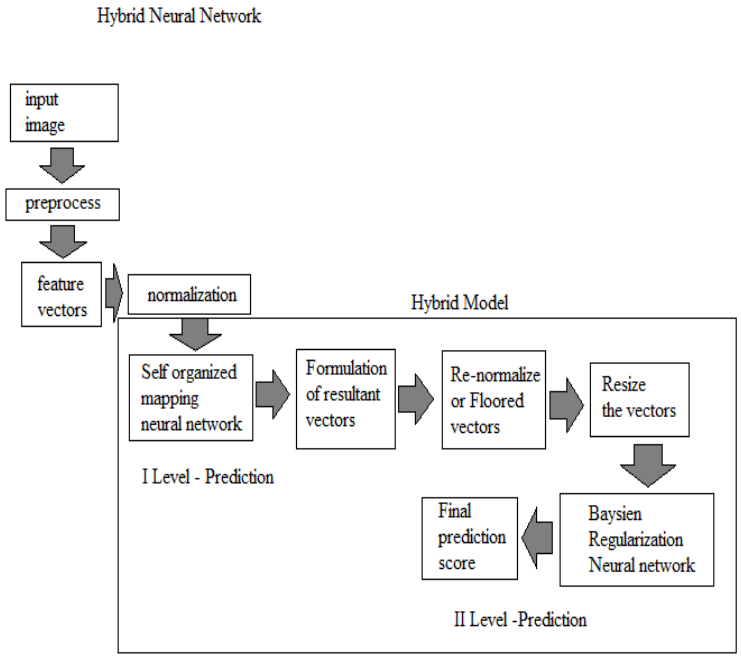

Fig.4 HBRNN Block Diagram

The input degraded image y ∈ R ny first goes through a linear layer parameterized by the degradation matrix A ∈ R ny×mx for an initial estimate x (0) . x (0) is then fed into the linear layer parameterized by matrix A^{-} ∈ R mx×mx , whose output is added with x (0) weighted by δ1 via a shortcut connection. The updated x (1) is fed into the denoiser module, whose structure is shown in Fig4. The denoised signal v (1) is fed into the linear layer parameterized by A^{-}, whose output is further added with x (0) and v (1) via two shortcut connections for the updated x (2) . Such a process is repeated K times. In our implementation, K = 6 was always used. Instead of using fixed weights, all the weights (δ1, δk,1, δk,2, k = 1, 2, · · · , K) involved in the K recurrent stages can be discriminatively learned through endto-end training. Regarding the denoising module, as we are using a BAYSIEN-based denoiser that contains a large number of parameters, we enforce all the denoising modules to share the same parameters to avoid over-fitting.

Inspired by the recent advances on semantical segmentation and object segmentation, the architecture of the denoising network is illustrated in Fig.4. Similar to the U-net and the sharpMask net, the proposed network contains two parts: the feature encoding and the decoding parts. In the feature encoding part, there are a series of convolutional layers followed by pooling layers to reduce the spatial resolution of the feature maps. The pooling layer helps increase the receipt field of the neurons. In the feature encoding stage, all the convolutional layers are grouped into L feature extraction blocks (L = 4 in our implementation), as shown by the blue blocks in Fig.4. Each block contains four convolutional layers with ReLU nonlinearity and 3 × 3 kernels. The first three layers generate 64-channel feature maps, while the last layer doubles the number of channels followed by a pooling layer to reduce the spatial resolution of the feature maps with scaling factor 0.5. In the pooling layers, the feature maps are first convoluted with 2 × 2 kernels and then subsampled by a

ISSN- 2394-5125 VOL 8, ISSUE 02, 2021

scaling factor of 2 along both axes. The feature decoding part also contains a series of convolutional layers, which are also grouped into four blocks followed by an upsampling layer to increase the spatial resolution of the feature maps. As the finally extracted feature maps lose a lot of spatial information, directly reconstructing images from the finally extracted features cannot recover fine image details. To address this issue, the feature maps of the same spatial resolution generated in the encoding stage are fused with the upsampled feature maps generated in the decoding stage, for obtaining newly upsampled feature maps. Each reconstruction block also consists of four convolutional layers with ReLU nonlinearity and 3 × 3 kernels. In each reconstruction block, the first three layers produce 128-channels feature maps and the fourth layer generate 512-channels feature maps, whose spatial resolutions are upsampled with a scaling factor of 2 by a deconvolution layer. The upsampled feature maps are then fused with the feature maps of the same spatial resolution from the encoding part. Specifically, the fusion is conducted by concatenating the feature maps. The last feature decoding block reconstructed the output image. A skip connection from the input image to the reconstructed image is added to enforce the denoising network to predict the residuals, which has be verified to be more robust.

3. EXPERIMENTAL RESULTS

In this section, by comparing our HBRNN with ARCNN image restoration algorithms, we could see the huge advantage of our Hybrid Bayesian Regularization neural network of PSNR and structural similarity index (SSIM) . We only focus on the restoration of luminance channel (in YCrCb space) in this paper. Given a reference image f and a test image g, the SSIM can be defined as the following:

$$SSIM(f, g) = l(f, g)c(f, g)s(f, g)$$

We compute PSNR and SSIM to evaluate quantitatively the denoising results, which can be seen in Table I and Table II. From the comparing results on noise level $\sigma = 25$ and 50, we will have the following observations. Firstly, HBRNN holds the overwhelming superiority on PSNR than other ARCNN methods with the provided two noise levels in average. Especially, the superiority over the second best method reaches to 19 dB and 27 dB on $\sigma = 25$ and 50 respectively. It should be noticed that our method is always the best, but the second best changes when noise level varies. This reflects the robustness of HBRNN. Secondly, with regards to each image, HBRNN still has absolute advantage on PSNR..

Table 1. PSNR (dB) results

Image	PSNR-ARCNN	Proposed Method
River_Lake.jpg	42.7dB	70.83 dB
Peppers.png	51.2 dB	70.8 dB
Lena.jpg	47.12dB	70.13 dB

Table 2. SSIM results

Image	SSIM-ARCNN	Proposed Method
River_lake.jpg	0.74	0.77
Peppers.png	0.80	0.83
Lena.jpg	0.71	0.74

ISSN- 2394-5125 VOL 8, ISSUE 02, 2021

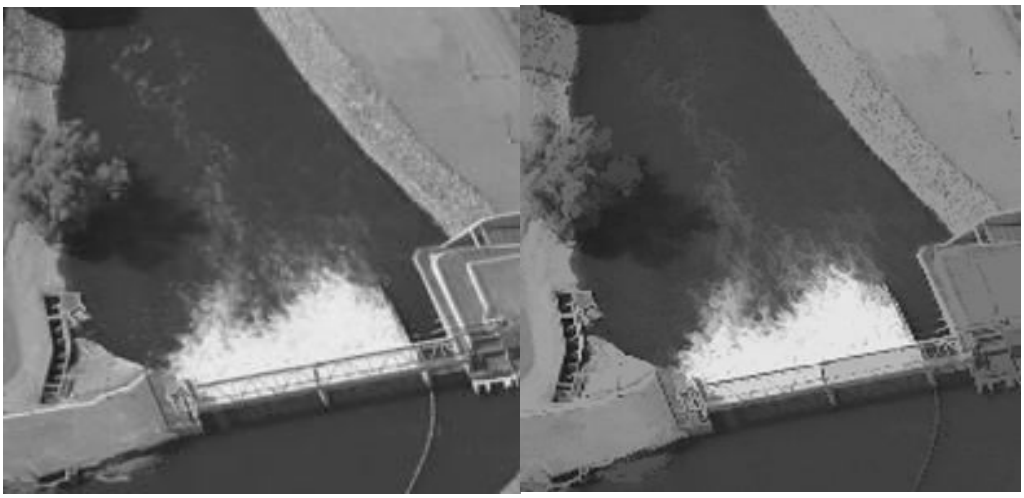

Fig. 5: Output of ARCNN for river lake (left). Output of HBRNN for river lake (right)

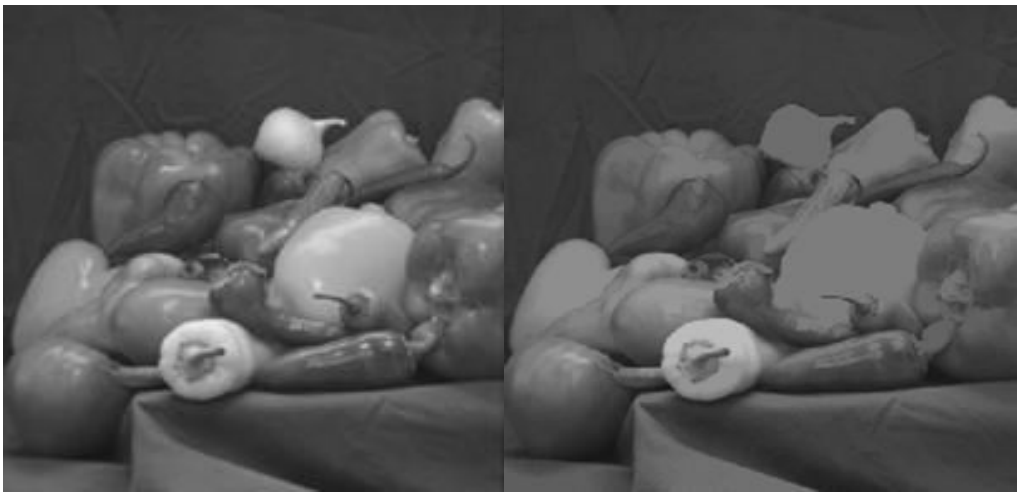

Fig.6:Output ofHBRNN for peppers (left)Output of ARCNN for peppers (right).

Fig.7: output of ARCNN for Lena (left).Output of HBRNN for lena (right).

Among 3 images, HBRNN has the highest PSNR values on almost all images regardless of the increase of noise variance. Hence, no matter on the whole or individuals, our HBRNN both shows the great preponderance than other compared methods in terms of PSNR. From Table II, it can be seen that HBRNN achieves significantly better SSIM indices than any other method. To be specific, ARN

ISSN- 2394-5125 VOL 8, ISSUE 02, 2021

exceeds the second best method 0.02 and 0.03 on noise level σ = 25 and 50 respectively. Similar to the PSNR metric, under the metric of SSIM, our HBRNN always shows the best performance while the other methods shake when the noise level changes, which also claims the robustness of HBRNN

Fig.8:Output of HBRNN for industry.

Fig.9:Output of ARCNN for industry.

4.CONCLUSION

In this paper, we attempted to improve the performance of image blind denoising by exploiting deep learning based methods in the absence of paired training data. The proposed HBRNNcan improve the blind denoising performance. Extensive experiments demonstrate the superiority of our method. To the best of our knowledge, we are the first to explore the potential of HBRNNin noise modeling and employ it in denoising tasks.

REFERENCES

[1] M. Elad and M. Aharon, "Image denoising via sparse and redundant representation over learned dictionaries," IEEE Transactions on Image Processing, vol. 15, no. 12, pp. 3736–3745, 2006.

[2] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," IEEE Transactions on image processing, vol. 16, no. 8, pp. 2080–2095, 2007.

[3] W. Dong, X. Li, L. Zhang, and G. Shi, "Sparsity-based image denoising via dictionary learning and structural clustering," in Proc. of the IEEE CVPR, 2011, pp. 457–464.

ISSN- 2394-5125 VOL 8, ISSUE 02, 2021

[4] W. Dong, G. Shi, and X. Li, "Nonlocal image restoration with bilateral variance estimation: a low-rank approach," IEEE Transactions on image processing, vol. 22, no. 2, pp. 700–711, 2013.

[5] W. Dong, X. Li, L. Zhang, and G. Shi, "Weighted nuclear norm minimization with application to image denoising," in Proc. of the IEEE CVPR, 2014, pp. 2862–2869.

[6] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and superresolution by adaptive sparse domain selection and adaptive regularization," IEEE Transactions on image processing, vol. 20, no. 7, pp. 1838–1857, 2011.

[7] A. Danielyan, V. Katkovnik, and K. Egiazarian, "Bm3d frames and variational image deblurring," IEEE Transactions on image processing, vol. 21, no. 4, pp. 1715–1728, 2012.

[8] W. Dong, L. Zhang, G. Shi, and X. Li, "Nonlocally centralized sparse representation for image restoration." IEEE Transactions on Image Processing, vol. 22, no. 4, pp. 1620–1630, 2013.

[9] W. Dong, G. Shi, Y. Ma, and X. Li, "Image restoration via simultaneous sparse coding: Where structured sparsity meets gaussian scale mixture," International Journal of Computer Vision, pp. 1–16, 2015.

[10] A. Marquina and S. J. Osher, "Image super-resolution by tvregularization and bregman iteration," Journal of Scientific Computing, vol. 37, no. 3, pp. 367–382,