International Journal of Future Generation Communication and Networking
Vol. 13, No. 2s, (2020), pp. 25–32

Deep Joint Denoising and Demosaicking Using Convolutional Neural Network

Aishwarya Arbuj[1], Sonal Jagtap[2*], Akash Agrawal[3], Rushikesh Dalpe[4], Pallavi Sawant[5]

[1,2,3,,4,5] *Dept. of E & TC Engg., Smt. Kashibai Navale College of Engineering, Savitribai Phule Pune University, Pune*

[1]*arbujaishwarya@gmail.com*

[2*]*skjagtap.skncoe@sinhgad.edu*

[3]*akashach08@gmail.com*

[4]*rushikeshdalpe333@gmail.com*

[5]*pallavi.msawant_skncoe@sinhgad.edu*

Abstract

Present day computerized cameras depend on consecutive execution of independent image preparing steps to create sensible pictures. The initial two stages are normally identified with denoising and demosaicking where the previous intends to decrease Noise from the sensor also, the last changes over a progression of light power readings to shading pictures. Present day approaches attempt to mutually take care of these issues, i.e. joint denoising-demosaicking which is an innately badly presented issue given that 66% of the power data are missing and the rest are bothered by Noise. While there are a few AI frameworks that have been as of late presented to take care of this issue, proposed a novel calculation which is enlivened by amazing old style image regularization strategies, enormous scale advancement and profound learning procedures. The inferred neural system has a straightforward and clear translation contrasted with other discovery information driven approaches. The broad experimentation line illustrates that the proposed system beats any past approaches on both boisterous and Noise free information crosswise over a wide range of datasets. The improvement in remaking quality is credited to the principle way structure system design, which as a result requires less trainable parameters than the current state of-the-workmanship arrangement and besides can be proficiently prepared by utilizing an altogether more modest number of preparing information than existing profound demosaicking systems.

Keywords*-Deep learning; denoising; demosaicking; convolutional neural network (CNN).*

I. INTRODUCTION

In order to create realistic images, modern digital cameras depend on sequential execution of distinct image processing steps. The first two steps are generally associated with denoising and demosaicking where the former aims to decrease the sensor's noise and the latter transforms a sequence of light intensity measurements into colour images. Modern methods attempt to address these issues collectively, i.e. joint denoising-demosaicking, which is an inherently ill-posed problem considering that two-thirds of the data on intensity is missing and the remainder are disturbed by noise. While there are several machine learning systems lately implemented to fix this issue. A new algorithm inspired by strong methods of classical image regularization, large-scale optimization, and deep learning techniques. Proposed neural network has a clear and transparent interpretation compared to other methods powered by black-box information. Proposed network perform in both noisy and noise-free information across multiple datasets. This improvement in the quality of restoration is ascribed to the principled manner in which proposed network architecture is designed, which needs fewer trainable parameters than the present state-of - the-art solution and can also be trained effectively using considerably fewer training information than existing profound demosaicking networks.

II. RELATED WORK

The first network which was basically, additive white Gaussian noise(AWGN) includes convolutionary layers as a key element, while the second, IR-based relies instead on non-local filtering layers, enabling the intrinsic non-local self-similarity properties of natural pictures to be exploited. Not at all like most existing profound system approaches requiring the preparation of a particular model for each considered noise level, the proposed models can deal with a wide scope of noise levels using a single set of learned parameters, while they are very robust when the noise degrading the latent image does not match the noise statistics used during training. The latter argument is backed by outcomes that we report corrupted by unidentified noise on publicly accessible pictures and compare them with alternatives acquired through competing techniques. At the same time, the networks introduced achieve excellent results under additive white Gaussian noise (AWGN), which are comparable to those of the current state-of - the-art network, while relying on a shallower architecture with the number of trained parameters being one order of magnitude smaller [1].

Taking the reference of [1], Proposed an IR-based denoising algorithm whose iterative steps can be effectively calculated. The iterative procedure is then unfurled into a profound neural system comprising of different denoisers interleaved with back- projection (BP) modules to guarantee consistency of perception. It proposes a convolutionary neural network (CNN) based denoiser that can exploit natural images ' multi-scale redundancies. As such, not only does the proposed network exploit the powerful denouncing ability of DNNs, it also leverages the observer models prior. Through end-to-end training, it is possible to jointly optimize both denoisers and BP modules. Experimental results on several IR tasks, such as image denoising, super- resolution and debl, show that the proposed method can lead to very competitive and often state-of - the-art results on multiple IR tasks, including image denoising, deblurring, and super-resolution[2].

As mentioned above, the structural insights of traditional methods of optimization and the speed of recent network-based methods. Specifically, we are proposing a novel profound system organized, named ISTA-Net, propelled by the Iterative Shrinkage-Thresholding Algorithm (ISTA) to streamline a general l 1 standard CS remaking model. We are developing an effective strategy to solve the proximal mapping associated with the sparsity-inducing regularize using nonlinear transforms in order to cast ISTA into deep network form. All ISTA-Net parameters (e.g. nonlinear transformations, shrinkage thresholds, step sizes, etc.) are learned end-to-end rather than hand-crafted. In addition, given that natural image residuals are more compressible, an improved version of ISTA-Net in the residual domain, dubbed {ISTA-Net} +, is obtained to further enhance CS reconstruction. Extensive CS experiments show that, while retaining quick computational velocity, the suggested ISTA-Nets outperform current state-of - the-art optimization and network-based CS techniques by big margins [3].

After the CS techniques by big margins, the recent revival of interest in artificial neural networks was fuelled in various image processing and computer vision tasks by their successful applications. In this work, we use the rotational variance of the natural image patch distribution and propose a multilayer-based multilayer-based multilayer neural network to demonstrate image. Compared to state-of - the-art approaches requiring much larger neighbourhoods, we show that it does surprisingly

well [4].
And it also trains a large corpus of images with a deep neural network instead of using hand-tuned filters. While profound learning has demonstrated extraordinary achievement, it's innocent application utilizing existing preparing datasets doesn't give good outcomes to our concern because of the absence of hard cases. We present metrics to identify difficult patches and techniques for photographs of the mining community for such patches in order to create a better training set. Our experiments show that both noisy and noise-free data are outperformed by this network and training procedure. Besides, our calculation is a request for extent quicker than the past best performing systems [5].

III. METHODOLOGY

Neural network used by previous researchers for solving the joint denoising-demosaicking problem that yielded state-of-the-art results on various datasets, both real and synthetic, without the need of millions of training images. There is a close connection between proposed algorithm and some instances of the proximal gradient descent algorithm, such as the iterative residual network algorithm.

In fact, the denoising of our approach is equivalent to computing the proximal operator. However, between the two algorithms above and our approach, there are two differences that we would like to highlight. Firstly, a distinctive difference between CNN the based approach and the proximal operator is the fact that the denoiser can only approximate the solution, so in a sense, it is an inexact proximal solution. Thus the discussed algorithm acts as an inexact Proximal Gradient Descent (IPGD). Secondly, the network algorithm requires the exact form of the employed regularizer, such as total variation. In contrast, the method implicitly learns the regularizer from available data as a part of the proximal approximation. Of course, there is no straightforward way to derive the type of regularization that the deep learning denoiser has learned during training.

Noise in the image might be presented either during its development or during its chronicle. For example the Noise that is presented in a photographic film is because of the capacity of the development which is the recording mechanism.

The two separate procedures that are probably going to add to the noise in the images are as:- a)Arbitrary changes in the quantity of photons and the photoelectrons on the photoactive surface of the finder.

b)The irregular warm Noise which is created in the circuit that detects the picture, gains it and procedures the sign from the identifier's photoactive surface. Noise may likewise be presented during the transmission of brilliant vitality.

It is obvious that the Noise presented in the gaining procedure of the image is incompletely signal ward and somewhat signal autonomous added substance noises.

ISSN: 2233-7857 IJFGCN

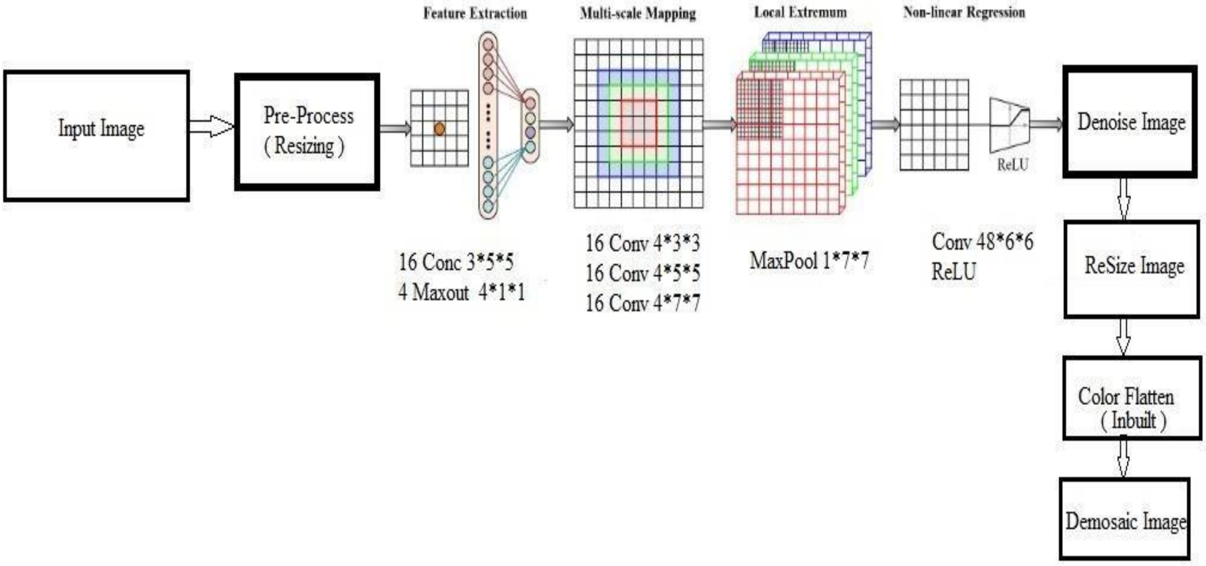

Fig 1 System architecture

The images are degraded by noise which is some random error. Noise could be picked along with the image during capturing, transmitting or during the processing. It could be dependent or independent of the image content. Variety of noise models are used to create different types of noise for the images. The probabilistic characteristics of the noise can be used to describe it. The often used idealized noise, which is also called white noise, has the intensity that does not decrease with increasing frequency. A special case of white noise is Gaussian noise. A random variable with Gaussian distribution has its probability density given by the Gaussian curve. Noise which is normally dependent of the images signal occurs when an image is transmitted through some channel. This signal independent degradation is called additive noise and this can be described by the model.

$$f(x, y) = g(x, y) + v(x, y) \qquad (1)$$

Where the noise ν and the input image g are independent variables. In many cases the magnitude of the noise depends on the very signal magnitude. When the noise magnitude is much higher in comparison with the signal, this model describes multiplicative noise. Quantization noise (Uniform noise) occurs when insufficient quantization levels are used, for instance, 50 levels for a monochromatic image. Presence of impulse noise implies that an image is corrupted with individual noisy pixels whose brightness differs significantly from that of its neighbourhood. The term salt-and-pepper noise is used to describe saturated impulse noise, i.e., an image corrupted with white and/or black pixels. Salt-and-pepper noise can corrupt the binary images. The efficiency issues for the classification the blur type is monitored by using an edge detector to obtain binary input values for the CNN training which benefits the blur analysis task.

A convolutional neural network (CNN) could likewise be a chosen sort of counterfeit neural system that utilizations perceptron's, an machine learning unit calculation, for supervised learning, to data analyses CNNs apply to image processing, tongue processing and other forms of cognitive tasks.

CNN includes six components: Convolutional layer, Sub-sampling layers, Rectified linear unit (ReLU), Fully connected layer, Output layer and Softmax layer

1) *Convolution Layer:* Convolutional layers are determined by number of generated maps and kernel's size. The kernel is moved over the valid area of the given image (perform a convolution) for generating the map. If f_k be a filter with a kernel size n × m and is supposed io applied into the given image x, output of the layers can be calculated as follows:

$$C(x_{u,v}) = \sum_{i=-\frac{n}{2}}^{\frac{n}{2}} \sum_{i=-\frac{m}{2}}^{\frac{m}{2}} f_k(i,j) x_{u-i,v-j} \qquad (2)$$

Where each CNN neuron has n × m number of input connections.

2) *Sub-sampling layers:* the map size of previous layer in order to increase the invariance of the kernels. Sub-sampling includes two types of average pooling and maximum-pooling . By applying maximum function in the Max-Pooling, input value is reduced at the x_i. If m be the size of the kernel, output of max-pooling can be calculated as follows:

3) *Rectified linear unit:* A rectified linear unit is a activation function which it simply thresholded at zero and can be calculated as follows:

$$R(x) = max(0, x) \qquad (3)$$

ReLU has advantages over tanh/sigmoid function in which it can be implemented by simple Thresholding at zero, while in tanh/sigmoid there are expensive operations like exponentials. ReLU is also prevents loosing gradient error, and extremely accelebrate the stochastic gradient descent convergence compared with the lanh/sigmoid functions.

4) *Fully connected layer:* Fully connected layers are similar to neurons in general neural networks which its neurons are fully connected with every neurons in the prior layer. In the be input with size W and the number of neurons represented by $x$in the fully connected layer, the layer can be calculated as follows:

$$F(x) = \sigma(W * x) \qquad (4)$$

Where σ is activation function.

5) *Output Layer***:** The output layer represent class of the input image which its size equal to number of classes. Output vector x produce resulting class as follows:

$$C(x) = \{i | \exists i \forall j \neq i : x_j \leq x_i\} \qquad (5)$$

6) *Softmax layer:* The error of the network is propagated back through a softmax layer. If N be the size of the input

vector, a mapping can be calculated by softmax such that:

$$S(x) : R \rightarrow [0,1]^N . \qquad (6)$$

IV. RESULTS AND DISCUSSION

In the following system we pre-train the denoiser images on simple case where M=I. The pre-

ISSN: 2233-7857 IJFGCN

training of dataset has proven that it vastly reduces the time required for training. These images were split in two sets, 400 were used to form a train set and the rest 100 formed a validation set. All the images were randomly cropped into patches of size 180*180 pixels. Using the pre-trained denoiser, the overall network is further trained end-to-end to minimize the averaged loss.

The demonizing step of our approach is equivalent to the well-known proximal operator. Emphasize. Firstly, a distinctive difference between our CNN based approach and the proximal operator is the fact that the denoiser can only approximate the solution, so in essence it is an inexact proximal solution. Even in the non-convex cases like ours, an inexact proximal solution can converge in the same convergence rate as the original PGD algorithms, provided that certain assumptions apply. Fig. 1(a), 2(a), 3(a) is obtained by using an adaptive homogeneity- directed demosaicking algorithm and 1(b), 2(b), 3(b) by using DNN algorithm whereas 1(c), 2(c), 3(c) is obtained by using convolutional neural network by using IT Demosaicking Dataset.

Fig 1 (a) AHD (b) DNN (c) CNN

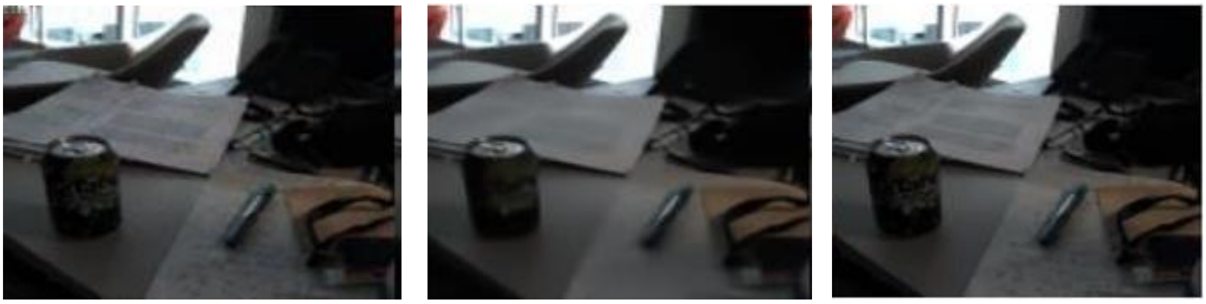

Fig 2(a) AHD (b) DNN (c) CNN

Fig 3(a) AHD (b) DNN (c) CNN

Table I: - Accuracy comparison

Sr.No	**Algorithm**	**Accuracy**
1	An adaptive homogeneity- directed demosaicking algorithm [7].	60%-70%

ISSN: 2233-7857 IJFGCN

2	DNN Algorithms [2]	75%-85%
3	Convolutional Neural Network Algorithm	80%-90%

K. Hirakawa and Ep. al.[7] used An adaptive homogeneity- directed demosaicking algorithm in fig 1(a),2(a) and 3(a) where he suggested a system where Metric neighborhood modeling techniques were used to compare the level of color artifacts that are present in images and to select the direction for interpolation. Filter bank interpolation techniques were developed to cancel aliasing, and interpolation artifact reduction iterations suppressed color artifacts. The discussed demosaicking algorithm estimates missing pixels by interpolating in the direction with fewer color artifacts. The aliasing problem is addressed by applying filter bank techniques to 2-D directional interpolation. The interpolation artifacts are reduced using a nonlinear iterative procedure. Experimental results using digital images confirm the effectiveness of this approach where as M. Gharbi and Ep.al. [2] used demosaicking algorithms and have achieved result as shown in in fig 1(b), 2(b) and 3(b) where they have suggested an adaptive approach as well as a new moire´ detection metric to tackle these challenges. The method outperforms state-of-the-art solutions in terms of both perceptual and statistical visual quality, while being an order of magnitude faster.

V. CONCLUSION

A new deep learning system for high-quality pictures for the raw sensor pictures compared to the present state of the art network, our demosaick network delivers superior outcomes both quantitatively and qualitatively. In the meantime, even when trained on tiny data sets, strategy is able to generalize well and the amount of the network parameters compared to other competing alternatives is kept low. Finally, introducing CNN as an effective way to train networks with an arbitrary number of sets, which hope will pave the way for the successful training of learning-based approaches for other similar tasks of image restored.

REFERENCES

[1] Y. Wang, “A multilayer neural network for image demosaicking,” pp. 1852–1856, Oct 2014.

[2] M. Gharbi, G. Chaurasia, S. Paris, and F. Durand, “Deep Joint Demosaicking and Denoising,” ACM Trans. Graph., vol. 35, no. 6, pp. 191:1–191:12, Nov. 2016.

[3] W. Dong, P. Wang, W. Yin, G. Shi, F. Wu, and X. Lu, “Denoising Prior Driven Deep Neural Network for Image Restoration,” CoRR, vol. abs/1801.06756, 2018.

[4] J. Zhang and B. Ghanem, “Ista-net: Iterative Shrinkage-Thresholding Algorithm Inspired Deep Network for Image Compressive Sensing,” CoRR, vol. abs/1706.07929, 2017.

[5] S. Lefkimmiatis, “Universal denoising networks: A Novel CNN Architecture for Image Denoising,” in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2018.

[6] Filippos Kokkinos and Stamatios Lefkimmiatis, “Iterative Residual Network for Deep Joint Image Demosaicking and Denoising,” Journal of Transactions in Image Processing IEEE,2019.

[7] K. Hirakawa and T. W. Parks, “Adaptive homogeneity-directed demosaicing algorithm,” IEEE

ISSN: 2233-7857 IJFGCN

Transactions on Image Processing, vol. 14, no. 3, pp. 360–369, March 2005.

[8] Sarbjit Kaur and Er. Ram Singh,“ Image De-noising Techniques: A Review Paper”, International Journal For Technological Research In Engineering Volume 2, Issue 8, April-2015.

[9] Shan Gai1 , Long Wang, Guowei Yang and Peng Yang,“Sparse representation based on vector extension of reduced quaternion matrix for multiscale image denoising”, IET Image Processing Research Article,2016.

[10] Alisha P B, Gnana Sheela K ,“Image Denoising Techniques-An Overview”, IOSR Journal of Electronics and Communication Engineering (IOSR- JECE),2016

ISSN: 2233-7857 IJFGCN