JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.2043-2052
https://publishoa.com
ISSN: 1309-3452

Image Deblurring using Repeating Gradient Descent Model

1. **Mr. Gangireddy Prabhakar Reddy**, Asst. Professor,St.Peter's Engineering College, Telangana, India, prabhakar.sp17@gmail.com
2. **Mohammed Afraaz**, Student(B.Tech.-CSE), St. Peter's Engineering College, Telangana, India, mohammedpsvita@gmail.com
3. **Amit Ranjan Bej**, Student(B.Tech. -CSE), St. Peter's Engineering College, Telangana, India, arb11459@gmail.com
4. **Y. Lokesh Kumar**, Student(B.Tech. -CSE), St. Peter's Engineering College, Telangana, India, lokeshkumaryelukanti@gmail.com
5. **Bandaru Chintu**, Student(B.Tech. -CSE), St. Peter's Engineering College, Telangana, India, isaacchintu333@gmail.com

Received 2022 April 02**; Revised** 2022 May 20**; Accepted** 2022 June 18

Abstract: Technology has made it easier for people to take pictures, and every single day over 10 billion pictures are taken. A greater part of which is taken through the lens of a smartphone which is prone to motion blur. In this paper, we implement a neural network architecture to recover a sharp image by deconvolving the blur kernel from the motion-affected image. Visual deconvolution eliminates blurriness with a specific blurred kernel, that's necessary and challenging as a result of the inverse problem. The prevalent method relies on optimization subjects to regularisation algorithms which are constructed or learnt from past observations. Previous eager approaches have demonstrated higher reconditioningcalibre, and their limited and static model design makes them impractical. They are exclusively concerned with obtaining a reference and must be aware of the noise level for deconvolution. By developing a universally applicable optimizer that uses a specific gradient descent approach, we bridge the divide between optimization-based and learning-based techniques. We present a Repeating Gradient Descent Model (RGDM)using methodically integrating deep learning based advanced networks into a fully specified gradient descent scheme. Using a convolutional neural network, a dynamic variable update unit shared across stages is employed to create updates from results that are recent. The RGDM acquires an embeddedpicture reference and a update method that can be used universally by training on various samples and recursive supervision. The learnt optimizer may be employed multiple times to increase the quality of various debased observations. The suggested strategy is highly interpretable and very generalizable. Extensive tests based acrossartificialreference points and demanding realisticimages reveal that the suggested universally applicable strategy is successful, resilient, and applicable to image deblurring applications in the real world.

Keywords: Image deconvolution, Deblurring,Universal Update Rules, Gradient Descent,Optimized Network

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.2043-2052
https://publishoa.com
ISSN: 1309-3452

I. Introduction:

The objective of image deblurring, is to restore a clear picture from a fuzzy one. the blurry picture y $\in R^m$ is represented as a latent image convolution x $\in R^n$ and a blurry kernel k $\in R^l$:

$$\boldsymbol{y} = \boldsymbol{k} * \boldsymbol{x} + \boldsymbol{n}(\mathbf{1})$$

where ***** represents the operator representing convolution, a white Gaussian noise term represented by n $\in$ R^m with an indefinite standard deviation known as noise level. Given **y**, **k** namely a blurred image and its accompanying blurred kernel, recovering a clearer **x** is referred to as reverse convolution, which is frequently employed as a sub-component ofdeblurring of blind images [1]-[3].

As a result of unidentified noise and the loss ofinformation that is of high frequency, single picture deconvolution is difficult and mathematically ill-posed. Numerous traditional approaches use various manually created empirical statistics [4]-[6]images or trained generative models (e.g., G.M. Models [7])which often result in nonconvex andoptimizationthat is slow. The optimization techniques are used in iterative updating the pictures using the references and the model for imaging in (1). For effectiveness, discriminative approaches [8]-[10] for learning are examined to train mapping functions from a blurrypictureto a clear picture, that are often confined to certain blurand noise kernels at different levels.

Deep neural networks (DNNs) are increasingly employed for developing image restoration models as a result of their effectiveness in several computer vision applications. As it is impossible to directly implement Deep Neural Nets to the de-convolution for distinct blur kernels [10], several techniques try unwrapping an optimization algorithm as a predefined waterfall system with a predetermined no. of stages with multiple simple networks that are merged at each step [9], [11], [13], [15].

Typically, the components of Deep Neural Networksjust model the prior operators. [12]In the structure of these static models, the operators based on Deep Neural Networks are trained particularly using theincomplete result of the preceding phase. Consequently, these models often need tailored training formanual parameter adjustment for different levels of noise or for a certain blurred picture which restricts their practical uses. Although learning-based solutions use optimization strategies for the deblurring software, they are limited to learning a staticfunction and don't implement dynamic solutions throughout the optimization process.

We overcome the aforementioned concerns by learning a global picture deconvolution optimizer. RGDMis a recurrent Deep Neural Network architecture inspired from gradient descent optimization techniques. Using a universal image updating unit that replicates the gradient descent optimization approach, the unknown variable x is repeatedly updated by RGDM. In order to do this, we manually provide parameters and train a universally applicable gradient descent optimizer that can be dynamically used to update the input image depending on the past changes.

We provide parameters and train key components of the generic gradient descent algorithm, unlike earlier techniques that primarily focused on image prior learning. In prior techniques, CNNs were mostly used as

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.2043-2052
https://publishoa.com
ISSN: 1309-3452
a denoiser for picture gradients in certain splitting-based optimization techniques. During the implementation of the suggested model for the optimizer, we note that adding the conventional optimization technique for the deep neural network architecture is advantageous, as it is better able to exploit the problem's structure.

The suggested model learns the optimizing processes, and also the items related to the regulariser, which represents theprior image. In addition, the optimizer shared between steps is trained to resolvevariable update statues dynamically, which makes it more adaptable and generic to resolve observations with varying levels of noise and varied blur kernels. For the input pictures with varying degrees of degeneration, the trained optimizer may generate precise and clear outcomes through many iterations in an adaptable manner.

In conclusion, this paper's primary contributions are:

• This paper offers a novel discriminative learning model, RGDM, for learning a deconvolutionoptimizer of animage. The RGDM includes a number of Convolutional Neural Networks into the overall gradient descent approach. RGDM uses a parameter sharing method that tends to develop a global optimizer that can be repeatedly applied multiple times to improve the performance on insights gained, making it a practically applicable technique.

• Training a single RGDM model enables it to deal with several forms of blurriness and noises. Repeated trials on both simulatedand actual photos demonstrate that the constraint-free RGDM trained from a simulated dataset may achieve comparable or superior results compared to other reputed algorithms.

II. Research Background:

A. Observation 1:

For text image deblurring,a simple but effective regularised model based on intensity is presented [1]. The suggested image prior is driven by the observation of unique text image qualities. On the basis of this prior, we create a technique for efficient optimization that generates credible intermediate results for estimating kernels.

The suggested technique in [1] does not need elaborate filtering procedures to pick prominent edges, which are essential to current deblurring algorithms. We examine the link with various edge-selection-based deblurring algorithms and suggest a method to choose edges in a more systematic manner. In the last phase of picture restoration, a straightforward technique is devised a for removing artefacts and rendering improved deblurred images. The experimental findings reveal that the proposed approach outperforms current text picture deblurring techniques. In addition, the authors demonstrate that the suggested approach can efficiently deblur photos with low light.

B. Observation 2:

For text image deblurring, the authors present an effective method based ongradient [2]. The suggested image prior is driven by the observation of unique text image qualities. On the basis of this prior, they create a technique for efficient optimization that generates credible intermediate results for estimating kernels. The suggested technique does not need elaborate filtering procedures to pick prominent edges, which are essential to current deblurring algorithms. The authors examine the link with various edge-selection-based deblurring algorithms.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.2043-2052
https://publishoa.com
ISSN: 1309-3452

C. Observation 3:

It is difficult to eliminate pixel-wise uneven motion blur. The prevalent option is to guess the blur kernel by adding a reference image, however it is difficult to select a sufficiently informative and generic prior. Instead of imposing a theory-based prior, the authors suggest learning one [3].

Learning a reference image over a latent image would necessitate modelling all potential image content. The technique is founded on the important fact that understanding the motion flow enables the model to concentrate on the source of blur regardless of the picture content. This is a considerably simpler job, but itskips the repeated procedure that is generally used to apply latent image priors.

Using a fully-convolutional network (F.C.N.), this method calculates the motion flow from the blurred picture and finds an unblurred image from this calculated flow. Their F.C.N. is the first global mapping from a blurry picture to a dense network that predicts a motion flow. They construct synthetic blurred-image-motion-flow pairings by simulating motion flows, so eliminating the requirement for human labelling.

III. Existing Systems:

The use of image deblurring based on a non-blind blur kernel has been heavily researched in the field of computer vision and deep learning. The existing techniques can be broadly divided into two categories: manually constructedgeneric methods and the learning-based techniques.

Empirically constructed deconvolutiontechniques

Many manually-designed approaches use empirical statistics on natural image gradients as the regularization or prior termsuch as the total variation (TV) regulatorsparsity prior on second-order image gradientsand approximate hyper-Laplacian distribution. Various optimization approaches have been researched and discussed for solving the image deblurring problem. One of which is the alternating direction method of multipliers (ADMM). These genericapproaches are often sensitive to the hyperparameter settings and are computationally costly [8], [9].

The learned algorithms also needwell-adjustedhyperparameters for specific noise levels. A fewmethods address deconvolution by straightforwardly learning a differentiating function for efficiency. Schuler et al. imposed a regularized inversion of the blur in the Fourier domain and then remove the noise using a learned multi-layer perceptron (MLP). Schmidt and Roth proposed shrinkage fields (CSF), an efficient discriminative learning procedure based on a random field structure. Schmidt et al. proposed an approach based on Gaussian conditional random field, in which the parameters are calculated through regression trees [4], [5], [17].

IV. Proposed Methodology:

This section will begin with a quick review of the deconvolution algorithm that is based on a non-blind model and the generic gradient descent technique. Then, we present the RGDM model that uses a gradient descent technique which is completely specified. We conclude by discussing the methodology fortraining and deblurring.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.2043-2052
https://publishoa.com
ISSN: 1309-3452

A. The Deconvolution Equation

As stated in the introduction, given a blurred image y and its blur kernel represented by k, a non-blind deconvolution algorithm using a Gaussian noise assumption should minimise the data exactness term $f(x) = 1/2\lambda y{-}Ax22$, where the noise level of y is represented by λ > 0. Given a regulator Ω(x) and taking into account the difficulty of the issue,deconvolution is implemented by reducing the minimization problem

$$\min_x \frac{1}{2\lambda} \|y - Ax\|_2^2 + \gamma\Omega(x) \qquad (2)$$

Here Ω(x) corresponds to the picture reference and the term $\gamma \geq 0$determines the regularisation intensity. In general, Ω(x) may have any shape [6].

Beginning at a basic level, we use the gradient descent algorithm. Let t be the step index. Thegradient descent is solved for $\hat{x}$which is an estimate of the input image afterupdating it an arbitrary number of times [39].

$$d^t = -\left(\nabla f(x^t) + \gamma\nabla\Omega(x^t)\right)$$

$$x^{t+1} = x^t + a^t d^t \qquad (3)$$

Here d^t represents the direction of gradient descent, a^t the learning rate, and$\nabla f(x^t)$ and $\nabla\Omega(x^t)$ the differentials of $f(\cdot)$ and$\Omega(\cdot)$ at t.During traditional gradient approaches, t is often found using a fixed or approximative line search. The solution to the reverse convolution problem:

$$f(x^t) = 1/\lambda(A^T Ax^t - A^T y) \qquad (4)$$

To improve the speedof optimization, we dispose d^tusing a scaling matrix D^twhich provides the information about the curvature.D^t is an inverse Hessian matrix when the second order information is used.

We finally obtain a standard equation for updating at a given step t:

$$\mathbf{x}^{t+1} = \mathbf{x}^t - \alpha^t \mathbf{D}^t \left(1/\lambda(\mathbf{A}^\top \mathbf{A}\mathbf{x}^t - \mathbf{A}^\top \mathbf{y}) + \gamma\nabla\Omega(\mathbf{x}^t)\right). \quad (5)$$

Given an initial valuex^0, the gradient descent solves equation (3) by repeatedly updating equation (5) until a stoppingcriterionisreached. The formula provided by (5) is an optimized equation that can be used to arrive at the global minimum and implements an optimizer that is applicable globally.

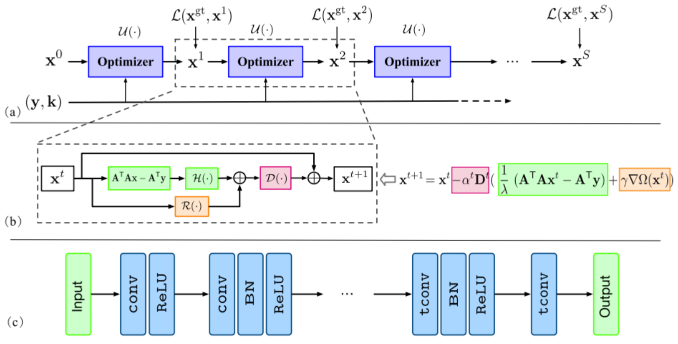

***Figure 1*.a).Overall architecture of our RGDM, b) Architecture of Optimizer Unit, c) Structure of CNN Blocks**

B. Neural Network Architecture

The model that can solve this issue is proposed as a Repeated Gradient Descent Model. We consider that the updates of x^t from using a repeated optimization method forms a series of variable length and constructs$U()$ using a globally applicable gradient descent unit (GDU) and implements it in a repeatedmanner.

In Gradient Descent Unit, the producer $G()$ takes the output of the previous iteration$x^t$and creates a differential of similar dimensions.$U()$contains subcomponents $R(), H(), and\ D()$which map the input to an output of the same dimensions.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.2043-2052
https://publishoa.com
ISSN: 1309-3452

Considering that convolutional neural networks are often used to represent comparable mapping functions, $R(), H(), and\ D()$are constructed using three CNNs that have the topology illustrated in Fig.1. (c). As determining the optimal structure for each $R(), H(), and\ D()$is not the primary objective, we employ an architecturally similar structure for all three. However, they are trainedusing different hyperparameters which provide diverse equations that are merged together resulting in the model.

C. Deconvolution Using Learned Optimizer

Though the model is trained quickly the suggested optimizer is able to provide a sustained improvement in picture quality with many testing rounds. Consequently, we use the learnt optimizer for deconvolution by using a varied number of steps and halts the procedure based on certain criteria, which is the common behaviour of a traditional algorithm.

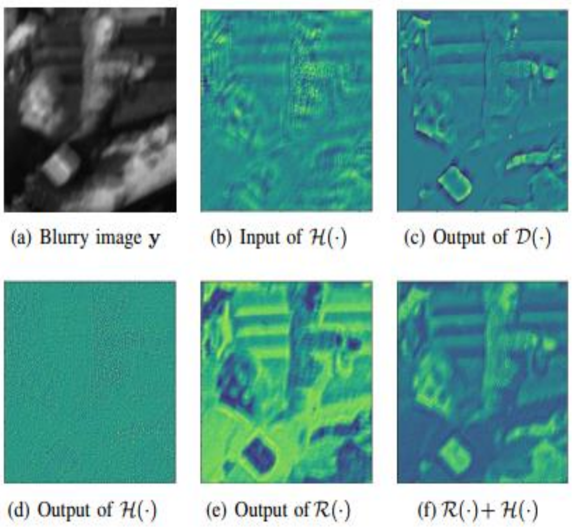

Figure 2. **Visualization of$R(), H(), and\ D()$.**

The suggested model utilises similar hyperparameters for several phases. By concentrating on each training iteration individually,the optimizer which can be applied globally is taught to improve various semi processed pictures. Thus, the optimizer can observe photographs that havedifferent noise levels. The recursive implementationby the optimizer improves the picture quality of the image in each step. Therefore, though we are training the algorithm using lesser layers it provides a foundation for more iterations as each iteration takes lesser time. In addition, the image previously learnt in the initial phases may be utilised to further stages of picture restoration.

V. Results Analysis:

1)Training:

To generate the dataset used for training and testing,40,960 RGB images of256×256 pixels are selected from the PASCAL VOC dataset which is collected by NYU, as the ground-truth images x_i. The simulated blurred images are created as per a standard approach. We generate five blur kernels for each input image$x_i$and produce the blurred image y_i, which gives 204,800 images in total. [11]

Then a noise term is added from N (0, σ2I), 8-bit quantization is used to implement a Gaussian noise term. As per our objective we train the data on a single model instead of creating a unique model for each noise level or blur kernel. We extract varied noise levels from the interval [0.3%, 1.5%] and also varied blur kernels from a pre-established set, which evaluates the ability of the network in outlying cases with highly varied data. The model is then tested with a 10% training sample to study the behaviour of the determined approach.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.2043-2052
https://publishoa.com
ISSN: 1309-3452

σ	Mea.	FD	Levin	EPLL	MLP	CSF	TNRD*	IRCNN	GradNet*	FDN	RGDN
0.59%	PSNR	32.36	33.60	34.35	31.55	32.08	–	33.35	–	**36.15**	35.04
	SSIM	0.917	0.934	0.941	0.876	0.916	–	0.884	–	**0.965**	0.954
1%	PSNR	30.85	32.01	32.45	30.68	28.12	28.88	33.14	31.43	33.62	**33.68**
	SSIM	0.892	0.913	0.930	0.882	0.828	0.854	0.896	0.912	0.949	**0.954**
2%	PSNR	28.84	29.92	30.03	28.16	21.68	28.10	30.09	28.88	29.70	**31.01**
	SSIM	0.851	0.877	0.883	0.841	0.594	0.824	0.887	0.841	0.896	**0.899**

Table 1. **Comparison with other algorithms at different noise levels**

2)Testing:

The actual testing is performed irrespective of the training data. We use several different datasets used in Image Processing to test our model. Varied experimental circumstances such as the use of varied noise levels isapplied to validate the strength of our approach. We apply certain stopping criterion and set the max number of iterations as 50.

We first conduct a complete mathematical comparison with other existing methods that were discussed previously. The conventional optimization-based methods usually rely on pre-determined notions that are implemented using a static principle. Existing systems are optimized for a fixed issue instead of a dynamic and universal solution of the problem. They rely on fixed hyperparameters for different test cases to overcome their static behaviour [4], [5], [7], [8].

3) Comparative Analysis:

As previously mentioned, we compare to a baseline of Levin with bilateral filtering for noise reduction. It is important to note that for Levin, we assumed the SNR to be equal to the average SNR of the entire set of test images. The results are shown in Table 1.

We see that our method performs about 4% better than the baseline on our test set for higher noise levels. You can see in Fig. 3. from our network, the background is much cleaner. That being said, thereare very significant colour artifacts that appear in the black region of our image. We believe this due to the fact that our inputs were not normalized. We therefore run into the issue in which some of ourinputs have really low dynamic range.

Imagine, for sake of argument, that one of our input images has values ranging from 0-0.05 and another has values ranging from 0-1. In this scenario, the contribution from the input with values ranging from 0-1 will dominate our loss function. This makes it vital that we re-scale our inputs so that their variability reflect their importance to our model. Since we don't know these prior, a good step would have been to normalize all of our inputs to the same standard deviation beforehand so that we guarantee that their variability is at least not the inverse of their importance. Since the extremely black and white regions lie at the ends of the visible

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.2043-2052
https://publishoa.com
ISSN: 1309-3452

colour spectrum, the images they match with the most have low effective resolution.

The model trained on the dataset performed well as given by the images below:

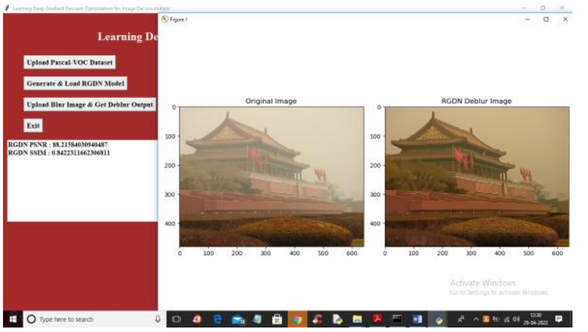

Figure 3. **Presence of Fog in image removed**

As observed, the model does well to recognize natural occurrences of blurriness such as fog and haze and also fixes blur caused by focusing on a subject.

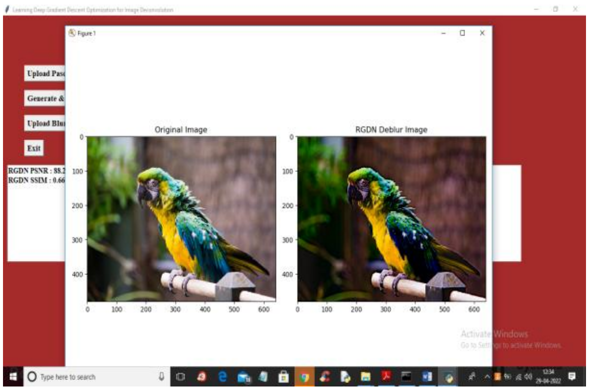

Figure 4. **Background blur removed**

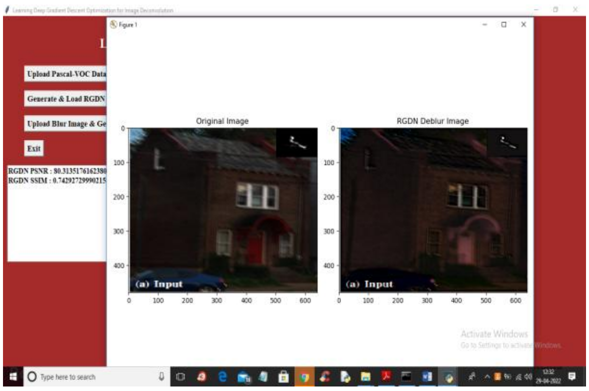

Figure 5. **Blurriness due to shaking fixed**

As we can see in Fig.4 the blurriness caused by shaking the camera is mostly fixed but as a result the image has been darkened. This issue can be observed among other dark samples as the information present in the image is too low to comprehend.

VI. Scope of Future Work:

In general, we have shown that our methodprovides results that are slightly better than the baseline of previous works. However, there are significant colour artifacts that can arise from it. Our output is also still not a completely sharp image reconstruction. For this reason, we believe that there are several modifications we would like to check in our next iteration of the project.

One method is the addition of a Generative Adversarial Network (GAN) section into the loss function. Adding this to the networkinforms it what a sharp image should look like, thereby improving the understanding of the network of better output and how to reach it.

Another method to expand the network is by enabling the GDU to simulate other sorts of sounds and losses. Considering the tests and comparisons demonstrated the potential of the proposed approach for tasks outside picture deconvolution, we may also apply it to further image restoration applications.

VII. Conclusion:

We built a RGDM that functions as an image reverse convolution optimizer. This network's components are influenced by principal structures of the gradient descent algorithm and constructed appropriately. The suggested RGDM learns a reference and adjusts adaptive constraints using a convolutional neural network's components. The suggested algorithm is trained on a dataset that consists of many diverse images including real-life examples and is thus capable of restoring a wider variety of blurred pictures than earlier methods. Our Gradient Descent Unit is intended for accommodating the Gaussian noise.

Acknowledgement:

We are thankful to St.Peter's Engineering College for helping usand supporting usin preparing this paper. We are also thankful to our guide Mr. Gangireddy Prabhakar Reddy for providing us this opportunity to improve ourselves.

References:

[1] J. Pan, Z. Hu, Z. Su, and M.-H. Yang, "Deblurring text images via l0-regularized intensity and gradient prior," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 2901– 2908.

[2] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in European Conference on Computer Vision (ECCV), 2010, pp. 157–170.

[3] D. Gong, J. Yang, L. Liu, Y. Zhang, I. Reid, C. Shen, A. v. d. Hengel, and Q. Shi, "From motion blur to motion flow: a deep learning solution for removing heterogeneous motion blur," The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[4] A. Levin, R. Fergus, F. Durand, and W. T. Freeman, "Image and depth from a conventional camera with a coded aperture," ACM transactions on graphics (TOG), vol. 26, no. 3, p. 70, 2007.

[5] D. Krishnan and R. Fergus, "Fast image deconvolution using hyperlaplacian priors," in Advances in Neural Information Processing Systems (NIPS), 2009, pp. 1033–1041.

[6] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," SIAM Journal on Imaging Sciences, vol. 1, no. 3, pp. 248–272, 2008.

[7] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," in The IEEE International Conference on Computer Vision (ICCV), 2011, pp. 479–486.

[8] U. Schmidt and S. Roth, "Shrinkage fields for effective image restoration," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 2774–2781.

[9] C. J. Schuler, H. Christopher Burger, S. Harmeling, and B. Scholkopf, "A machine learning approach for non-blind image deconvolution," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013, pp. 1067–1074.

[10] L. Xu, J. S. Ren, C. Liu, and J. Jia, "Deep convolutional neural network for image deconvolution," in Advances in Neural Information Processing Systems (NIPS), 2014, pp. 1790–1798.

[11] J. Zhang, J. Pan, W.-S. Lai, R. Lau, and M.-H. Yang, "Learning fully convolutional networks for iterative non-blind deconvolution," CVPR, 2017.

[12] J. R. Chang, C.-L. Li, B. Poczos, B. V. Kumar, and A. C. Sankaranarayanan, "One network to solve them all solving linear inverse problems using deep projection models," 2017.

[13] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep cnn denoiser prior for image restoration," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 3929–3938.

[14] R. Liu, X. Fan, M. Hou, Z. Jiang, Z. Luo, and L. Zhang, "Learning aggregated transmission propagation networks for haze removal and beyond," IEEE transactions on neural networks and learning systems, no. 99, pp. 1–14, 2018.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.2043-2052
https://publishoa.com
ISSN: 1309-3452
[15] J. Kruse, C. Rother, and U. Schmidt, "Learning to push the limits of efficient fft-based image deconvolution," in IEEE International Conference on Computer Vision (ICCV), 2017, pp. 4586–4594.

[16] U. Schmidt, Q. Gao, and S. Roth, "A generative perspective on mrfs in low-level vision," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013, pp. 1067–1074.

[17] D. Gong, M. Tan, Q. Shi, A. van den Hengel, and Y. Zhang, "MPTV: Matching pursuit based total variation minimization for image deconvolution," IEEE Transactions on Image Processing, pp. 1–1, 2018.

[18] T. Goldstein, B. O'Donoghue, S. Setzer, and R. Baraniuk, "Fast alternating direction optimization methods," SIAM Journal on Imaging Sciences, vol. 7, no. 3, pp. 1588–1623, 2014.

[19] L. Sun, S. Cho, J. Wang, and J. Hays, "Good image priors for nonblind deconvolution: Generic vs specific," in European Conference on Computer Vision (ECCV), 2014, pp. 231–246.

[20] U. Schmidt, J. Jancsary, S. Nowozin, S. Roth, and C. Rother, "Cascades of regression tree fields for image restoration," IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), vol. 38, no. 4, pp. 677–689, 2016.

[21] Y. Chen, W. Yu, and T. Pock, "On learning optimized reaction diffusion processes for effective image restoration," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 5261– 5269.

[22] S. V. Venkata krishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model-based reconstruction," in Global Conference on Signal and Information Processing (Global SIP), 2013, pp. 945–948.

[23] F. Heide, S. Diamond, M. Nießner, J. Ragan-Kelley, W. Heidrich, and G. Wetzstein, "Proximal: Efficient image optimization using proximal algorithms," ACM Transactions on Graphics (TOG), vol. 35, no. 4, p. 84, 2016.

[24] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," IEEE Transactions on Image Processing, vol. 4, no. 7, pp. 932–946, 1995.

[25] M. Jin, S. Roth, and P. Favaro, "Noise-blind image deblurring," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.