JES
Journal of Engineering Sciences

Vol 10, Issue 10, Oct /2019
ISSN NO:0377-9254

IMAGE DE-NOISING BASED ON CONVOLUTION NEURAL NETWORKS IN VIDEO'S

K. Vishnupriya (M.Tech) [1]
A. Hazarathaiah (Professor) [2]
Department of Digital Electronics and Communication Engineering[1]
Narayana Engineering College, Guduru, Spsr Nellore, Andhrapradesh, INDIA[1, 2]
kvishnupriyha@gmail.com[1] , a:hazarath@gmail.com[2]

ABSTRACT

Because of the quick induction and great execution, discriminative learning strategies have been broadly considered in picture denoising. Notwithstanding, these strategies for the most part get familiar with a particular model for each clamor level, and require numerous models for denoising pictures with various clamor levels. They additionally need adaptability to manage spatially variation clamor, restricting their applications in down to earth denoising. To address these issues, we present a quick and adaptable denoising convolutional neural organize, in particular FFDNet, with a tunable commotion level guide as the info. The proposed FFDNet takes a shot at down sampled sub images, accomplishing a decent exchange off between derivation speed and denoising execution. Rather than the current discriminative denoisers, FFDNet appreciates a few alluring properties, including (i) the capacity to deal with a wide scope of commotion levels (i.e., [0, 75]) adequately with a solitary system, (ii) the capacity to evacuate spatially variation commotion by indicating a non-uniform clamor level guide, and (iii) quicker speed than benchmark BM3D even on CPU without relinquishing denoising execution. Broad tests on engineered and genuine loud pictures are directed to assess FFD Net in correlation with cutting edge denoisers. The outcomes demonstrate that FFDNet is compelling and proficient, making it exceptionally alluring for pragmatic denoising applications.

Index Terms—Image denoising, convolutional neural networks, Gaussian noise, spatially variant noise

I. INTRODUCTION

The importance of image denoising in low level vision can be revealed from many aspects. First, noise corruption is inevitable during the image sensing process and it may heavily degrade the visual quality of acquired image. Removing noise from the observed image is an essential step in various image processing and computer vision tasks [1], [2]. Second, from the Bayesian perspective, image denoising is an ideal test bed for evaluating image prior models and optimization methods [3], [4], [5]. Last but not least, in the unrolled inference via variable splitting techniques, many image restoration problems can be addressed by sequentially solving a series of denoising sub problems, which further broadens the application fields of image denoising [6], [7], [8], [9].

As in many previous literature of image denoising [10], [11], [12], [13], in this paper we assume that the noise is additive white Gaussian noise (AWGN) and the noise level is given. In order to handle practical image denoising problems, a flexible image denoiser is expected to have the following desirable properties: (i) it is able to perform denoising using a single model; (ii) it is efficient, effective and user-friendly; and (iii) it can handle spatially variant noise. Such a denoiser can be directly deployed to recover the clean image when the noise level is known or can be well estimated. When the noise level is unknown or is difficult to estimate, the denoiser should allow the user to adaptively control the trade-off between noise reduction and detail preservation. Furthermore, the noise can be spatially variant and the denoiser should be flexible enough to handle spatially variant noise.

However, state-of-the-art image denoising methods are still limited in flexibility or efficiency. In general, image denoising methods can be grouped into two major categories, model based methods and discriminative learning based ones. Model based methods such as BM3D [11] and WNNM [5] are flexible in handling denoising problems with various noise levels, but they suffer from several drawbacks. For example, their optimization algorithms are generally time-consuming, and cannot be directly

used to remove spatially variant noise. Moreover, model-based methods usually employ hand-crafted image priors (e.g., sparsity [14], [15] and nonlocal self similarity [12], [13], [16]), which may not be strong enough to characterize complex image structures.

As an alternative, discriminative denoising methods aim to learn the underlying image prior and fast inference from a training set of degraded and ground-truth image pairs. One approach is to learn stage-wise image priors in the context of truncated inference procedure [17]. Another more popular approach is plain discriminative learning, such as the MLP [18] and convolutional neural network (CNN) based methods [19], [20], among which the DnCNN [20] method has achieved very competitive denoising performance. The success of CNN for image denoising is attributed to its large modeling capacity and tremendous advances in network training and design. However, existing discriminative denoising methods are limited in flexibility, and the learned model is usually tailored to a specific noise level. From the perspective of regression, they aim to learn a mapping function $x = F(y; \Theta\sigma)$ between the input noisy observation y and the desired output x. The model parameters $\Theta\sigma$ are trained for noisy images corrupted by AWGN with a fixed noise level σ, while the trained model with $\Theta\sigma$ is hard to be directly deployed to images with other noise levels. Though a single CNN model (i.e., DnCNN-B) is trained in [20] for Gaussian denoising, it does not generalize well to real noisy images and works only if the noise level is in the preset range, e.g., [0, 55]. Besides, all the existing discriminative learning based methods lack flexibility to deal with spatially variant noise.

To overcome the drawbacks of existing CNN based denoising methods, we present a fast and flexible denoising convolutional neural network (FFDNet). Specifically, our FFDNet is formulated as $x = F(y, M; \Theta)$, where M is a noise level map. In the DnCNN model $x = F(y; \Theta\sigma)$, the parameters $\Theta\sigma$ vary with the change of noise level σ, while in the FFDNet model, the noise level map is modeled as an input and the model parameters Θ are invariant to noise level. Thus, FFDNet provides a flexible way to handle different noise levels with a single network.

By introducing a noise level map as input, it is natural to expect that the model performs well when the noise level map matches the ground-truth one of noisy input. Furthermore, the noise level map should also play the role of controlling the trade-off between noise reduction and detail preservation. It is found that heavy visual quality degradation may be engendered when setting a larger noise level to smooth out the details. We highlight this problem and adopt a method of orthogonal initialization on convolutional filters to alleviate this. Besides, the proposed FFDNet works on downsampled sub-images, which largely accelerates the training and testing speed, and enlarges the receptive field as well.

Using images corrupted by AWGN, we quantitatively compare FFDNet with state-of-the-art denoising methods; including model-based methods such as BM3D [11] and WNNM [5] and discriminative learning based methods such as TNRD [17] and DnCNN [20]. The results clearly demonstrate the superiority of FFDNet in terms of both denoising performance and computational efficiency. In addition, FFDNet performs favorably on images corrupted by spatially variant AWGN. We further evaluate FFDNet on real-world noisy images, where the noise is often signal-dependent, non-Gaussian and spatially variant. The proposed FFDNet model still achieves perceptually convincing results by setting proper noise level maps. Overall, FFDNet enjoys high potentials for practical denoising applications.

The main contribution of our work is summarized as follows:

• A fast and flexible denoising network, namely FFDNet, is proposed for discriminative image denoising. By taking a tunable noise level map as input, a single FFDNet is able to deal with noise on different levels, as well as spatially variant noise.

• We highlight the importance to guarantee the role of the noise level map in controlling the trade-off between noise reduction and detail preservation.

• FFDNet exhibits perceptually appealing results on both synthetic noisy images corrupted by AWGN and real world noisy images, demonstrating its potential for practical image denoising.

The remainder of this paper is organized as follows. Sec. II reviews existing discriminative denoising methods. Sec. III presents the proposed image denoising model. Sec. IV reports the experimental results. Sec. V extention work of proposed work Sec. VI reports the extention results Sec.VII concludes the paper.

II.RELATED WORK

In this section, we briefly review and discuss the two major categories of relevant methods to this work, i.e., maximum a posteriori (MAP) inference guided discriminative learning and plain discriminative learning.

A. MAP Inference Guided Discriminative Learning

Instead of first learning the prior and then performing the inference, this category of methods aims to learn the prior parameters along with a compact unrolled inference through minimizing a loss function [21]. Following the pioneer work of fields of experts [3], Barbu [21] trained a discriminative Markov random field (MRF) model together with a gradient descent inference for image denoising. Samuel and Tappen [22] independently proposed a compact gradient descent inference learning framework, and discussed the advantages of discriminative learning over model-based optimization method with MRF prior. Sun and Tappen [23] proposed a novel nonlocal range MRF (NLR-MRF) framework, and employed the gradient-based discriminative learning method to train the model. Generally speaking, the methods above only learn the prior parameters in a discriminative manner, while the inference parameters are stage-invariant. With the aid of unrolled half quadratic splitting (HQS) techniques, Schmidt et al. [24], [25] proposed a cascade of shrinkage fields (CSF) framework to learn stage-wise inference parameters. Chen et al. [17] further proposed a trainable nonlinear reaction diffusion (TNRD) model through discriminative learning of a compact gradient descent inference step. Recently, Lefkimmiatis [26] and Qiao et al. [27] adopted a proximal gradient-based denoising inference from a variational model to incorporate the nonlocal self-similarity prior. It is worth noting that, apart from MAP inference, Vemulapalli et al. [28] derived an end-to-end trainable patch-based denoising network based on Gaussian Conditional Random Field (GCRF) inference. MAP inference guided discriminative learning usually requires much fewer inference steps, and is very efficient in image denoising. It also has clear interpretability because the discriminative architecture is derived from optimization algorithms such as HQS and gradient descent [17], [21], [22], [23], [24]. However, the learned priors and inference procedure are limited by the form of MAP model [25], and generally perform inferior to the state-of-the-art CNN-based denoisers. For example, the inference of CSF [24] is not very flexible since it is strictly derived from the HQS optimization under the field of experts (FoE) framework. The capacity of FoE is however not large enough to fully characterize image priors, which in turn makes CSF less effective. For these reasons, Kruse et al. [29] generalized CSF for better performance by replacing some modular parts of unrolled inference with more powerful CNN.

B. Plain Discriminative Learning

Instead of modeling image priors explicitly, the plain discriminative learning methods learn a direct mapping function to model image prior implicitly. The multi-layer perceptron (MLP) and CNNs have been adopted to learn such priors.

The use of CNN for image denoising can be traced back to [19], where a five-layer network with sigmoid nonlinearity was proposed. Subsequently, auto-encoder based methods have been suggested for image denoising [30], [31]. However, early MLP and CNN-based methods are limited in denoising performance and cannot compete with the benchmark BM3D method [11]. The first discriminative denoising method which achieves comparable performance with BM3D is the plain MLP method proposed by Burger et al. [18]. Benefitted from the advances in deep CNN, Zhang et al. [20] proposed a plain denoising CNN (DnCNN) method which achieves state-of-the-art denoising performance. They showed that residual learning and batch normalization [32] are particularly useful for the success of denoising. For a better trade-off between accuracy and speed, Zhang et al. [9] introduced a 7-layer denoising network with dilated convolution [33] to expand the receptive field of CNN. Mao et al. [34] proposed a very deep fully convolutional encoding-decoding network with symmetric skip connection for image denoising. Santhanam et al. [35] developed a recursively branched de convolutional network (RBDN) for image denoising as well as generic image-to-image regression. Tai et al. proposed a very deep persistent memory network (MemNet) by introducing a memory block to mine persistent memory through an adaptive learning process. Plain discriminative learning has shown better performance than MAP inference guided

discriminative learning; however, existing discriminative learning methods have to learn multiple models for handling images with different noise levels, and are incapable to deal with spatially variant noise. To the best of our knowledge, it remains an unaddressed issue to develop a single discriminative denoising model which can handle noise of different levels, even spatially variant noise, in a speed even faster than BM3D.

III. Proposed Fast and Flexible Discriminative CNN Denoiser

We present a single discriminative CNN model, namely FFDNet, to achieve the following three objectives:

• Fast speed: The denoiser is expected to be highly efficient without sacrificing denoising performance.

• Flexibility: The denoiser is able to handle images with different noise levels and even spatially variant noise.

• Robustness: The denoiser should introduce no visual artifacts in controlling the trade-off between noise reduction and detail preservation.

In this work, we take a tunable noise level map M as input to make the denoising model flexible to noise levels. To improve the efficiency of the denoiser, a reversible downsampling operator is introduced to reshape the input image of size W×H×C into four downsampled sub-images of size W/2 × H/2 ×4C. Here C is the number of channels, i.e., C = 1 for grayscale image and C = 3 for color image. In order to enable the noise level map to robustly control the trade-off between noise reduction and detail preservation by introducing no visual artifacts, we adopt the orthogonal initialization method to the convolution filters.

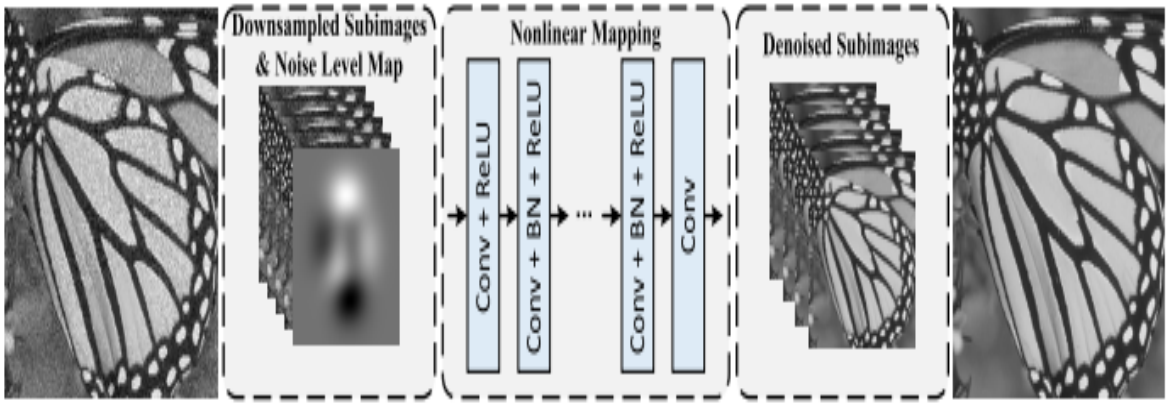


Fig.1. The architecture of the proposed FFDNet for image denoising. The input image is reshaped to four sub-images, which are then input to the CNN together with a noise level map. The final output is reconstructed by the four denoised sub-images.

A. Network Architecture

Fig. 1 illustrates the architecture of FFDNet. The first layer is a reversible downsampling operator which reshapes a noisy image y into four downsampled sub-images. We further concatenate a tunable noise level map M with the downsampled sub-images to form a tensor ˜y of size W 2 × H 2 × (4C + 1) as the inputs to CNN. For spatially invariant AWGN with noise level σ, M is a uniform map with all elements being σ. With the tensor ˜y as input, the following CNN consists of a series of 3 × 3 convolution layers. Each layer is composed of three types of operations: Convolution (Conv), Rectified Linear Units (ReLU), and Batch Normalization (BN) [32]. More specifically, "Conv+ReLU" is adopted for the first convolution layer, "Conv+BN+ReLU" for the middle layers, and "Conv" for the last convolution layer. Zero-padding is employed to keep the size of feature maps unchanged after each convolution. After the last convolution layer, an up scaling operation is applied as the reverse operator of the downsampling

operator applied in the input stage to produce the estimated clean image ˆx of size W × H × C. Different from DnCNN; FFDNet does not predict the noise. The reason is given in Sec. III-F. Since FFDNet operates on downsampled sub-images, it is not necessary to employ the dilated convolution [33] to further increase the receptive field.

By considering the balance of complexity and performance, we empirically set the number of convolution layers as 15 for grayscale image and 12 for color image. As to the channels of feature maps, we set 64 for grayscale image and 96 for color image. The reason that we use different settings for grayscale and color images is twofold. First, since there are high dependencies among the R, G, and B channels, using a smaller number of convolution layers encourage the model to exploit the inter-channel dependency. Second, color image has more channels as input, and hence more features (i.e. more channels of feature map) is required. According to our experimental results, increasing the number of feature maps contributes more to the denoising performance on color images. Using different settings for color images, FFDNet can bring an average gain of 0.15dB by PSNR on different noise levels. As we shall see from Sec. IV-F, 12-layer FFDNet for color image runs slightly slower than 15-layer FFDNet for grayscale image. Taking both denoising performance and efficiency into account, we set the number of convolution layers as 12 and the number of feature maps as 96 for color image denoising.

B. Noise Level Map

Let's first revisit the model-based image denoising methods to analyze why they are flexible in handling noises at different levels, which will in turn help us to improve the flexibility of CNN-based denoiser. Most of the model-based denoising methods aim to solve the following problem

$$\hat{x} = \arg\min_x \frac{1}{2\sigma^2}\|y - x\|^2 + \lambda\Phi(x), \qquad (1)$$

where $\frac{1}{2\sigma^2}\|y - x\|^2$ is the data fidelity term with noise level σ, Φ(x) is the regularization term associated with image prior, and λ controls the balance between the data fidelity and regularization terms. It is worth noting that in practice λ governs the compromise between noise reduction and detail preservation. When it is too small, much noise will remain; on the opposite, details will be smoothed out along with suppressing noise. With some optimization algorithms, the solution of Eqn. (1) actually defines an implicit function given by

$$\hat{x} = \mathcal{F}(y, \sigma, \lambda; \Theta). \qquad (2)$$

Since λ can be absorbed into σ, Eqn. (2) can be rewritten as

$$\hat{x} = \mathcal{F}(y, \sigma; \Theta). \qquad (3)$$

In this sense, setting noise level σ also plays the role of setting λ to control the trade-off between noise reduction and detail preservation. In a word, model-based methods are flexible in handling images with various noise levels by simply specifying σ in Eqn. (3).

According to the above discussion, it is natural to utilize CNN to learn an explicit mapping of Eqn. (3) which takes the noise image and noise level as input. However, since the inputs y and σ have different dimensions, it is not easy to directly feed them into CNN. Inspired by the patch based denoising methods which actually set σ for each patch, we resolve the dimensionality mismatching problem by stretching the noise level σ into a noise level map M. In M, all the elements are σ. As a result, Eqn. (3) can be further rewritten as

$$\hat{x} = \mathcal{F}(y, \boldsymbol{M}; \Theta). \qquad (4)$$

It is worth emphasizing that M can be extended to degradation maps with multiple channels for more general noise models such as the multivariate (3D) Gaussian noise model N (0, Σ) with zero mean and covariance matrix Σ in the RGB color space. As such, a single CNN model is expected to inherit the flexibility of handling noise model with different parameters, even spatially variant noises by noting M can be non-uniform.

C. Denoising on Sub-images

Efficiency is another crucial issue for practical CNN-based denoising. One straightforward idea is to reduce the depth and number of filters. However, such a strategy will sacrifice much the modeling capacity and receptive field of CNN [20]. In [9], dilated convolution is introduced to expand receptive field without the increase of network depth, resulting in a 7- layer denoising CNN. Unfortunately, we empirically find that FFDNet with dilated

convolution tends to generate artifacts around sharp edges. Shi et al. proposed to extract deep features directly from the low-resolution image for super-resolution, and introduced a sub-pixel convolution layer to improve computational efficiency. In the application of image denoising, we introduce a reversible downsampling layer to reshape the input image into a set of small sub-images. Here the downsampling factor is set to 2 since it can largely improve the speed without reducing modeling capacity. The CNN is deployed on the sub-images, and finally a sub-pixel convolution layer is adopted to reverse the downsampling process. Denoising on downsampled sub-images can also effectively expand the receptive field which in turn leads to a moderate network depth. For example, the proposed network with a depth of 15 and 3 × 3 convolution will have a large receptive field of 62 × 62. In contrast, a plain 15-layer CNN only has a receptive field size of 31×31. We note that the receptive field of most state-of-the-art denoising methods ranges from 35×35 to 61×61. Further increase of receptive field actually benefits little in improving denoising performance [40]. What is more, the introduction of sub sampling and sub-pixel convolution is effective in reducing the memory burden. Experiments are conducted to validate the effectiveness of downsampling for balancing denoising accuracy and efficiency on the BSD68 dataset with $\sigma = 15$ and 50. For grayscale image denoising, we train a baseline CNN which has the same depth as FFDNet without downsampling. The comparison of average PSNR values is given as follows: (i) when σ is small (i.e., 15), the baseline CNN slightly outperforms FFDNet by 0.02dB; (ii) when σ is large (i.e., 50), FFDNet performs better than the baseline CNN by 0.09dB. However, FFDNet is nearly 3 times faster and is more memory-friendly than the baseline CNN. As a result, by performing denoising on sub-images, FFDNet significantly improves efficiency while maintaining denoising performance.

D. Examining the Role of Noise Level Map

By training the model with abundant data units (y, M; x), where M is exactly the noise level map of y, the model is expected to perform well when the noise level matches the ground-truth one (see Fig. 2(a)). On the other hand, in practice, we may need to use the learned model to smooth out some details with a higher noise level map than the ground-truth one (see Fig. 2(b)). In other words, one may take advantage of the role of λ to control the trade-off between noise reduction and detail preservation. Hence, it is very necessary to further examine whether M can play the role of λ.

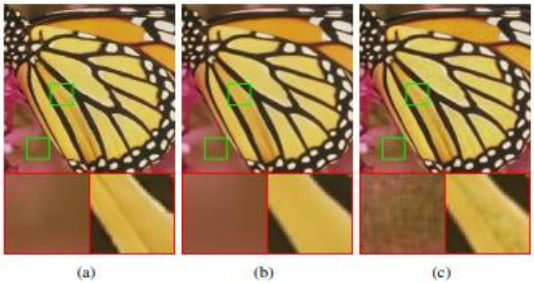


Fig.2. an example to show the importance of guaranteeing the role of noise level map in controlling the trade-off between noise reduction and detail preservation. The input is a noisy image with noise level 25. (a) Result without visual artifacts by matched noise level 25. (b) Result without visual artifacts by mismatched noise level 60. (c) Result with visual artifacts by mismatched noise level 60.

Unfortunately, the use of both M and y as input also increases the difficulty to train the model. According to our experiments on several learned models, the model may give rise to visual artifacts especially when the input noise level is much higher than the ground-truth one (see Fig. 2(c)), which indicates M fails to control the trade-off between noise reduction and detail preservation. Note that it does not mean all the models suffer from such problem. One possible solution to avoid this is to regularize the convolution filters. As a widely-used regularization method, orthogonal regularization has proven to be effective in eliminating the correlation between convolution filters, facilitating gradient propagation and improving the compactness of the learned model. In addition, recent studies have demonstrated the advantage of orthogonal regularization in enhancing the network generalization ability in applications of deep hashing and image classification. According to our experiments, we empirically find that the orthogonal initialization of the convolution filters, works well in suppressing the above mentioned visual artifacts. It is worth emphasizing that this section aims to highlight the necessity of guaranteeing the role of M in controlling the trade-off between noise reduction and detail preservation

rather than proposing a method to avoid the possible visual artifacts caused by noise level mismatch. In practice, one may retrain the model until M plays its role and results in no visual artifacts with a lager noise level.

IV.EXPERIMENTS&RESULTS

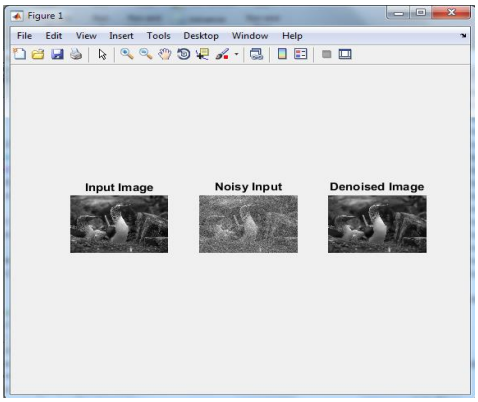


Fig.1: Denoising For Grayscale Images

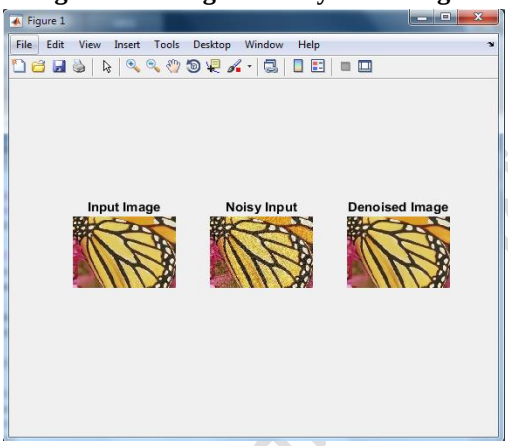


Fig.2: Denoising For Color Images

V.EXTENTION METHOD

For this project I have done extention work on videos by using a fast and flexible denoising convolutional neural network. In this process first the video will be dividing into frames and then store that frames into one file. Then the video frames will be taking as input then apply a fast and flexible denoising convolutional neural network, namely FFDNet, with a tunable noise level map. This extention work FFDNet on downsampled, achieving a good trade-off between inference speeds and denoising performance. Compare to proposed work the extention work give better denoising performance. In proposed psnr value will be 29db and extention work psnr value will be 33db.based on the psnr value the extention work will give better results compare to proposed work.

VI.EXTENTION RESULTS

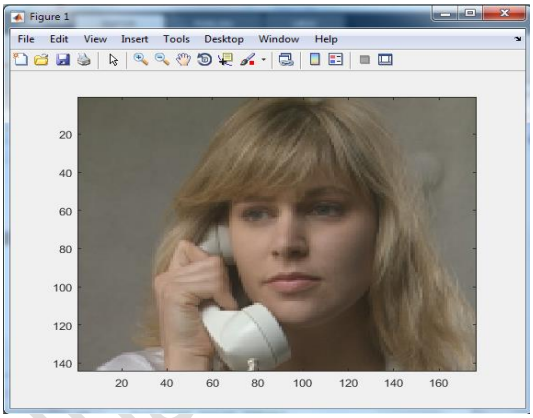


Fig.1: Input Video

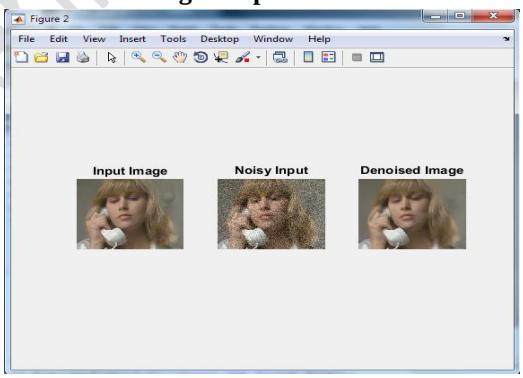
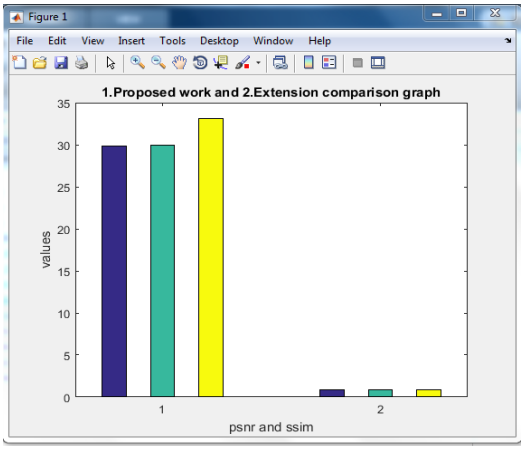


Fig.2: Denoising On Video Frames

Comparison Graph



Comparison Table

S.NO	PARAMETERS	GRAY IMAGES FOR PROPOSED METHOD	COLOR IMAGES FOR PROPOSED METHOD	EXTENTION MEHOD
1	PSNR	29.8995	29.9847	33.1184
2	SSIM	0.8316	0.9175	0.8927

VII. Conclusion&Future work

In this paper, I proposed another CNN model, specifically FFDNet, for quick, successful and adaptable discriminative denoising. To accomplish this objective, a few strategies were used in system plan and preparing, for example, the utilization of commotion level guide as input and denoising in down sampled sub-pictures space. The results on engineered pictures with AWGN showed that FFDNet cannot just deliver cutting edge results when information commotion level matches ground-truth clamor level, yet in addition can heartily control the exchange off between commotion decrease and detail safeguarding. The outcomes on pictures with spatially variation AWGN approved the adaptability of FFDNet for giving inhomogeneous commotion. The outcomes on genuine boisterous pictures additionally showed that FFDNet can convey perceptually engaging denoising results. At long last, the running time correlations demonstrated the quicker speed of FFDNet over other contending strategies, for example, BM3D. Considering its adaptability, productivity and adequacy, FFDNet gives a down to earth answer for CNN denoising applications.

References

[1] H. C. Andrews and B. R. Hunt, "Digital image restoration," Prentice Hall Signal Processing Series, Englewood Cliffs: Prentice-Hall, 1977, vol. 1, 1977.

[2] P. Chatterjee and P. Milanfar, "Is denoising dead?" IEEE Transactions on Image Processing, vol. 19, no. 4, pp. 895–911, 2010.

[3] S. Roth and M. J. Black, "Fields of experts: A framework for learning image priors," in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, 2005, pp. 860–867.

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

[5] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2862–2869.

[6] M. V. Afonso, J. M. Bioucas-Dias, and M. A. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," IEEE Transactions on Image Processing, vol. 19, no. 9, pp. 2345–2356, 2010.

[7] F. Heide, M. Steinberger, Y.-T. Tsai, M. Rouf, D. Pajak, D. Reddy, O. Gallo, J. Liu, W. Heidrich, K. Egiazarian et al., "FlexISP: A flexible camera image processing framework," ACM Transactions on Graphics, vol. 33, no. 6, p. 231, 2014.

[8] Y. Romano, M. Elad, and P. Milanfar, "The little engine that could: Regularization by denoising (RED)," submitted to SIAM Journal on Imaging Sciences, 2016.

[9] 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, 2017, pp. 3929–3938.

[10] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," IEEE Transactions on Image processing, vol. 12, no. 11, pp. 1338–1351, 2003.

[11] 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.

[12] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in IEEE International Conference on Computer Vision, 2009, pp. 2272–2279.

[13] 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.

[14] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," IEEE Transactions on Image Processing, vol. 15, no. 12, pp. 3736–3745, 2006.
[15] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," IEEE Transactions on Image Processing, vol. 17, no. 1, pp. 53–69, 2008.
[16] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, 2005, pp. 60–65.
[17] Y. Chen and T. Pock, "Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1256– 1272, 2017.
[18] H. C. Burger, C. J. Schuler, and S. Harmeling, "Image denoising: Can plain neural networks compete with BM3D?" in IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 2392–2399.
[19] V. Jain and S. Seung, "Natural image denoising with convolutional networks," in Advances in Neural Information Processing Systems, 2009, pp. 769–776.
[20] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," IEEE Transactions on Image Processing, vol. 26, no. 7, pp. 3142–3155, July 2017.
[21] A. Barbu, "Training an active random field for real-time image denoising," IEEE Transactions on Image Processing, vol. 18, no. 11, pp. 2451– 2462, 2009.
[22] K. G. Samuel and M. F. Tappen, "Learning optimized MAP estimates in continuously-valued MRF models," in IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 477–484.
[23] J. Sun and M. F. Tappen, "Learning non-local range markov random field for image restoration," in IEEE Conference on Computer Vision and Pattern Recognition, 2011, pp. 2745–2752.
[24] U. Schmidt and S. Roth, "Shrinkage fields for effective image restoration," in IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2774–2781.
[25] U. Schmidt, "Half-quadratic inference and learning for natural images," Ph.D. dissertation, Technische Universitat, Darmstadt, 2017. [Online]. ¨ Available: http://tuprints.ulb.tu-darmstadt.de/6044/.
[26] S. Lefkimmiatis, "Non-local color image denoising with convolutional neural networks," in IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 3587–3596.
[27] P. Qiao, Y. Dou, W. Feng, R. Li, and Y. Chen, "Learning non-local image diffusion for image denoising," in Proceedings of the 2017 ACM on Multimedia Conference, 2017, pp. 1847–1855.
[28] R. Vemulapalli, O. Tuzel, and M.-Y. Liu, "Deep gaussian conditional random field network: A model-based deep network for discriminative denoising," in IEEE Conference on Computer Vision and Pattern Recognition, June 2016.
[29] 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, Oct 2017.
[30] J. Xie, L. Xu, and E. Chen, "Image denoising and in painting with deep neural networks," in Advances in Neural Information Processing Systems, 2012, pp. 341–349.
[31] F. Agostinelli, M. R. Anderson, and H. Lee, "Robust image denoising with multi-column deep neural networks," in Advances in Neural Information Processing Systems, 2013, pp. 1493–1501.
[32] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in International Conference on Machine Learning, 2015, pp. 448–456.
[33] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," in International Conference on Learning Representations, 2016.
[34] X. Mao, C. Shen, and Y.-B. Yang, "Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections," in Advances in Neural Information Processing Systems, 2016, pp. 2802–2810.
[35] V. Santhanam, V. I. Morariu, and L. S. Davis, "Generalized deep image to image regression," in IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 5609–5619.