Covid-19 Images and Video Denoising Algorithms Based on Convolutional Neural Network CNNs

Zoubeida MESSALI(1,2)*, Salsabil SAAD SAOUD(1), Amira LAMRECHE(1)
(1) Electronics Department, Faculty of Technology, University of Mohamed El Bachir El Ibrahimi - Bordj Bou Arréridj, El-Anasser, BBA, Algeria
(2) ETA Laboratory, El-Anasser, BBA, Algeria
*messalizoubeida@univ.bba

Abstract: In this paper, the most sophisticated denoising algorithms of images and video are applied and implemented. More precisely, we study and implement the video denoising algorithms "VBM3D", "VBM4D", "DVDNet" and "FastDVDnet". Much attention is given to the latest DVDNet and Fast DVDNet algorithms, which are based on CNN. We carry out a detailed quantitative and qualitative comparative study between the considered algorithms. Two assessments are adapted; the first is a qualitative comparison based on the quality of the images / videos and the second is quantitative in terms of PSNR and running time criteria. To see the direct impact of our study on the current pandemic, and to show the importance of image and video pre-processing algorithms in the field of medical imaging; we apply the considered denoising algorithms based on CNN on our built COVID- 19 dataset and TEST_PCR videos.

Keywords: Denoising, Deep learning, CNN, Covid-19, Gaussian noise, PSNR.

1. INTRODUCTION

Video denoising appears to be a largely under-explored field. Recently, new image and video denoising methods based on deep learning techniques have gained attention because of their outstanding performance.

Machine Learning especially deep learning, is the science of art that studies how to develop algorithms of Learning from big data.

In this study, we are interested in improving video denoising algorithms such as VBM3D and VBM4D [1] which are based on researching similar patches in the same window. we adapt Convolutional Neural Networks (CNNs), CNN and FastCNN algorithms.

Their performance compares favorably with the considered state-of-the-art image denoising algorithms [1,2], both quantitatively and visually.

One of the remarkable features of deep learning that we want to exploit is the ability to denoise multiple levels of noise by training a single model rather than training multiple models.CNN denoising algorithm will be evaluated in terms of PSNR, running time and details preservation.

This paper is organized as follows: In Section 2, a brief review of VBM3D and VBM4D video denoising algorithms, is given. In Section 3 we describe the implemented denoising scheme based on CNNs. Experimental results and Performance assessment study are reported in section 4. Section 5 concludes the paper.

2. RELATED WORKS

Image denoising has enjoyed consistent and large use in past years. However, video denoising has received less attention than still image denoising in the literature. Furthermore, CNNs in image and video applications have taken over the state-of-the-art [3-10].The first denoising approaches using neural networks were proposed in the mid and late 2000s.

Deep video denoising state-of-the-art based on a convolutional neural network (CNN) architecture has been introduced in [6,9]. The approach introduced in [6] is an enhanced version of DVDnet algorithm with reduced computing times.

One of the most effective and reliable methods in video and image denoising are the methods based on researching similar patches in the same window, like VBM3D and VBM4D denoising algorithms [1,2], which we have implemented on our dataset COVID_19 and TESTPCR to establish a quantitative comparison with the proposed CNN algorithms.

2.1 Video Block Matching and 3D Filtering (VBM3D)

VBM3D algorithm has three important concepts, namely: clustering, collaborative filtering, and aggregation. The general procedure has two steps. Figure 1 shows the details of the algorithm [3].

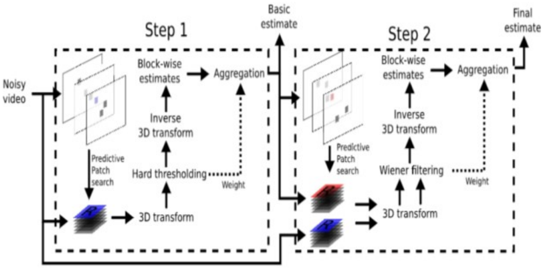


Fig.1 Diagram of the heart of the VBM3D algorithm [3]

2.2VBM4D: Video Block Matching and 4D Filtring

VBM4D [1] is an improved version of VBM3D. Unlike VBM3D, VBM4D does not group blocks, but Spatio-temporal volumes that are mutually similar according to a non-local search procedure. Therefore, the groups in VBM4D are 4D stacks of 3D volumes, and the collaborative filtering is then performed via a separable 4D space-time transform [1]. We should note that collaborative filtering consists on using several types of filtering for the same algorithm. These non-local patch-based methods were until recently at the forefront of image denoising but are now surpassed by CNNs.

3.CONTRIBUTION

In this paper, we propose a new and efficient way to feed video auto-similarities to CNN. The non-locality is incorporated into the network via a first non-scalable layer which finds for each patch in the input image its most similar patches in a search region. The central values of these patches are then gathered into a feature vector which is assigned to each image pixel. This information is presented to a CNN trained to predict the denoised image. We apply the proposed architecture to both image denoising and video denoising. For the latter, the corrective measures are sought in a 3D spatiotemporal volume. The proposed architecture obtains cutting edge results based on [4].

We have implemented two algorithms (DVDnet [4] and Fast DVDnet [5]) applied on our dataset COVID_19 and TESTPSR with their two versions: RGB and grayscale.

The noisy images and videos are modeled as follows:

$$\boldsymbol{V_i} = \boldsymbol{U_i} + \varepsilon_i \qquad (1)$$

Where : $\boldsymbol{V_i}$ **:**the noisy observations

$\boldsymbol{U_i}$ **:** the original signal

ε_i **:** the noise.

In our work, two types of noise are considered: Gaussian and Poisson noise. Five standard deviations σ of noise are tested (10, 20, 30,40, and 50).

Video Data set

In our dataset, we have chosen two videos (.MP4) on the topic of CORONAVIRUS news. The first one we called COVID-19 and the second is TESTPCR. We extracted four frames numbered from 1 to 4 (extension jpg) for use with VBM3D and DVDnet, then we gathered them in a short video of one second (1s) and converted this last one to .avi extension with the software Format Factory 5.4.5.1. The video test sequence contains 30 frames of 720×1280 pixels for RGB and 480×720 pixels for Gray Scale.

3.1 A Fast Network for Deep Video Denoising DVDnet

This video denoising algorithm is based on the Convolutional Neurol Network CNN, called DVDnet. The algorithm achieves results comparable to or better than those of state-of-the-art algorithms based on searching similar patches in the same window, butwith fast running time. The noise is considerably reduced with preserving the details in the frames [4]. Figure 2 illustrates the main steps of this algorithm.

To implement this algorithm, we performed the following steps:

1. Look for similar patches throughout the noisy video footage.
2. Divide the denoising process into two steps: $2T + 1$ frames after $2T$ frames,
3. Perform spatial denoising (for $2T + 1$ and $2T$ frames).
4. Alignment of the frames.
5. Perform temporal denoising (for $2T + 1$ and $2T$ frames).

Temporal and SpatialDenoising

The characteristics of spatial and temporal denoising blocks result from the trade-off between denoising and detail preservation. The two blocks are implemented as feed-forward networks, as shown in Figures3 and 4 respectively [4].

3.2 Fast DVDnet Towards Real-Time DeepVideo Denoising Without Flow Estimation

This algorithm is based on DVDnet [6], but at the same time introduces some important changes compared to its predecessor [4]. Figure 5 illustrates the detailed flowchart of the Fast DVDnet algorithm.

Algorithm:

1. Five consecutive images are used to denoise the middle image.
2. The encoder (denoising blocks 1 and 2 as shown in Figure 6) has been adapted to take three images and a noise map as inputs.
3. Upsampling in the decoder is done with a Pixel Shuffle [7] layer, which helps reduce grid artifacts.
4. The merging of the characteristics of the encoder with those of the decoder is done with pixel-by-pixel addition operation instead of a channel concatenation. This translates into reduced memory requirements
5. Blocks implement residual learning with a residual connection between the central noisy input frame and the output, which has been observed to facilitate the learning process [4]

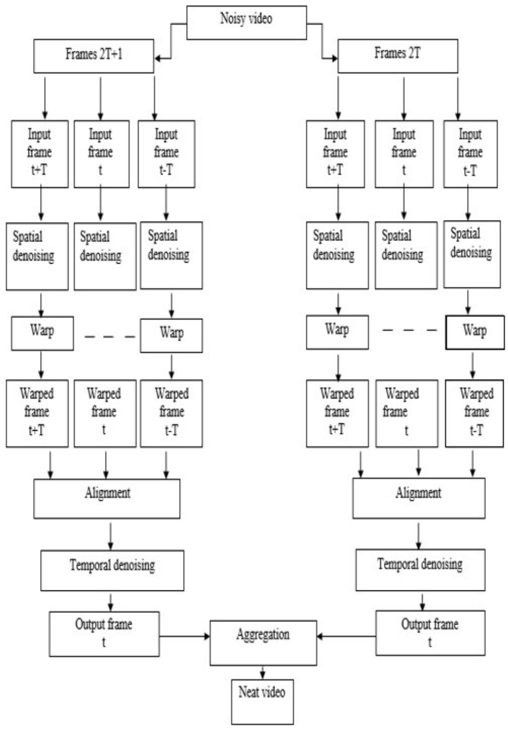


Fig. 2 Flowchart of the of the DVDnet algorithm

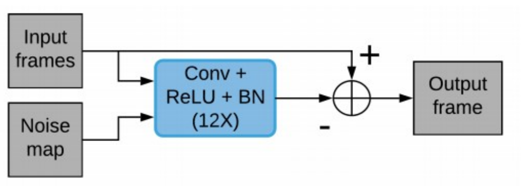


Fig. 3 Spatial denoising [4]

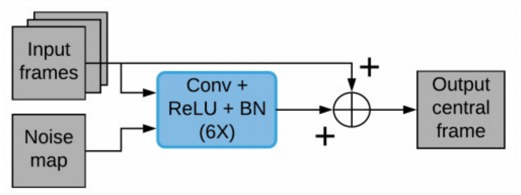


Fig. 4 Temporal denoising [4]

ReLU : point-wise activation functions [8,9], **ReLU**(·) = max(·, 0)
BN : normalisation per batch
12X : convolutional layers
6X : convolutional layers
For the sake of clarity, Figures 6 and 7 summarize the main steps of FastDVDnet and the architecture of the denoising block 1 and 2 respectively.

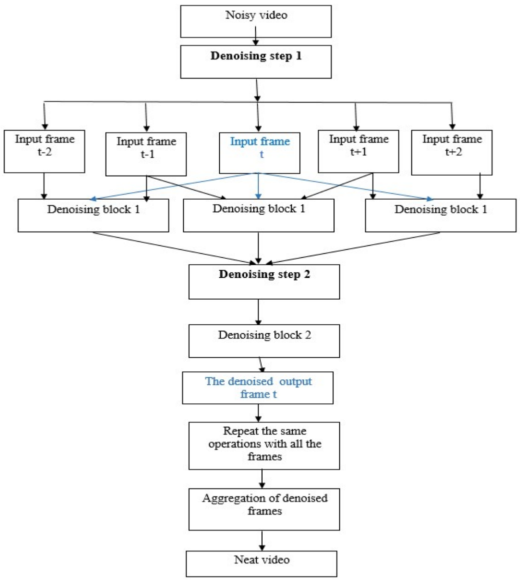


Fig. 5 Flowchart of the FastDVDnet algorithm

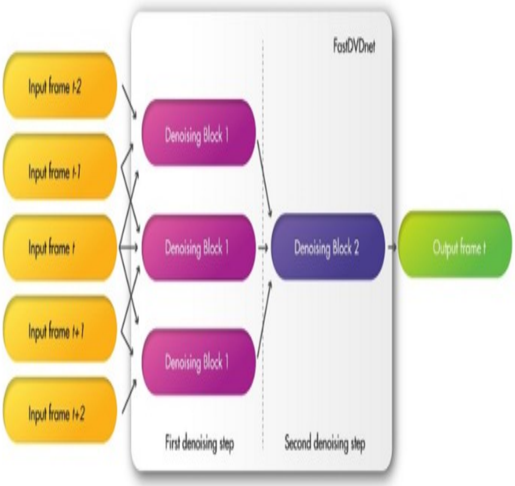


Fig.6 High-level diagram of FastDVDnet architecture

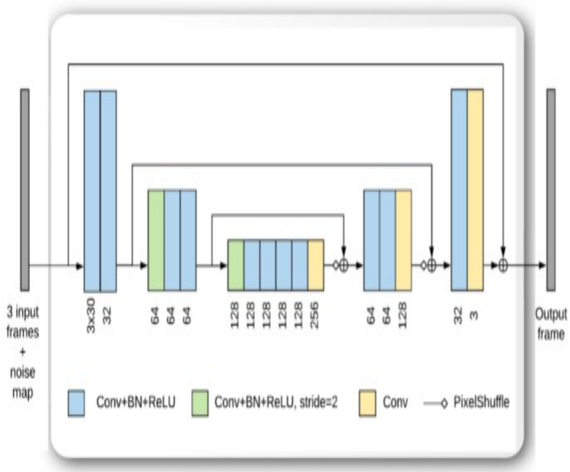


Fig.7 The architecture of the denoising block 1 and 2 [6]

4. EXPERIMENTAL RESULTS

4.1 Quantitative Assessment DVDnet/VBM3D/VBM4D with Gaussian Noise

In order to evaluate quantitatively the considered video CNN denoising algorithms, we have computed PSNR criterium.
DVDnet, VBM3D and VBM4D are applied on our COVID-19 dataset. The three denoising algorithms are compared in terms of PSNR for different noise levels. The obtained results are summarized in Table1.

Table 1 PSNR (dB) vs σ ofVBM3D, VBM4D and DVDnet, test video: COVID-19 dataset

σ	Method	COVID-19	
		RGB	grayscale
10	V-BM3D	24.81	24.47
	V-BM4D	**42.85**	**44.46**
	DVDnet	41.10	38.99
20	V-BM3D	22.94	24.16
	V-BM4D	**39.12**	**39.75**
	DVDnet	38.23	35.54
30	V-BM3D	21.77	23.97
	V-BM4D	**36.57**	**36.76**
	DVDnet	36.40	33.56
40	V-BM3D	20.79	23.77
	V-BM4D	34.60	**34.60**
	DVDnet	**35.02**	32.22
50	V-BM3D	20.08	23.62
	V-BM4D	33.03	**32.91**
	DVDnet	**33.89**	31.10

The corresponding curve is given in Figure 8.

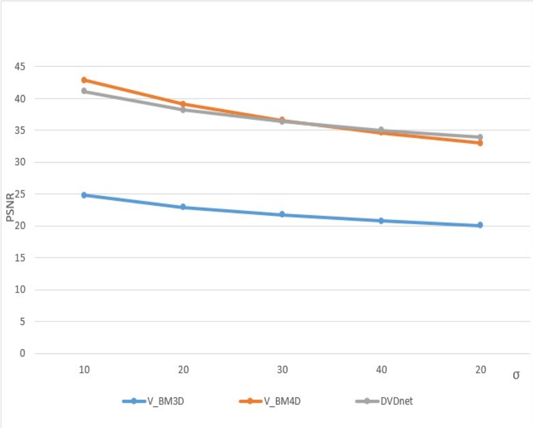


Fig.8 PSNR vs σ of DVDnet, V-BM3D, and V-BM4D algorithms applied to COVID-19 dataset

We have also calculated the running time of the three algorithms applied on both noisy Gray-Scale (GS) and RGB TESTPCR dataset with σ=10. Note that, we ran our experiments on a simple computer (Intel I3-5005U with 4Gb of ram). The obtained results are reported in Table 2.

Table 2 Running time (s) of VBM3D, VBM4D and DVDnet algorithms, σ = 10, test videos:GS & RGB TESTPCR dataset.

Method	V-BM3D		V-BM4D		DVDnet	
	grayscale	RGB	grayscale	RGB	grayscale	RGB
Running time (s)	80.23	9	128	375.2	24.30	142.98

The corresponding obtained images/videos are illustrated in Figures9 and 10 respectively.

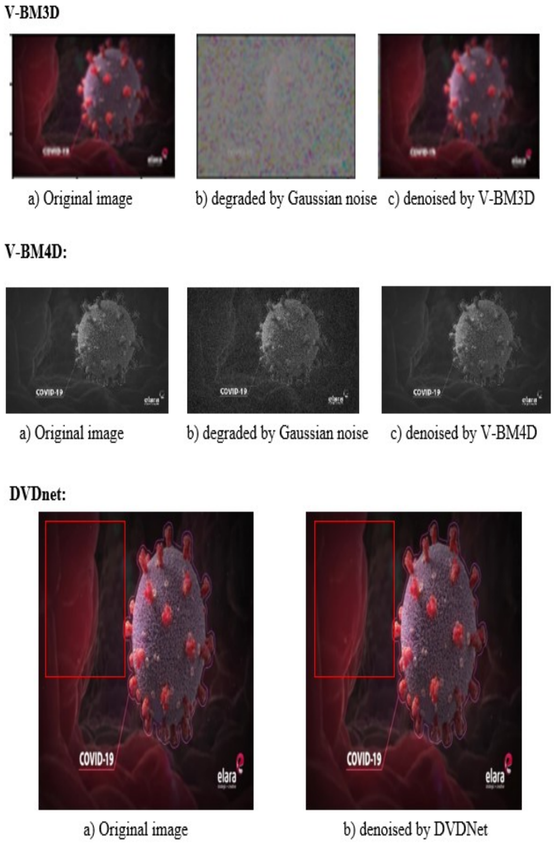


Fig. 9 Denoised frames by VBM3D, VBM4D, and DVDnet Gaussian noise level σ = 10, RGB COVID-19 sequence.

Fast DVDnet is applied on RGB TESTPCR dataset with Gaussian noise. Figure 11 shows the corresponding frame video.
Table 3 shows PSNR (dB) values vs noise level. The corresponding curves are illustrated in Figure 12, the running time is reported in Table 4.

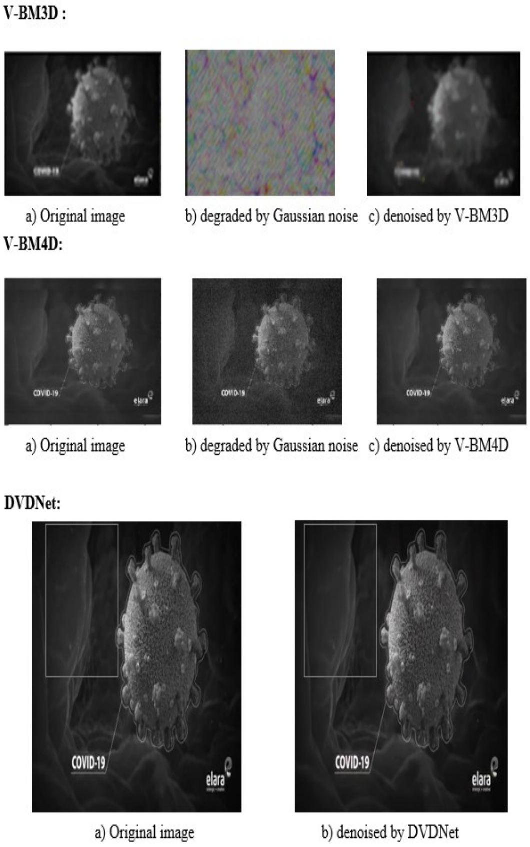


Fig. 10 Denoised frames by VBM3D, VBM4D, and DVDnet, Gaussian noise level σ = 10, GS COVID-19 sequence.

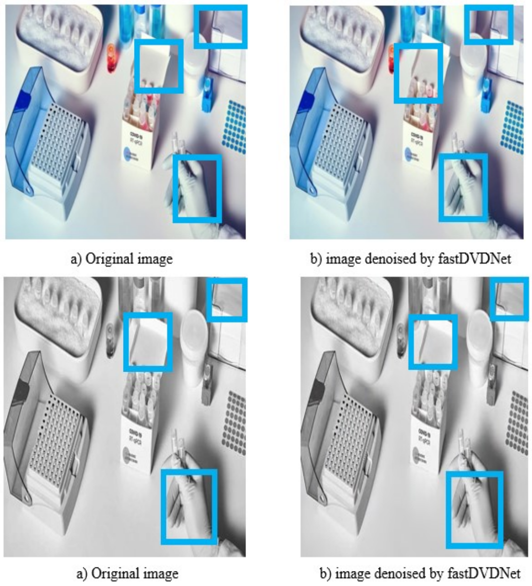


Fig. 11 Denoised frames by FastDVDnet σ = 10, Gaussian noisy RGB TESPCR dataset.

Table 3 PSNR vs σ of FastDVDnet , test video TESTPCR dataset

σ	Method	TESTPCR	
		RGB	grayscale
10	fastDVDNet	**42.67**	41.44
20	fastDVDNet	**39.57**	37.38
30	fastDVDNet	**37.61**	35.14
40	fastDVDNet	**36.18**	33.45
50	fastDVDNet	**34.95**	32.07

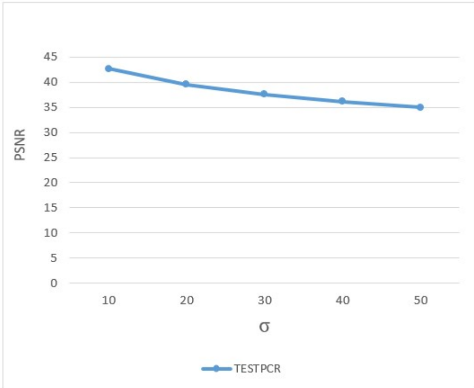


Fig.12 PSNR vs σ of FastDVDnet applied to RGB TESTPCR dataset

Table 4Running time of Fast DVDnet, σ = 10, test videos: GS & RGB TESTPCR dataset

Quantitative results	TESTPCR	
	RGB	grayscale
Running time(s)	159.8	**25.47**

DVDnet and FastDVDnet with Poisson noise

DVDnet and FastDVDnet are also tested in presence of Poisson noise, in terms of both PSNR and running time. Tables 5 and 6 report the obtained results respectively.

Table 5Performance of DVDnet applied on Poisson noisy GS/RGB COVID-19 data set

Quantitative results	COVID-19	
	RGB	grayscale
Running time(s)	198.94	**39.20**
PSNR (dB)	32.67	**33.19**

Table 6 Performance of FastDVDnet applied on Poisson noisy GS/RGB TESTPCR data set

Quantitative results	TESTPCR	
	RGB	grayscale
Running time(s)	206.11	**47.42**
PSNR (dB)	27.82	**29.30**

We have highlighted the best performance in bold.

4.2 Qualitative Assessment

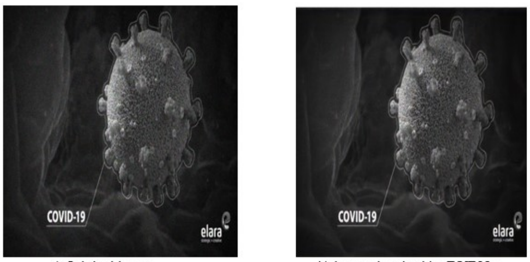
DVDnet/VBM3D/VBM4D/FastDVDnet with Gaussian noise

The visual quality of the obtained denoised frame videos by the considered denoising algorithms VBM3, VBM4D, DVDnet and FastDVDnet, is illustrated in Figures 9, 10 and 11 respectively.We have applied FastDVDnet on RGB TESTPCR frames.
A supplementary material is sent with this paper to show a demo of the denoised videos.
We note that COVID-19 dataset frames highlight results on stationary and moving parts of the video.
Since FastDVDnet is a too strong algorithm in denoising, no apparent difference appeared in the two images; this is why we have framed the differences.

DVDnet and FastDVDnet with Poisson noise

To complete the performance assessment of the CNN algorithms, we have also applied DVDnet and FastDVDnet on the frames corrupted by Poisson noise as shown in Figures 13 and 14 respectively.

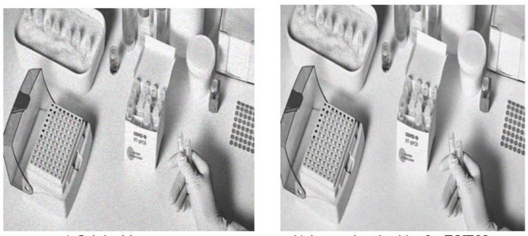
a) Original image b) image denoised by DVDNet



a) Original image b) image denoised by DVDNet

Fig. 13 Denoised frames by DVDnet, Poisson noisy GS and RGB COVID-19 sequences.

a) Original image b) image denoised by fastDVDNet



a) Original image b) image denoised by fastDVDNet

Fig. 14 Denoised frames by FastDVDnet, Poisson noisy GS and RGB TESPCR dataset.

Analyzing and Discussions

In this work, we have implemented denoising algorithms based on the convolutional network architecture compared with VBM3D and VBM4D algorithms, and we concluded the following:

1. BothDVDnet and FastDVDnet outperform VBM3D and VBM4D in terms of running time, PSNR, and preservation of image details.
2. Both DVDnet and FastDVDnet can handle a wide range of noise level(σ = 10, 20, 30, 40, 50)
3. Both DVDnet and FastDVDnet have the flexibility to deal with many types of noise
4. The running time of DVDnet is shorter than the running time of FastDVDnet
5. Both DVDnet and FastDVDnet give a high quality of frames but from the results of PSNR and the image details, FastDVDnetoutperforms DVDnet.

5. CONCLUDING REMARKS

This work aimed to present image and video denoising methods already existing in the literature especially new methods based on deep learning techniques that have attracted attention due to their exceptional performance, namely: DVDnet and Fast DVDnet.

We have also studied and implemented VBM3D and VBM4D denoising algorithms, which are based on the search for similar patches in a Spatio-temporal space. These two algorithms are implemented with COVID-19 dataset to enrich our comparative study but also to show the interest of CNN in denoising.

With COVID-19 dataset, the lower the σ, the greater the PSNR with all the algorithms. DVDnet outperforms VBM3D and VBM4D algorithms in terms of running time, PSNR, and preservation of image details.

Additive Poisson noise is also considered. We implemented Poisson noise that we generated using MATLAB on COVID-19 dataset; the results were satisfactory although the program is specialized for Gaussian noise.

Concerning the second TESTPCR dataset, we applied FastDVDnet algorithm, which is more improved than DVDnet.Both quantitative and qualitative denoising results of FastDVDnet show remarkable temporal consistency, very low jitter, and a satisfactory details preservation with very fast running time,

Finally, every architecture based on mimicking the human brain gives satisfactory results so we can make improvements on all the algorithms based on CNN by increasing the number of convolutional layers and looking for other activation and acceleration functions.

References

[1] M. Maggioni, G. Boracchi, A. Foi, and K. Egiazarian, “Video denoising, deblocking, and enhancement through separable 4-D nonlocal spatiotemporal transforms,” IEEE Transactions on Image Processing, vol. 21, no. 9, pp. 3952–3966, Sep. 2012.

[2] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “BM3D image denoising with shape adaptive principal component analysis,” in Proc. Signal Process. Adapt. Sparse Struct. Represent., Saint-Malo, France, Apr. 2009.

[3] T. Ethert, P. Arias "Implementation of the VBM3D Video Denoising Method and Some Variant "ArXiv Journal, vol.2001.01802, Janv. 2020.

[4] M. Tassano, J. Delon, and T. Veit “ An Analysis and Implementation of the FFDnet image denoising method. Image Processing On Line (IPOL), 9, pp.1–25, Jan 2019.

[5] K. Zhang, W. Zuo, and L. Zhang. FFDnet: Toward a Fast and Flexible Solution for CNN-Based Image Denoising. IEEE Transactions on Image Processing, 27(9), pp.4608–4622, Sep. 2018. 1, 2, 3, 5.

[6] M. Tassano, J. Delon, and T. Veit"FastDVDnet : Towards Real-Time Deep Video Denoising Without Flow Estimation", ArXiv preprint ArXiv : 1907.01361 2020.

[7] W. Shi, J. Caballero, F. Huszar, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang “Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1874–1883, Las Vegas, NV, USA, Jun 2016.

[8] A. Krizhevsky, I. Sutskever, and G.E. Hinton, “ ImageNet Classification with Deep Convolutional Neural Networks, Advances In Neural Information Processing Systems, pp. 1–9, 2012.

[9] M. Tassano, J. Delon, and T. Vei "DVDnet : A Fast Network for Deep Video Denoising",IEEE International Conference on Image Processing (ICIP),pp. 1-25, 22-25 Sept. 2019, Taipei, Taiwan, Taiwan.

[10] A.Davy Th.Ehret, G.Facciolo,J.Morel Pa.Arias, Non-local video denoising by CNN , , arXiv: 1811.12758v1 [cs. CV] 30 November 2018.