2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Conditional Generative Adversarial Networks for low-dose CT image denoising aiming at preservation of critical image content

Koen C. Kusters[1], Luis A. Zavala-Mondragón[1], Javier Oliván Bescós[2], Peter Rongen[2], Peter H.N. de With[1], Fons van der Sommen[1]

Abstract— **X-ray Computed Tomography (CT) is an imaging modality where patients are exposed to potentially harmful ionizing radiation. To limit patient risk, reduced-dose protocols are desirable, which inherently lead to an increased noise level in the reconstructed CT scans. Consequently, noise reduction algorithms are indispensable in the reconstruction processing chain. In this paper, we propose to leverage a conditional Generative Adversarial Networks (cGAN) model, to translate CT images from low-to-routine dose. However, when aiming to produce realistic images, such generative models may alter critical image content. Therefore, we propose to employ a frequency-based separation of the input prior to applying the cGAN model, in order to limit the cGAN to high-frequency bands, while leaving low-frequency bands untouched. The results of the proposed method are compared to a state-of-the-art model within the cGAN model as well as in a single-network setting. The proposed method generates visually superior results compared to the single-network model and the cGAN model in terms of quality of texture and preservation of fine structural details. It also appeared that the PSNR, SSIM and TV metrics are less important than a careful visual evaluation of the results. The obtained results demonstrate the relevance of defining and separating the input image into desired and undesired content, rather than blindly denoising entire images. This study shows promising results for further investigation of generative models towards finding a reliable deep learning-based noise reduction algorithm for low-dose CT acquisition.**

I. Introduction

X-ray Computed Tomography (CT) is an imaging modality routinely used in clinical practice in which patients are exposed to potentially harmful X-ray radiation [1]. With the aim of decreasing the radiation exposure, reduced-dose protocols are desirable. However, the reduced-dose images are corrupted by noise that is hard to model due to its diverse sources, as well as the CT image reconstruction process. The presence of noise may impair the visualization of clinically relevant features by clinicians. To make the application of reduced-dose protocols feasible, the inclusion of noise reduction algorithms within the processing chains is required.

There are diverse approaches for dealing with noise in CT. For example, the inclusion of noise reduction stages after reconstruction, such as the total-variation regularization approach proposed by Tian *et al.* [2]. Popular alternatives are to perform denoising through iterative reconstruction techniques [3], [4], [5], [6]. However, the sequential nature of these these approaches increases their execution time, which makes them less attractive for clinical applications.

More recently, deep learning techniques such as Generative Adversarial Networks (GAN) [7] have shown promising results in noise reduction for CT. Yang *et al.* [8] and Wolterink *et al.* [9] exploited the GAN framework in context of low-to-routine dose CT emulation. In the sequel of this paper we will denote low-to-routine dose as 'low'-dose. The results have shown that GAN frameworks excel in image-translation tasks and produce images that are both well-denoised and realistically looking. However, despite the visually appealing results achieved by such generative models, a point of concern is that they may alter critical image content [8], which can impede correct diagnosis by clinicians.

Several well-established GAN frameworks exist that are suitable for the image translation tasks, such as Cycle-Consistent Adversarial Networks (CycleGAN) [10], Wasserstein Generative Adversarial Networks (WGAN) [11] or conditional Generative Adversarial Networks (cGAN) [12]. Each of these frameworks has specific strengths. Cycle-GAN enables a consistent cycle of translation between domains, which makes it feasible for unpaired data applications. Furthermore, WGANs use a specialized loss term that improves training stability. Finally, the cGAN framework, proposed by Isola *et al.* [12], addresses the case of paired image-to-image translation and achieves excellent results. Given the nature of our dataset, namely paired low- and routine-dose CT images, and the success of the cGAN architecture, we have chosen to explore the cGAN framework for 'low'-dose image translation. cGAN frameworks are composed by two Convolutional Neural Networks (CNNs), the generator and discriminator respectively. The generator is in charge of translating an image from input domain to the target domain, while the discriminator has the task to distinguish whether the generated image is a 'real' or 'fake' image in the target domain. Adversarial feedback of the discriminator forces the generator to produce images that resemble the 'real' examples better.

The main contribution of this paper is the exploration of a generator architecture averting alteration of critical image content, based on the popular Residual Encoder-Decoder (RED) CNN proposed by Chen *et al.* [13], which is designed for noise reduction in CT. To improve the preservation of image content, we propose to decompose the signal with an undecimated Laplacian Pyramid with one decomposition level to separate high-frequency noise from low-frequency

[1] Koen C. Kusters, Luis A. Zavala-Mondragón, Peter H.N. de With and Fons van der Sommen are with the Faculty of Electrical Engineering, Eindhoven Univ. of Technology, The Netherlands - fvdsommen@tue.nl

[2] Javier Oliván Bescós and Peter Rongen are with Philips Healthcare, Best, The Netherlands.

image content in CT data. Afterwards, we apply the RED-CNN to remove noise exclusively on the high-frequency band components, as we hypothesize that a large part of the noise is contained in this frequency band. As the reader may notice, we do not process the low-frequency band, which typically contains most of the signal energy. This approach significantly limits the modification of critical image content contained in the low-frequency band. Hereafter, this approach is referred to as Laplacian RED-CNN.

Two main experiments are performed to prove that Laplacian RED-CNN manages to preserve details better than the standard RED in noise reduction applications within a cGAN framework. First, for reference, we train RED-CNN to perform noise reduction without cGAN regularization. In the second experiment, we train RED and Laplacian RED-CNN with adversarial feedback provided by a PatchGAN CNN [12]. The results show that our approach better preserves the content of the image, which is important in the medical context. To prove our claims, we present and discuss numerical metric values evaluated on full-CT slices, as well as on patches of interest. Furthermore, a discussion on visual results of generated images and suggestions for future research are provided.

The remainder of the paper is organized as follows. Section II gives an overview of the methodology, Section III elaborates on the experimental results with an extensive discussion. The conclusions are presented in Section IV.

II. Methods

A. CT Data

For our experiments, we use the brain scans provided with the "Low Dose CT Image and Projection Data" [14], [15] from The Cancer Imaging Archive (TCIA) [16]. This dataset consists of 50 head scans acquired from 50 different patients, providing a total of 1,782 routine-dose (NDCT) and low-dose (LDCT) slices. In this dataset, LDCT scans were derived from the NDCT ones, by adding synthetic noise to simulate acquisitions with 25% of routine-dose. Each slice is an image with a size of $1 \times 512 \times 512$ pixels. The available data was randomly divided into training, validation and test sets on a per-patient basis. A split of 80%–20% was taken for training and test sets, respectively, while the validation set was taken as 10% of the training set. This resulted in training, test and validation sets consisting of 1,291, 351 and 140 LDCT-NDCT image pairs, respectively.

B. Data pre-processing

In CT imaging, pixel intensities of the images are expressed in Hounsfield Units (HU), which cover a wide range of tissues. Since we are interested in improving the visualization of brain tissue, we constrain the HU window to the range [-9, +136] HU, which is close to the values observed by clinicians. After this operation, we normalized the pixel values to the range [0,1]. Furthermore, since the training set was relatively small, data augmentation was utilized to increase the size of the set. The augmentation operations employed for this purpose were rotation of $90°$, $180°$ and $270°$, horizontal flipping with $0°$ and $90°$ rotation and vertical flipping with $0°$ and $90°$ rotation. These augmentation transformations increased the training set size with a factor of 8. Thus, after data augmentation the training set consisted of 10,328 LDCT-NDCT image pairs.

C. Network Architectures

1) Generator: RED-CNN proposed by Chen *et al.* [13] serves as the basis for the generator architecture in the cGAN framework[1]. This 10-layer network employs an encoder-decoder structure with 3 residual skip connections. The architecture comprises of 5 blocks with convolutional layers for the encoder and 5 blocks with transposed convolutional layers for the decoder. All (transposed) convolutional layers employ 96 filters of size 5×5 with stride 1 over non-zero padded input, except for the last transposed convolutional layer with a single filter. The residual connections from nodes in the encoder are added to some feature maps in the decoder. The input and output size of the RED-CNN generator is $BS \times 1 \times 512 \times 512$, where BS represents the batch size.

In our experience, the use of CNNs such as RED, in a conditional GAN framework may lead to alteration of the image content with the objective to produce clear and realistically looking images. We hypothesize that the reason for these alterations is given by the fact that conventional encoding-decoding CNNs, attempt to learn simultaneously filtering and reconstruction of the signal, which may lead to alterations of the signal. To overcome this limitation, we propose to separate the spectrum of the signal in a low- and high-frequency band by applying an undecimated Laplacian Pyramid with one decomposition level. This separation forces RED to focus the processing exclusively on the high-frequency band, where generally the largest fraction of the energy is contained with noise, while leaving the low-frequency band unchanged. We call this approach the Laplacian RED-CNN.

The implementation of Laplacian RED-CNN is as follows. The first layer is a 9×9 Gaussian low-pass filter with $\sigma = 3$, which extracts low-frequency information from the input image. The next step of the Laplacian pyramid is to subtract the low-frequency image version from the original input. This yields high-frequency information, consisting mostly of noise, which will serve as input to the RED-CNN.

Due to the subtraction of low-frequency content, feature maps inside the architecture will be zero-mean. To avoid loss of information embedded in negative activations, the final activation layer is replaced by an hyperbolic tangent activation. The final operation in the modified generator is to add the low-frequency content back again to the processed high-frequency content. The structure of both RED-CNN as well as the Laplacian RED-CNN generator are presented in Fig. 1.

2) Discriminator: A multi-scale PatchGAN discriminator architecture, based on the paper of Isola *et al.* [12], was adopted for this research. This discriminator architecture

[1]Based on https://github.com/SSinyu/RED-CNN/blob/master/networks.py

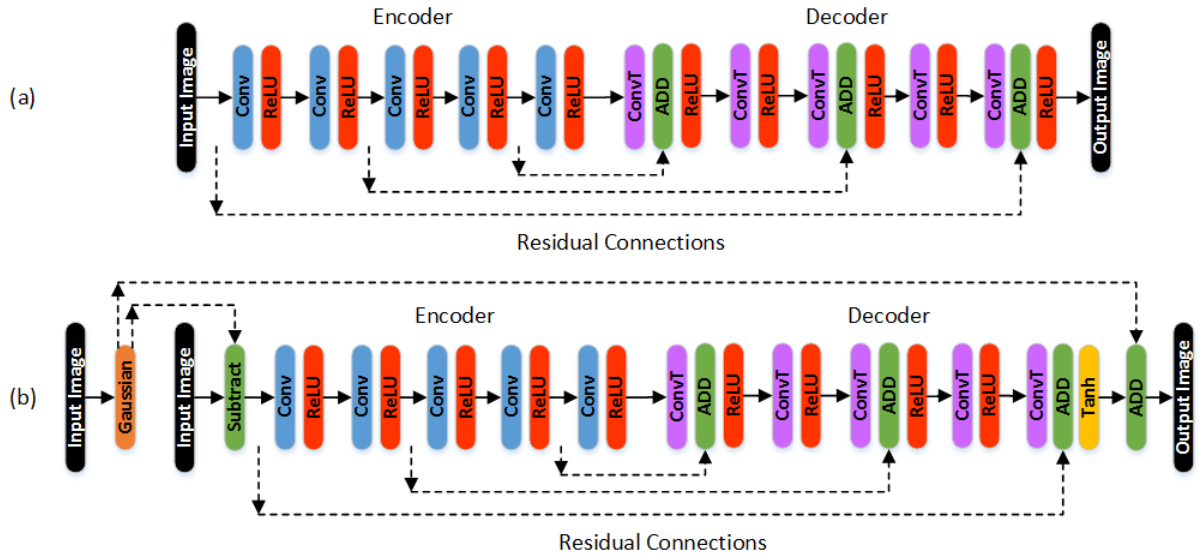

Fig. 1. Structure of the generator architectures. (a) RED-CNN generator. (b) Laplacian RED-CNN generator: RED-CNN generator adapted for image content preservation by an undecimated Laplacian Pyramid with one decomposition level and hyperbolic tangent activation.

restricts attention to structures in local image patches. The output size of this discriminator is $BS \times 1 \times 32 \times 32$, where BS stands for batch size. This implies that each 16×16 patch is classified as either 'real' or 'fake'. The architecture of the discriminator is defined by:

$$C64 - C_N128 - C_N256 - C_N512 - C_S1, \tag{1}$$

where Ck indicates a convolutional layer followed by a leaky Rectified Linear Unit (ReLU), C_Nk denotes a convolution block followed by a batch normalization and a leaky ReLU. Lastly, C_Sk indicates a convolutional layer followed by a Sigmoid activation. For the aforementioned building blocks, the number k indicates the number of filters. Convolutional layers employ kernels of size 4×4 applied with stride 2 and padding 1, except for C_S1 with a stride of 1. For all leaky ReLUs, parameter α is set to $\alpha = 0.2$.

D. Training Implementation

In cGAN frameworks, the generator and discriminator are trained in an adversarial manner. The discriminator indicates if the processed LDCT image looks 'real' or 'fake', thereby encouraging the generator to improve the texture of the processed images to better resemble the corresponding 'real' NDCT image. In each training iteration, the discriminator is trained on this task by providing a batch of NDCT-LDCT pairs and a batch of Processed-LDCT pairs. To this end, the following loss function is specified:

$$\begin{aligned}\mathcal{L}_D = &\frac{1}{2} L_{\text{BCE}}[D(I^{LD}, I^{ND}), 0] \\ &+ \frac{1}{2} L_{\text{BCE}}[D(I^{LD}, G(I^{LD})), 1],\end{aligned} \tag{2}$$

where D represents discriminator, G denotes generator, I^{LD} and I^{ND} are the LDCT and NDCT image, respectively. The first loss term reflects the binary cross-entropy (BCE) between the discriminator's output on 'real' target NDCT data and target label 0. The second loss term reflects the BCE between the discriminator's output on generated data and label 1. The generator is responsible for translating noisy LDCT images to the NDCT target domain, such that the discriminator is not able to distinguish processed images from real NDCT images. This process is guided by a loss function that is specified by:

$$\begin{aligned}\mathcal{L}_G = &L_{\text{BCE}}[D(I^{LD}, G(I^{LD})), 0] \\ &+ \lambda \, ||G(I^{LD}) - I^{ND}||_1,\end{aligned} \tag{3}$$

where $\lambda = 100$ is a weighting factor for the L_1 loss. The first loss term reflects the BCE between the discriminator's output on generated data and inverted target label, which represents the adversarial feedback of the discriminator to the generator. The second loss term indicates the L_1 loss between the generated and NDCT target image. This additional loss forces the generator to create images that are closely resembling the NDCT target image in terms of pixel intensity distribution.

The proposed cGAN models were trained for 16 epochs with a batch size of 4 image pairs. Parameters in both the generator and discriminator were updated using the Adam optimizer, with a learning rate of 0.0002, where $\beta_1 = 0.5$ and $\beta_2 = 0.999$. Furthermore, a learning rate scheduler was implemented, which reduced the maximum learning rate by a factor 0.2 with threshold 0.01 and a patience of 5 epochs, when the generator loss did not decrease. The proposed method has been implemented in Python using the PyTorch framework and experiments were executed on an GeForce GTX 1080 Ti.

E. Evaluation

For performance evaluation of the generated images as well as image patches of interest, three different metrics are employed. As a first metric, Peak Signal-to-Noise ratio

(PSNR) is utilized. PSNR is expressed in decibels [dB] and is defined as:

$$\text{PSNR} = 10 \log_{10} \left(\frac{R^2}{\text{MSE}} \right), \tag{4}$$

where R is the maximum pixel intensity and MSE is the mean-squared error between generated image and reference NDCT target image.

As a complementary metric to the PSNR, we compute the Structural Similarity Index (SSIM), which indicates the perceptual difference between two images in terms of structural changes. The SSIM is computed window-wise between generated and NDCT target image and yields a value in the unity interval, where unity means perfect similarity. The value of SSIM between windows x and y is computed by:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \tag{5}$$

where μ denotes the mean of the corresponding window, σ is the variance of the corresponding window and σ_{xy} is the covariance between the windows. Furthermore, $c_1 = (0.01 \cdot L)^2$ and $c_2 = (0.03 \cdot L)^2$ with L being the dynamic range of pixel values.

As third and last metric, Total Variation (TV) is used. This metric gives information about the smoothness of images, which is useful since noisy images typically have a higher TV than images without noise. The TV for each pixel in an image is calculated by:

$$\text{TV}(x, y) = \sqrt{\sum_{i=-1}^{+1} (I_{x,y} - I_{x+i,y})^2 + (I_{x,y} - I_{x,y+i})^2}, \tag{6}$$

where $I_{x,y}$ represents the pixel intensity of pixel (x, y). In order to obtain a single value for each image, all values within the resulting 2D matrix are accumulated.

TABLE I
REFERENCE VALUES OF LDCT AND NDCT FOR PATCHES OF INTEREST AND VALUES OF FULL-CT IMAGES AVERAGED OVER THE TEST SET

	LDCT			**NDCT**
	PSNR [dB]	SSIM [-]	TV [-]	TV [-]
Patch A	17.0585	0.5252	1,205.10	642.61
Patch B	19.7134	0.7287	293.41	183.10
Patch C	15.4060	0.6322	394.71	224.64
Patch D	17.0426	0.5727	615.25	330.22
Test set	32.6730	0.9354	11,443.87	8,894.50

III. Experimental results and Discussion

To demonstrate the improved performance of our approach, we evaluate cGAN frameworks employing RED-CNN and Laplacian RED-CNN as generator, as well as a RED-CNN without adversarial feedback of an auxiliary CNN. The performances of the methods are measured with the defined metrics and validated on full-CT images in the test set as well as on patches of interest. In order to assess the performance of the methods more precisely, the metrics are also evaluated on the corresponding LDCT and NDCT images, to serve as a reference.

These reference values are depicted in Table I, while the results for the described methods are listed in Table II. The results from Tables I and II are illustrated in Fig. 2 by showing the denoised images with highlighted patches of interest, as well as the corresponding source LDCT and target NDCT images. As shown in Table II, RED-CNN without auxiliary CNN yields results that are superior to the results of the cGAN models, in terms of PSNR and SSIM values. Furthermore, it is shown that the cGAN framework with RED-CNN as generator produces results that are the best in terms of TV, since the values are closest to the reference TV values of corresponding target NDCT images depicted in Table I.

During evaluation of the results, we have observed that the metric values may not always reflect the perceptual observations, so that visual inspection of the generated images remains the most important criterion for quality assessment. From the illustrations in Fig. 2, it can be observed that RED-CNN without adversarial feedback generates images that are severely smoothed. Hence, generated images do not resemble target images in terms of texture, while preserving most of the fine structural details, as highlighted in Patches A and B. However, due to aggressive smoothing, low-contrast structures in images are attenuated, which can be seen in Patches C and D. We hypothesize that the smoothing behavior is related to the fact that in the setup without auxiliary CNN, the network needs to perform both denoising and reconstruction simultaneously. The network focuses only on the denoising part, since it does not succeed in performing both.

A way to improve on texture is the usage of an auxiliary discriminator architecture in a cGAN framework, which is illustrated in the third and fourth columns in Fig. 2. The cGAN framework with RED-CNN as generator produces images that resemble the target images significantly better in terms of generated texture. This is a direct consequence of the adversarial feedback of the discriminator architecture, which encourages the generator to create such texture. However, Patches A, C and D and to a lesser extent B, show a major limitation of the usage of a conventional CNN as generator. As can be observed, fine structures highlighted in the patches are severely deformed or even removed. The cause of this phenomenon is that the generator prioritizes to mimic the texture of the NDCT target image, rather than preserving critical image content.

In order to avert the destruction of essential image content, the content that is desired in the output image should be preserved by excluding it from the input to the generator. The results of the cGAN framework with Laplacian RED-CNN as generator show that this generator is able to preserve these fine structural details, while also resembling the target texture to a reasonable extent. The highlighted patches show that this method outperforms the others in terms of preservation of fine structures highlighted in these patches. However, the generated texture is slightly less realistic compared to the

TABLE II

RESULTS OF RED-CNN, CGAN (RED-CNN) AND CGAN (LAPLACIAN RED-CNN) ON PATCHES OF INTEREST AND VALUES OF FULL-CT IMAGES AVERAGED OVER THE TEST SET

	RED-CNN			**cGAN (RED-CNN)**			**cGAN (Laplacian RED-CNN)**		
	PSNR [dB]	SSIM [-]	TV [-]	PSNR [dB]	SSIM [-]	TV [-]	PSNR [dB]	SSIM [-]	TV [-]
Patch A	**22.0562**	**0.5874**	365.70	19.0191	0.4175	**616.28**	19.6535	0.5233	679.55
Patch B	**24.2092**	**0.8186**	139.08	21.3418	0.7437	197.36	21.7002	0.7413	**195.49**
Patch C	**18.8080**	**0.6556**	126.29	16.0615	0.5173	**227.87**	16.7848	0.5859	228.34
Patch D	**21.9552**	**0.6611**	207.65	18.6532	0.4950	**346.38**	19.8486	0.6042	363.69
Test set	**35.8688**	**0.9494**	7,648.08	32.9276	0.9256	**8,924.82**	33.1356	0.9346	9,210.62

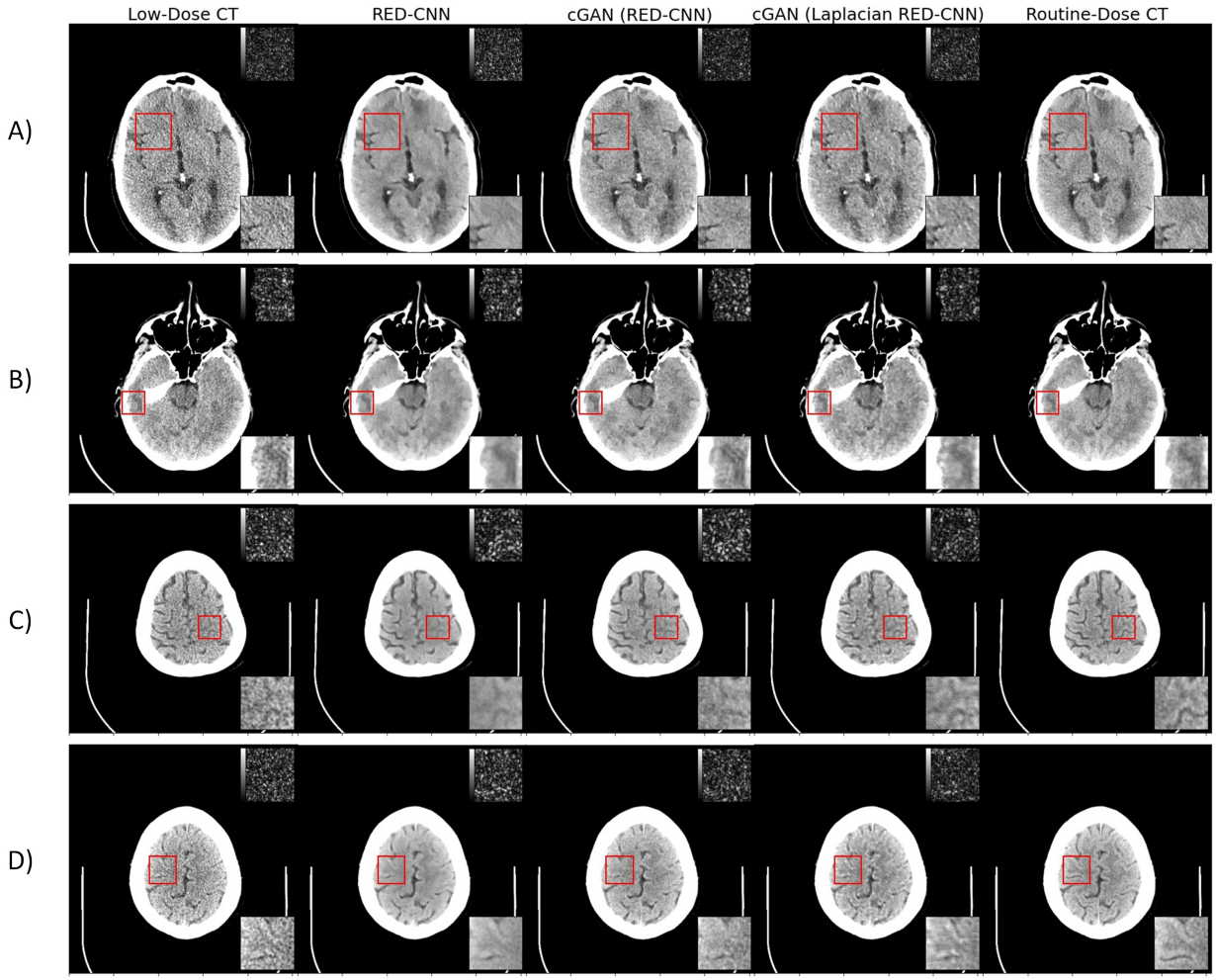

Fig. 2. Results of three different methods for 4 image examples, referred to as Sub-figures A), B), C) and D). For each image in these sub-figures, the contents of the red box is presented as magnified view at the right bottom. These zoomed versions are referred to as Patches A, B, C and D. Similarly, at the right top, the absolute difference is shown with the target patch. The display window is ranged within [-9, +56] HU.

results of the cGAN with RED-CNN as generator. Thus, it can be concluded that the separation of the signal, to restrict the attention of the generator architecture on only high-frequency content, can be successfully used to avoid alteration of image content. On the other hand, due to this restriction, the generator experiences difficulties in generating realistic textures.

When combining the results in Table II and Fig. 2, it can be concluded that PSNR and SSIM favor smooth images, since RED-CNN without auxiliary CNN obtains superior results. TV accurately reflects the smoothness and inherently the *realisticness* of images. Comparing the results of both cGAN experiments at the scale of full-CT images as well as local patches, SSIM correlates with the preservation of structural details.

In conclusion, the results of the proposed method are still sub-optimal, in the sense that generated texture should still improve, while preserving the low level of alterations of critical image content. A potential improvement could be realized when optimizing the choice of the Gaussian kernel

parameters, since the algorithm is very sensitive to the size of the kernel and the value of σ. The size of the kernel determines the size of structures that will be captured by this filtering operation, while the standard deviation (σ) is highly important for the amount of residual noise on the low-frequency content that is preserved. Furthermore, the generator has to reconstruct texture based only on the high-frequency information that it is provided with. Hence, the choice of these parameters is a very delicate trade-off.

Therefore, future research should be focused on possible optimization methods for these parameters, e.g. exploration of trainable Laplacian Pyramid parameters. Furthermore, to exploit the benefits of a Laplacian Pyramid even more, it could be applied at multiple scales instead of only one. Ensuring that high-frequency information will be extracted at multiple scales, can be beneficial for both the high-frequency noise removal process as well as the texture reconstruction.

IV. Conclusions

Denoising techniques are required in the reconstruction processing chain of CT images acquired with a reduced-dose protocol. In this paper, we have presented the Laplacian RED-CNN, as an alternative to the traditional RED-CNN generator in cGANs for noise reduction. The proposed approach separates the input image into low- and high-frequency bands, which potentially avoids the loss of critical image content by concentrating the learning of the cGAN on the high-frequency band. In this way, most high-frequency undesired content will be removed by the generative model, while preserving low-frequency structural image content.

The obtained results show the feasibility of this approach for the low-to-routine CT image-translation task and helps in avoiding the common pitfall of generative models. However, the results on texture generation can still be further improved, by e.g. investing in optimization of parameters for the separation process with the Gaussian kernel. Furthermore, real clinical quarter-dose datasets are envisioned as part of future work in order to provide additional evidence supporting our claims.

In summary, this study presents a promising point of engagement for further investigation of generative models towards finding a reliable deep learning-based noise reduction for low-dose CT acquisition.

V. Acknowledgment

The authors would like to thank the National Institute of Biomedical Imaging and Bioengineering for funding provided for the generation of the dataset used in this work through the grants EB017095 and EB017185 (Cynthia McCollough, PI). In addition, the authors thank the European Union for the funding received for this research through the Horizon 2020 Grant No. 780026 "NEXIS".

References

[1] D. Brenner and E. Hall, "Computed tomography–an increasing source of radiation exposure." *The New England journal of medicine*, vol. 357 22, pp. 2277–84, 2007.

[2] Z. Tian, X. Jia, K. Yuan, T. Pan, and S. Jiang, "Low-dose ct reconstruction via edge-preserving total variation regularization." *Physics in medicine and biology*, vol. 56 18, pp. 5949–67, 2011.

[3] M. Beister, D. Kolditz, and W. Kalender, "Iterative reconstruction methods in x-ray ct," *Physica medica : PM : an international journal devoted to the applications of physics to medicine and biology : official journal of the Italian Association of Biomedical Physics (AIFB)*, vol. 28, pp. 94–108, 02 2012.

[4] A. Hara, R. Paden, A. Silva, J. Kujak, H. Lawder, and W. Pavlicek, "Iterative reconstruction technique for reducing body radiation dose at ct: Feasibility study," *AJR. American journal of roentgenology*, vol. 193, pp. 764–71, 09 2009.

[5] M. Willemink, T. Leiner, P. Jong, L. de Heer, R. Nievelstein, A. Schilham, and R. Budde, "Iterative reconstruction techniques for computed tomography part 2: Initial results in dose reduction and image quality," *European radiology*, vol. 23, 01 2013.

[6] D. Wu, K. Kim, G. El Fakhri, and Q. Li, "Iterative low-dose ct reconstruction with priors trained by artificial neural network," *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2479–2486, 2017.

[7] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," 2014.

[8] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, and G. Wang, "Low dose CT image denoising using a generative adversarial network with wasserstein distance and perceptual loss," *CoRR*, vol. abs/1708.00961, 2017. [Online]. Available: http://arxiv.org/abs/1708.00961

[9] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, "Generative adversarial networks for noise reduction in low-dose ct," *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2536–2545, 2017.

[10] J. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," *CoRR*, vol. abs/1703.10593, 2017. [Online]. Available: http://arxiv.org/abs/1703.10593

[11] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein gan," 2017.

[12] P. Isola, J. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *CoRR*, vol. abs/1611.07004, 2016. [Online]. Available: http://arxiv.org/abs/1611.07004

[13] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, and G. Wang, "Low-dose ct with a residual encoder-decoder convolutional neural network," *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2524–2535, 2017.

[14] C. McCollough, B. Chen, D. Holmes, X. Duan, Z. Yu, L. Yu, S. Leng, and J. Fletcher, "Data from low dose ct image and projection data [data set]," *The Cancer Imaging Archive*, 2020. [Online]. Available: https://doi.org/10.7937/9npb-2637

[15] T. R. Moen, B. Chen, D. R. Holmes III, X. Duan, Z. Yu, L. Yu, S. Leng, J. G. Fletcher, and C. H. McCollough, "Low-dose ct image and projection dataset," *Medical Physics*, vol. 48, no. 2, pp. 902–911, 2021. [Online]. Available: https://aapm.onlinelibrary.wiley.com/doi/abs/10.1002/mp.14594

[16] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Philips, D. Maffitt, M. Pringle, L. Tarbox, and F. Prior, "The cancer imaging archive (tcia): Maintaining and operating a public information repository," *Journal of Digital Imaging*, vol. 26, no. 6, pp. 1045–1057, December 2013.