The Gated Recurrent Conditional Generative Adversarial Network (GRC-GAN): application to denoising of low-dose CT images

Mateus Baltazar de Almeida
Universidade Federal do Agreste de Pernambuco
Email: mateusbaltazar9@gmail.com

Luis F. Alves Pereira
Universidade Federal do Agreste de Pernambuco
Email: luis-filipe.pereira@ufape.edu.br

Tsang Ing Ren
Universidade Federal de Pernambuco
Email: tir@cin.ufpe.br

George D. C. Cavalcanti
Universidade Federal de Pernambuco
Email: gdcc@cin.ufpe.br

Jan Sijbers
University of Antwerp
Email: jan.sijbers@uantwerpen.be

***Abstract*—The ionizing radiation that propagates through the human body at Computed Tomography (CT) exams is known to be carcinogenic. For this reason, the development of methods for image reconstruction that operate with reduced radiation doses is essential. If we reduce the electrical current in the electrically powered X-ray tubes of CT scanners, the amount of radiation that passes through the human body during a CT exam is reduced. However, significant image noise emerges in the reconstructed CT slices if standard reconstruction methods are applied. To estimate routine-dose CT images from low-dose CT images and thus reduce noise, the *Conditional Generative Adversarial Network* (cGAN) was recently proposed in the literature. In this work, we introduce the *Gated Recurrent Conditional Generative Adversarial Network* (GRC-GAN) that is based on the usage of network gates to learn the specific regions of the input image to be updated using the cGAN denoising operation. Moreover, the GRC-GAN is executed recurrently in multiple time steps. At each time step, different parts of the input image are denoised. As a result, our GRC-GAN better focus on the denoise criterium than the regular cGAN in the LoDoPaB-CT benchmark.**

I. Introduction

X-ray Computed Tomography (CT) is a key medical imaging technology. Chest CT, for instance, is currently playing a decisive role in the detection of COVID-19 [1]. However, the ionizing radiation to which CT examinations expose the patients are associated with severe health risks in the long term [2]. This scenario is even more dramatic for patients with health conditions that require recurrent CT scans [3], [4].

Low-Dose CT (LDCT) scans can be produced by reducing the electrical current of the X-ray tube. However, severe noise arises in the images generated from LDCT scans if conventional reconstruction methods are used. To eliminate the noise from the LDCT reconstructions and to preserve the structural information of the tissues, several Deep Learning methods have been proposed in the literature [5], [6]. Such approaches guide the future of medical imaging since the US Food & Drugs Administration (FDA) recently approved six algorithms based on Deep Learning for improving different imaging modalities [7]: the Deep Learning Image Reconstruction (GE Medical Systems, LLC.), the Advanced Intelligent Clear-IQ Engine (Canon Medical Systems Corporation), the SubtleMR (Subtle Medical, Inc.), the AI-Rad Companion - Pulmonary (Siemens Medical Solutions USA, Inc.), the AI-Rad Companion - Cardiovascular (Siemens Medical Solutions USA, Inc.), and the SubtlePET (Subtle Medical, Inc.).

The *Conditional Generative Adversarial Networks* (cGANs) are one of the most suited architectures for recovery of LDCT images [8]. In this work, we propose the *Gated Recurrent Conditional Generative Adversarial Network* (GRC-GAN), which incorporates concepts of *Convolutional Gated Recurrent Unit* (ConvGRU) [9] to improve the prediction performance of regular cGAN. While previously proposed recurrent GAN architectures are designed to learn spatio-temporal dependencies within a 3D volume [10]–[12], we apply recurrent calls of a cGAN followed by convolutional gates that learn which parts of the current solution should be suppressed or preserved in the next GRC-GAN iteration.

The main contributions of this paper are: *(i)* we show how to use recurrent networks driven by gates for updating partial solutions within an image denoising problem; *(ii)* we show an indication that new gated recurrent architectures have the potential to outperform many encoder-decoder networks previously proposed in the literature for image denoise since this new model can change its focus at each iteration; *(iii)* we present a method that improves the noise reduction obtained by the cGAN [8] at the low-dose CT problem by more than 35% in terms of the Normalized Root Mean Squared Error (NRMSE) measure; and *(iv)* our GRC-GAN, in contrast to the cGAN, follows the recent theory of eXplainable Artificial Intelligence (XAI) [13] since - by monitoring the 2D signals at the gates - it can be observed that different parts of the input image are the focus of the denoise operation at each time step.

II. Background

A. Conditional Generative Adversarial Networks (cGANs)

The GAN architecture [14] is based on the interaction of two neural networks: a Generator (G) and a Discriminator (D). While G is trained to generate new data distributions, D is trained to predict whereas a given input data is *(i)* a sample from the real-world or *(ii)* a simulation from G. Both G and D are then trained to compete to each other. At the end of this process, it is expected that G generates outputs almost indistinguishable from real samples. Mathematically, the GAN training stage is defined by the following *minimax* problem in terms of the value function $V(D,G)$:

$$\min_G \max_D V(D,G) = \mathbb{E}\left[\log\, D(y)\right] + \mathbb{E}\left[\log\left(1 - D(G(z))\right)\right], \tag{1}$$

where $\mathbb{E}(Z)$ is the expected value for the random variable Z, z is an input noise, and y is a real sample, $D(k) \to 1$ if the signal k comes from a *real* data distribution, and $D(k) \to 0$ if k is a simulated signal.

The Conditional GAN (cGAN) [15], is a GAN extension where both generator (G) and discriminator (D) are conditioned on prior information x. In this case, the cGAN training stage is extended as:

$$\min_G \max_D V(D,G) = \mathbb{E}\left[\log\, D(y|x)\right] + \mathbb{E}\left[\log\left(1 - D(G(z|x))\right)\right], \tag{2}$$

Fig. 1 illustrates the cGAN architecture proposed in [8] for noise reduction of LDCT images. In such approach, x is the LDCT image associated to the normal dose CT image y, and $z|x \to x$.

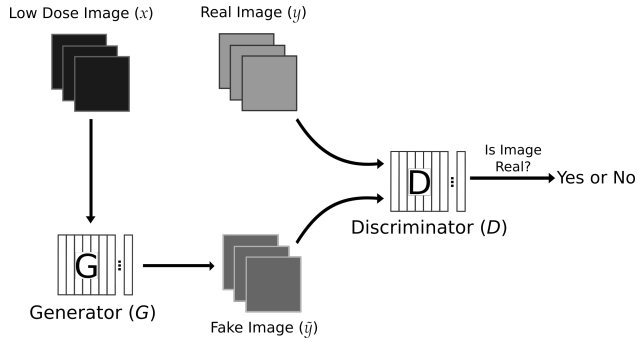

Fig. 1. The cGAN architecture for noise reduction of LDCT images [8]: G is trained to estimate a normal dose reconstruction $\tilde{y}$ from the LDCT image x given as input, and D is trained to distinguish whereas a given input is a real sample y or an estimation $\tilde{y}$.

B. Convolutional Gated Recurrent Unit (ConvGRU)

Initially proposed for processing sequential data, the GRU [16] is a gated recurrent architecture where *(i)* a hidden state h_t is propagated through the network execution along t time steps to hold information about the $t-1$ previous executions, and *(ii)* gates control whether the previous hidden state should be updated or ignored during the current computation.

From the original GRU, the convolutional GRU (ConvGRU) [9] was developed to handle 2D images. In this work, we use a simplified ConvGRU (illustrated in Fig. 2) where h_t is the GRU output x_t itself. Furthermore, $R : \mathbb{R}^{m\times n\times s} \to \mathbb{R}^{m\times n}$, $U : \mathbb{R}^{m\times n\times s} \to \mathbb{R}^{m\times n}$, and $O : \mathbb{R}^{m\times n\times s} \to \mathbb{R}^{m\times n\times s}$ are convolutional neural networks that represent the *reset/forget*, *update*, and *output* gates, respectively.

In such architecture, the *reset/forget* gate defines how much of the information carried from all previous time steps should be forgotten. On the other hand, the *update* gate controls how much of the past information should be passed to future stages. Finally, the *output* gate introduces new information to the solution based on the parts of the input data defined as relevant by the *reset/forget* gate.

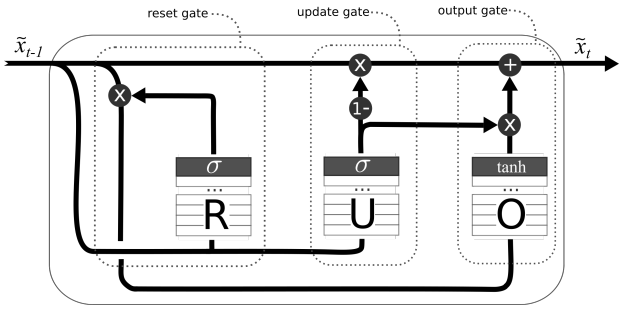

Fig. 2. The ConvGRU architecture used in this work: the hidden state h_t is equal to the output $\tilde{x}_t$. Furthermore, R, U, and O are the *reset*, *update*, and *output* gates, respectively.

Mathematically, the output $\tilde{x}_t$ in this ConvGRU architecture is computed as follows:

$$\tilde{x}_t = O(R(\tilde{x}_{t-1}) \cdot \tilde{x}_{t-1}) \cdot U(\tilde{x}_{t-1}) + (1 - U(\tilde{x}_{t-1})) \cdot \tilde{x}_{t-1} \tag{3}$$

III. Related Works

The works previously published in the state-of-the-art that are closely related to our research are divided into two main categories: *(i)* methods for recovering LDCT image reconstructions using GANs; and *(ii)* new recurrent architectures based on GANs. Starting from the regular cGAN architecture for LDCT [8], different solutions were proposed by changing the loss function [17], using a more complex *Generator* network G [18], and even including a third network into the architecture to estimate image sharpness [19]. Nevertheless, there is no recurrent GAN previously published in the literature to recover LDCT reconstructions to the best of our knowledge.

Despite being extensively used for learning spatio-temporal dependencies from sequential data [10], [11], recurrent methods can also embed the structure of traditional iterative algorithms. This way, Qin *et al.* [20] and Mardani *et al.* [21] proposed recurrent methods based on multiple iterations of GANs for artifact suppression in Magnetic Resonance (MR) imaging. However, our GRC-GAN differs from [20], [21] due to the use of *update* and *reset* gates that drive the cGAN

execution at each iteration t considering the computation of the previous $t-1$ time steps.

IV. Proposed Method

The main idea behind our *Gated Recurrent Conditional Generative Adversarial Network* (GRC-GAN) relies on the usage of network gates to drive recurrent executions of the cGAN for image denoising. More precisely, we define the expected behavior of such gates in Hypothesis 1:

Hypothesis 1. *considering the **reset/forget**, **update**, and **output** gates from the regular ConvGRU, we hypothesize that:*

(a) *the **reset/forget** gate can learn the location of the noise within a 2D input x_{t-1}, and subsequently attenuate the noise in those regions;*

(b) *the **update** gate is able to learn the regions of the solution computed by the **output** gate that are more accurate than the input x_{t-1}. Therefore, by constraining the focus of the **output** gate to a smaller area, its precision increases.*

The recurrence strategy to design the GRC-GAN does not involve learning spatio-temporal correlations within 3D reconstruction volumes. Instead, successive executions of the GRC-GAN over 2D slices individually intend to denoise image regions not recovered by the network on earlier executions. This idea guides our second hypothesis about the GRC-GAN:

Hypothesis 2. *the noise that remains after the first execution of the cGAN for image denoising can be solved by recurrent calls to the same model. At each execution, both the **reset/forget** and **update** gates are able to guide the processing into different features of the input data.*

Our proposed architecture is illustrated in Fig. 3. From the regular cGAN architecture, the GRC-GAN is obtained by replacing the cGAN generator (G) with a ConvGRU with G as its output gate. During the training stage, the weights of the networks R, U, G, and D are simultaneously updated according to:

$$\min_{R,G,U} \max_{D} V(R,U,G,D) = \mathbb{E}\left[\log\, D(y|\tilde{x}_T)\right] + \mathbb{E}\left[\log\left(1 - D(\tilde{x}_T)\right)\right], \tag{4}$$

where $\tilde{x}_T$ is the ConvGRU output defined in Equation (3) for T time steps, and x_0 is the LDCT image x.

V. Experiments

A. Database

The LoDoPaB-CT dataset [22] is composed of more than 40,000 CT scan slices of $\sim$ 800 patients selected from the LIDC/IDRI database [23]. In the LoDoPaB-CT dataset, the radiation dose attenuation is simulated by introducing Poisson noise into the projection data. Thus, low-dose and normal-dose reconstructions are obtained for each scan. All samples from the dataset were reconstructed using the regular Filtered Back-Projection (FBP) [24]. Fig. 4 shows examples of low-dose and normal-dose reconstructions for two scans.

Moreover, in favor of research reproducibility and comparability among solutions, the LoDoPaB benchmark is already divided into different subsets for training (35,820 images), validation (3,522 images), and testing (3,553 images).

B. Network Parameters

Our proposed GRC-GAN was built using the G and D networks from the cGAN described in [8], and the R and U as elementary networks composed of only three convolutional layers. The complete description of those networks are presented in the Tables I to III.

In the training stage, we used the Adam optimization algorithm to solve Equation (4) with learning rates $\alpha_G = 2 \times 10^{-4}$ for the network G and $\alpha_D = 10^{-4}$ for the network D, and number of GRC-GAN recurrences $r = 2$.

TABLE I
THE ARCHITECTURE OF THE NETWORKS R AND U IS COMPOSED OF THREE CONVOLUTIONAL BLOCKS STACKED.

Block	Layer	Kernel	Act. Function	Output Size
1	Conv 2D	3x3	LeakyReLU	362x362x32
	BatchNorm 2D			
2	Conv 2D	3x3	LeakyReLU	362x362x32
	BatchNorm 2D			
3	Conv 2D	3x3	Sigmoid	362x362x1

TABLE II
ARCHITECTURE OF THE NETWORK G USED IN THIS WORK. THIS IS THE GENERATOR NETWORK USED IN THE CGAN OF WOLTERINK *et al.* [8].

Block	Layer	Kernel	Act. Function	Output Size
1	Conv 2D	3x3	LeakyReLU	362x362x32
	BatchNorm 2D			
2	Conv 2D	3x3	LeakyReLU	362x362x32
	BatchNorm 2D			
3	Conv 2D	3x3	LeakyReLU	362x362x32
	BatchNorm 2D			
4	Conv 2D	3x3	LeakyReLU	362x362x32
	BatchNorm 2D			
5	Conv 2D	3x3	LeakyReLU	362x362x64
	BatchNorm 2D			
6	Conv 2D	3x3	LeakyReLU	362x362x128
	BatchNorm 2D			
7	Conv 2D	3x3	Tanh	362x362x1

C. Evaluation metrics

We used three metrics to quantify the accuracy of the predicted normal-dose CT image $\tilde{y}$ concerning the real image y, and they are: the Normalized Root Mean Squared Error (NRMSE), Peak-to-Signal Noise Ratio (PSNR), and Structural Similarity Index Measure (SSIM):

- The **Normalized Root Mean Squared Error (NRMSE)** quantifies the dissimilarity between two signals based on their subtraction. More precisely, it normalizes the square

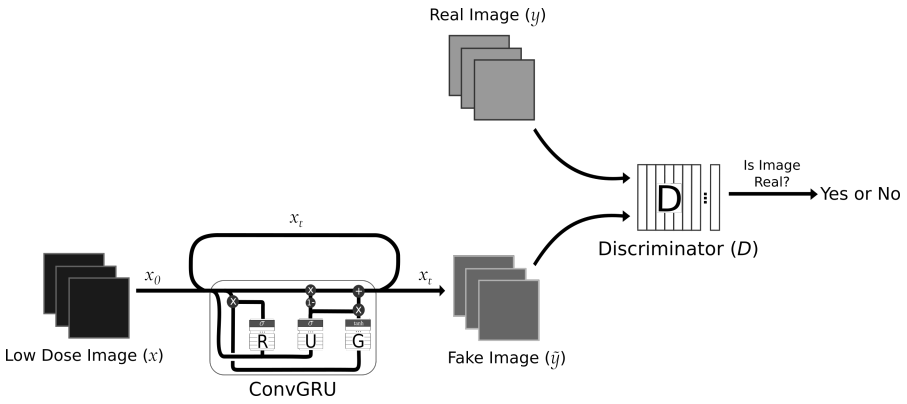

Fig. 3. The architecture of the proposed GRC-GAN: from the regular cGAN [8], the GRC-GAN is obtained by replacing the network G with the ConvGRU that has G as output gate. The ConvGRU output x_t is iterated for a fixed number of time steps. The current solution is updated by R, U, and G at each step.

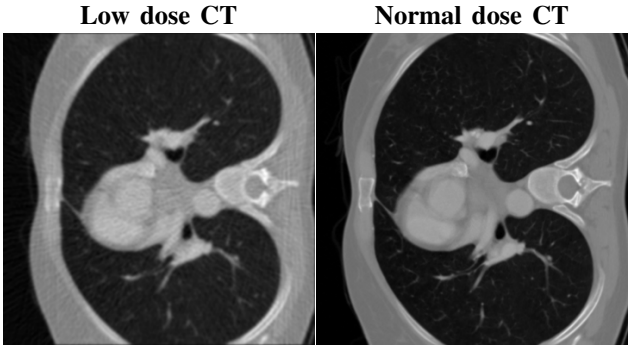

Fig. 4. Samples of low-dose (left) and normal-dose (right) chest CT scans from the LoDoPaB-CT dataset used in this work. Both samples were reconstructed with the regular Filtered Back-Projection (FBP) technique.

TABLE III
ARCHITECTURE OF THE NETWORK D USED IN THIS WORK. THIS IS THE DISCRIMINATOR NETWORK USED IN THE CGAN OF WOLTERINK *et al.* [8].

Block	Layer	Kernel	Act. Function	Output Size
1	Conv 2D	3x3	LeakyReLU	178x178x128
	BatchNorm 2D			
	Conv 2D	3x3	LeakyReLU	
	BatchNorm 2D			
	Conv 2D	3x3	LeakyReLU	
	BatchNorm 2D			
2	Conv 2D	3x3	LeakyReLU	86x86x64
	BatchNorm 2D			
	Conv 2D	3x3	LeakyReLU	
	BatchNorm 2D			
	Conv 2D	3x3	LeakyReLU	
	BatchNorm 2D			
3	Conv 2D	3x3	LeakyReLU	40x40x1
	BatchNorm 2D			
	Conv 2D	3x3	LeakyReLU	
	BatchNorm 2D			
	Conv 2D	3x3		
4	FullyConnected			256
5	FullyConnected		Sigmoid	1

root of the average of the squared errors between the signals $\tilde{y}$ and y:

$$NRMSE(\tilde{y}, y) = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\tilde{y}_i - y_i)^2}}{y_{max} - y_{min}} \tag{5}$$

where N is the dimension of both $\tilde{y}$ and y.

- The **Peak-to-Signal Noise Ratio (PSNR)** measures the signal fidelity with respect to corrupting noise. It computes the ratio between the highest possible signal power and the average power of the error:

$$PSNR(\tilde{y}, y) = 10 \cdot \log_{10}\left(\frac{MAX_y^2}{\frac{1}{N}\sum_{i=1}^{N}(\tilde{y}_i - y_i)^2}\right) \tag{6}$$

where MAX_y is the maximum possible value of the signal y.

- The **Structural Similarity Index Measure (SSIM)** [25] performs quality assessment based on the degradation of structural information. More precisely, the SSIM index combines the evaluation between $\tilde{y}$ and y in relation to luminance (l), contrast (c), and structure (s):

$$\begin{aligned} SSIM(\tilde{y}, y) &= [l(\tilde{y}, y)^{\alpha} \cdot c(\tilde{y}, y)^{\beta} \cdot s(\tilde{y}, y)^{\gamma}] \\ l(\tilde{y}, y) &= \frac{2\mu_{\tilde{y}}\mu_y + c_1}{\mu_{\tilde{y}}^2 + \mu_y^2 + c_1} \\ c(\tilde{y}, y) &= \frac{2\sigma_{\tilde{y}}\sigma_y + c_1}{\sigma_{\tilde{y}}^2 + \sigma_y^2 + c_2} \\ s(\tilde{y}, y) &= \frac{\sigma_{\tilde{y}y} + c_3}{\sigma_{\tilde{y}}\sigma_y + c_3} \end{aligned} \tag{7}$$

where μ represents the average of the signal $\tilde{y}$ or y, and σ^2 its variance. Furthermore, $\sigma_{\tilde{y}y}$ is the co-variance, and $c_1, c_2,$ and c_3 are variables that work to stabilize the division with weak denominator. Finally, $\alpha, \beta,$ and γ are weights usually set as 1.

D. Results and Discussion

From the trained GRC-GAN, we extracted the signals $R(x_0)$ and $U(x_0)$ computed for a given input data x_0 in order to

evaluate Hypothesis 1 about the role of each gate to compose the final solution. Fig. 5 shows the input image x_0 in (a), the position of the greatest errors in x_0, *i.e.*, $\{||x_0 - y|| > \epsilon_1 | P(||x_0 - y|| > \epsilon_1) = 0.15\}$ in (b), the pixels most likely to be forgotten according to the *reset/forget* gate, *i.e.*, $\{R(x_0) < \epsilon_2 | P(R(x_0) < \epsilon_2) = 0.15\}$ in (c), and pixels most likely to be updated according to the *update* gate, *i.e.* $\{U(x_0) > \epsilon_3 | P(U(x_0) > \epsilon_3) = 0.15\}$ in (d).

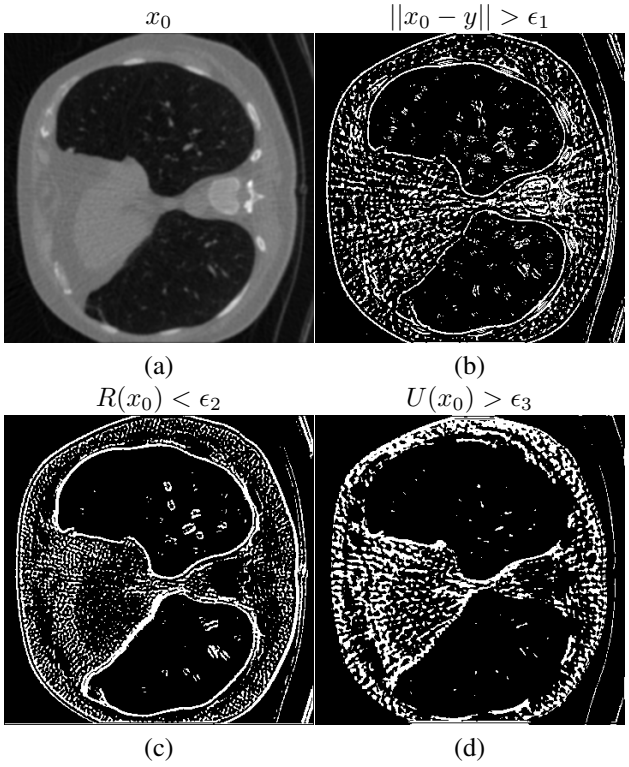

Fig. 5. Signals that illustrates the cGAN execution according to our Hypothesis 1 defined in section IV: an input low-dose reconstruction x_0 (a), the pixels of x_0 with the most significant noise with respect to its ground truth y (b), the pixels indicated by $R(x_0)$ to be mostly forgotten in the current time step (c), and the pixels indicated by $U(x_0)$ to be primarily updated in the current time step (d).

We can see that the pixels with the highest noise levels in x_0 are the most indicated to be forgotten by $R(x_0)$ and updated by $U(x_0)$. The gates output is not discrete, but we show them this way to help our visualization. Those results compose evidence in favor of our first hypothesis presented in section IV.

In order to evaluate Hypothesis 2 about the effect of recurrent calls of the network for different denoising regions of the image, Fig. 6 highlights the pixels with higher noise at (a) the input image x_0, (b) the partial solution after the first iteration, and (c) the final solution after the second iteration. The proposed model reduces the noise that remained after the first denoise iteration in the second iteration. Thus, we also have strong evidence on the validation of our Hypothesis 2.

In relation to the final reconstructions of the GRC-GAN versus the regular cGAN, Fig. 7 shows two samples of low-dose CT slices recovered by each method. We also show the

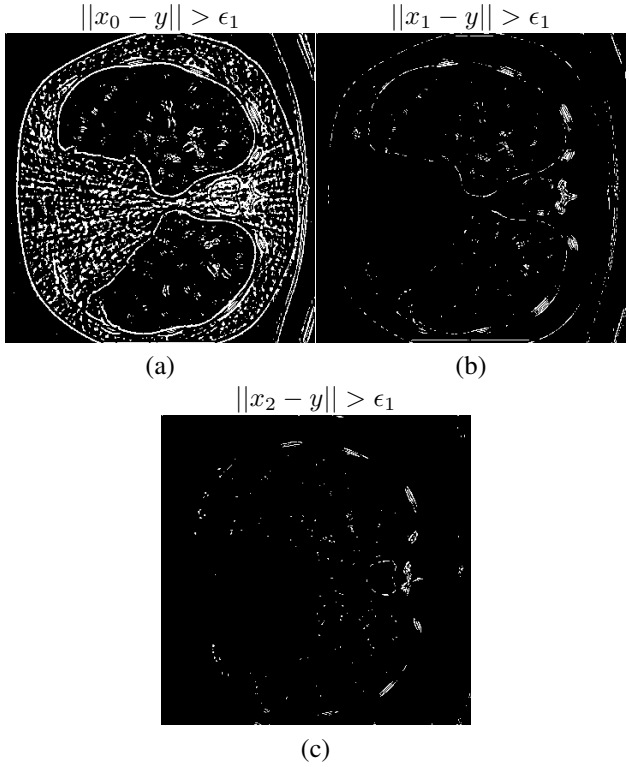

Fig. 6. Error maps that support our Hypothesis 2 defined in the section IV: the noise level in the input image (a) is highly decreased during the first network iteration (b). Furthermore, much of the noise that remains after the first iteration is reduced during the second denoise iteration (c).

regular FBP reconstruction (input) and the ground truth.

The error maps of each reconstruction shown previously are highlighted in Fig. 8. In order to help the human visualization, each error map in Fig. 8 was processed by the intensity transform $s = ||1 - r^{0.35}||$, where gray values of the input and output pixels are represented by r and s, respectively. Notice that such transform also inverted the gray values. Thus, higher pixel values are darker. The error maps of the cGAN reconstructions have areas much lighter than in the FBP error maps, and this represents a clear denoise of the cGAN in relation to the FBP input. However, the GRC-GAN error maps are even lighter than the ones from the cGAN, indicating a much higher denoise capacity of our proposed method.

Finally, we present the mean and standard deviation of SSIM, PSNR, and NRMSE in Table IV for all the methods evaluated over the test set comprising 3,553 CT slices. Those results indicate an average improvement on the denoise accuracy produced by the GRC-GAN concerning the regular cGAN of 17% and 36% when measured by PSNR and NRMSE, respectively. Not surprising, there was no significant difference in SSIM since the low-dose degradation observed in the LoDoPaB-CT dataset does not involve the presence of artifacts deforming the structural shape of the human internal tissues.

VI. Conclusion

In this paper, we described a new network architecture for image denoise: the GRC-GAN. Inspired by the success of the

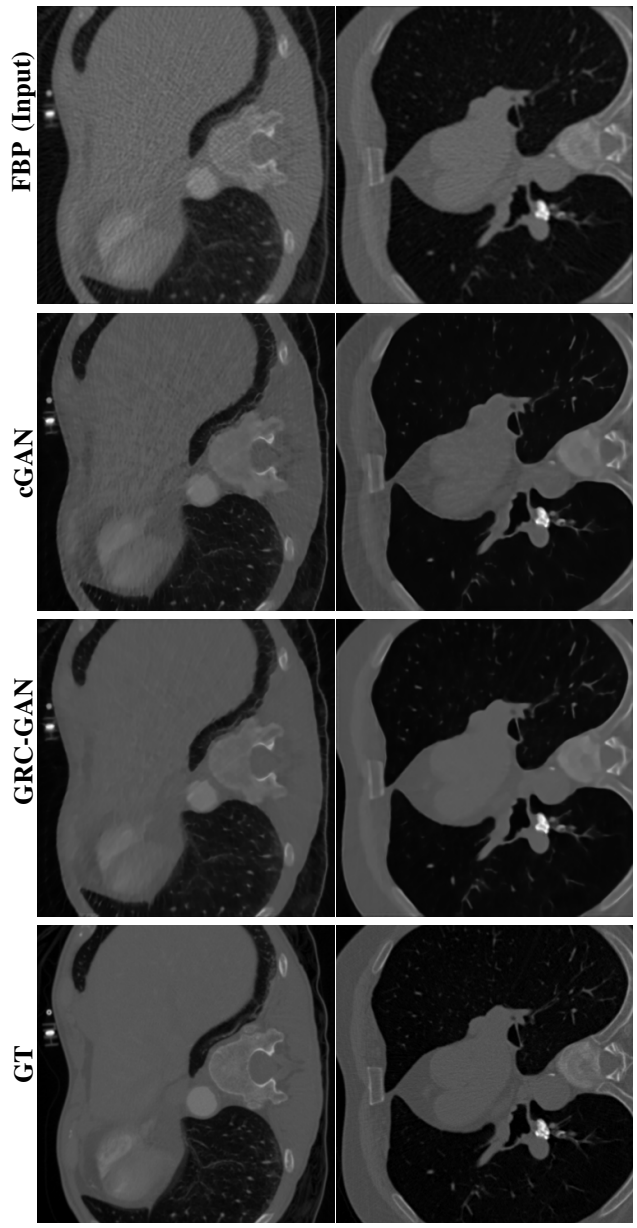

Fig. 7. Two samples of the regular FBP reconstruction from low-dose CT projections, the solutions obtained by the cGAN, our proposed GRC-GAN, and the ground truth.

TABLE IV
AVERAGE AND STANDARD DEVIATION OF IMAGE QUALITY METRICS FOR ALL THE METHODS EVALUATED OVER THE LODOPAB-CT TEST SET.

Method	SSIM	PSNR	NRMSE
FBP	0.7949 (0.091)	24.2 (3.7)	0.21 (0.107)
cGAN	0.8804 (0.09)	25.8 (3.2)	0.17 (0.057)
GRC-GAN	**0.8871 (0.094)**	**30.2 (3.9)**	**0.11 (0.069)**

original cGAN, we designed a recurrent cGAN that includes a *reset/forget* and an *update* gate to drive the model's attention during sequential executions for image denoising.

We conducted experiments using the LoDoPaB-CT dataset [22]. It is composed of more than 40,000 CT scanning slices of the human chest from 800 patients. Then, CT measurements of low photon count were created from the original X-ray projections with a parallel beam scan geometry. As a result, the LoDoPaB-CT is a standard dataset to train and benchmark learned low-dose CT reconstruction methods. According to our experimental analysis, the network gates can detect the regions of higher noise in the input image and then direct the network effort to update these regions first. Moreover, the recurrent executions showed themselves efficiently to improve the denoise accuracy obtained in previous steps. Finally, the average NRMSE of the reconstructions denoised by GRC-GAN is 36% lower than those denoised by the cGAN, and the average PSNR is 17% greater.

The results indicate that new gated recurrent architectures may overcome many encoder-decoder networks currently used in the state-of-the-art for image denoise. Moreover, our gated model is part of a new generation of explainable artificial intelligence since it can indicate its attention regions at each time step.

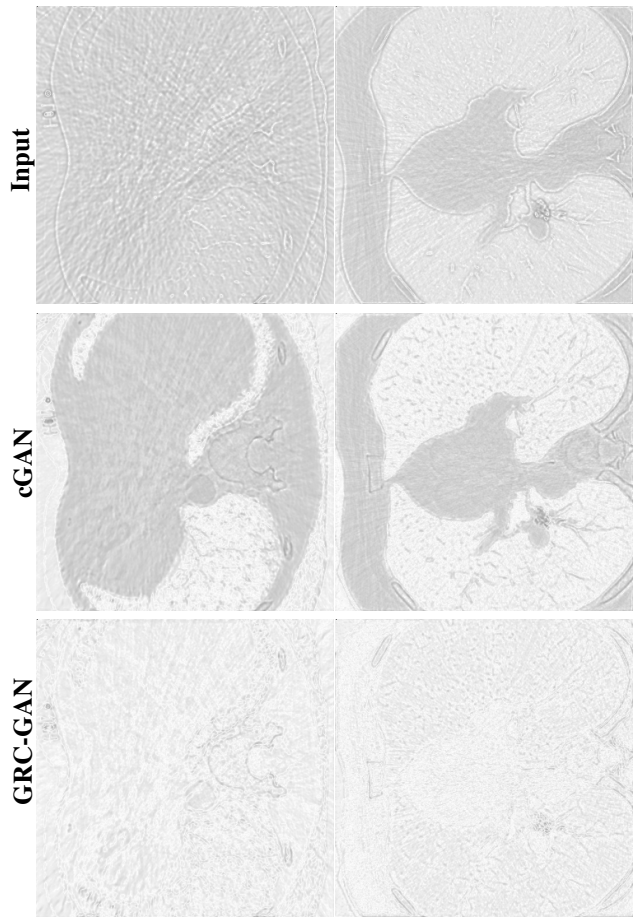

Fig. 8. Error maps of the reconstructions showed in Fig. 7 enhanced by $s = ||1 - r^{0.35}||$ to help human visualization. In such intensity transform, r is the gray value of the input pixel, and s its respective output. Such intensity transform also inverted the gray values to improve visualization. Thus, higher gray values are darker. We can see that the error maps produced by the GRC-GAN are much lighter than the ones produced by the cGAN. This indicates a much higher denoise accuracy of the proposed GRC-GAN in relation to the regular cGAN.

REFERENCES

[1] J. Liu, H. Yu, and S. Zhang, "The indispensable role of chest ct in the detection of coronavirus disease 2019 (covid-19)," *European journal of nuclear medicine and molecular imaging*, vol. 47, no. 7, pp. 1638–1639, 2020.

[2] L. Krille, G. P. Hammer, H. Merzenich, and H. Zeeb, "Systematic review on physician's knowledge about radiation doses and radiation risks of computed tomography," *European journal of radiology*, vol. 76, no. 1, pp. 36–41, 2010.

[3] M. M. Rehani, K. Yang, E. R. Melick, J. Heil, D. Šalát, W. F. Sensakovic, and B. Liu, "Patients undergoing recurrent ct scans: assessing the magnitude," *European radiology*, vol. 30, no. 4, pp. 1828–1836, 2020.

[4] M. M. Rehani, E. R. Melick, R. M. Alvi, R. D. Khera, S. Batool-Anwar, T. G. Neilan, and M. Bettmann, "Patients undergoing recurrent ct exams: assessment of patients with non-malignant diseases, reasons for imaging and imaging appropriateness," *European radiology*, vol. 30, no. 4, pp. 1839–1846, 2020.

[5] 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.

[6] Y. Han and J. C. Ye, "Framing u-net via deep convolutional framelets: Application to sparse-view ct," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1418–1429, 2018.

[7] S. Benjamens, P. Dhunnoo, and B. Meskó, "The state of artificial intelligence-based fda-approved medical devices and algorithms: an online database," *NPJ digital medicine*, vol. 3, no. 1, pp. 1–8, 2020.

[8] 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.

[9] N. Ballas, L. Yao, C. Pal, and A. Courville, "Delving deeper into convolutional networks for learning video representations," *arXiv preprint arXiv:1511.06432*, 2015.

[10] O. Mogren, "C-rnn-gan: Continuous recurrent neural networks with adversarial training," *arXiv preprint arXiv:1611.09904*, 2016.

[11] C. Esteban, S. L. Hyland, and G. Rätsch, "Real-valued (medical) time series generation with recurrent conditional gans," *arXiv preprint arXiv:1706.02633*, 2017.

[12] C. Qin, J. Schlemper, J. Caballero, A. N. Price, J. V. Hajnal, and D. Rueckert, "Convolutional recurrent neural networks for dynamic mr image reconstruction," *IEEE transactions on medical imaging*, vol. 38, no. 1, pp. 280–290, 2018.

[13] E. Tjoa and C. Guan, "A survey on explainable artificial intelligence (xai): Toward medical xai," *IEEE Transactions on Neural Networks and Learning Systems*, 2020.

[14] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[15] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *arXiv preprint arXiv:1411.1784*, 2014.

[16] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.

[17] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, and G. Wang, "Low-dose ct image denoising using a generative adversarial network with wasserstein distance and perceptual loss," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1348–1357, 2018.

[18] Z. Liu, T. Bicer, R. Kettimuthu, D. Gursoy, F. De Carlo, and I. Foster, "Tomogan: Low-dose synchrotron x-ray tomography with generative adversarial networks," *arXiv preprint arXiv:1902.07582*, 2019.

[19] X. Yi and P. Babyn, "Sharpness-aware low-dose ct denoising using conditional generative adversarial network," *Journal of digital imaging*, vol. 31, no. 5, pp. 655–669, 2018.

[20] C. Qin, J. Schlemper, J. Caballero, A. N. Price, J. V. Hajnal, and D. Rueckert, "Convolutional recurrent neural networks for dynamic mr image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 38, no. 1, pp. 280–290, 2019.

[21] M. Mardani, H. Monajemi, V. Papyan, S. Vasanawala, D. Donoho, and J. Pauly, "Recurrent generative adversarial networks for proximal learning and automated compressive image recovery," *arXiv preprint arXiv:1711.10046*, 2017.

[22] J. Leuschner, M. Schmidt, D. O. Baguer, and P. Maass, "Lodopab-ct, a benchmark dataset for low-dose computed tomography reconstruction," *Scientific Data*, vol. 8, no. 1, pp. 1–12, 2021.

[23] S. G. Armato III, G. McLennan, L. Bidaut, M. F. McNitt-Gray, C. R. Meyer, A. P. Reeves, B. Zhao, D. R. Aberle, C. I. Henschke, E. A. Hoffman *et al.*, "The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans," *Medical physics*, vol. 38, no. 2, pp. 915–931, 2011.

[24] R. Schofield, L. King, U. Tayal, I. Castellano, J. Stirrup, F. Pontana, J. Earls, and E. Nicol, "Image reconstruction: Part 1 – understanding filtered back projection, noise and image acquisition," *Journal of Cardiovascular Computed Tomography*, vol. 14, no. 3, pp. 219–225, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1934592519300607

[25] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.