Deep Learning based Metal Artifacts Reduction in post-operative Cochlear Implant CT Imaging

Zihao Wang, Clair Vandersteen, Thomas Demarcy, Dan Gnansia, Charles Raffaelli, Nicolas Guevara, Hervé Delingette

Deep Learning based Metal Artifacts Reduction in post-operative Cochlear Implant CT Imaging

Zihao Wang[1], Clair Vandersteen[2], Thomas Demarcy[3], Dan Gnansia[3], Charles Raffaelli[2], Nicolas Guevara[2], and Hervé Delingette[1]

[1] Université Côte d'Azur, Inria, Epione Team, France
zihao.wang@inria.fr
[2] Université Côte d'Azur, Nice University Hospital, France
[3] Oticon Medical, France

Abstract. To assess the quality of insertion of Cochlear Implants (CI) after surgery, it is important to analyze the positions of the electrodes with respect to the cochlea based on post-operative CT imaging. Yet, these images suffer from metal artifacts which often entail a difficulty to make any analysis. In this work, we propose a 3D metal artifact reduction method using convolutional neural networks for post-operative cochlear implant imaging. Our approach is based on a 3D generative adversarial network (MARGANs) to create an image with a reduction of metal artifacts. The generative model is trained on a large number of pre-operative "artifact-free" images on which simulated metal artifacts are created. This simulation involves the segmentation of the scala tympani, the virtual insertion of electrode arrays and the simulation of beam hardening based on the Beer-Lambert law.
Quantitative and qualitative evaluations compared with two classical metallic artifact reduction algorithms show the effectiveness of our method.

Keywords: Generative adversarial networks, Metal Artifacts Reduction

1 Introduction

The physical imaging process of spiral CT leads to the creation of artifacts in the reconstructed images when dense objects such as metallic parts are fully absorbing the X-rays. Such artifacts are for instance visible in CT post-operative images of the inner ear following the surgical insertion of Cochlea Implants (CI) due to the presence of a metallic electrode array. Fig. 1 (a) shows the image distortions around the metallic electrodes implanted along the scala tympani, one of three ducts constituting the cochlea with the scala vestibuli and scala media (see Fig. 1 Right). Those artifacts make the post-operative assessment of CI implantation very difficult in particular the estimation of the relative positions of the electrodes with respect to scala tympani and the scala vestibuli.

There is a vast body of research efforts for Metal Artifact Reduction (MAR) approaches, for instance based on iterative reconstruction or corrections of the physical effects. Deep learning MAR methods were also recently introduced.

Zhang *et al.* [11] introduced CNN as prior for filtered back projections based method. A simulation dataset was built for training the CNN. Huang *et al.* [3] developed CNN (*RL-ARCNN*) directly on the image domain for removing metal artifacts in cervical CT images based on a residual learning method in an end to end manner.A neural network named *DestreakNet* was proposed for streak artifacts reduction[2]. *DestreakNet* is used for quality improvement as post-processing after the application of the state of the art NMAR interpolation-based algorithm. Some details that were lost after the NMAR processing can be recovered by the *DestreakNet* network.

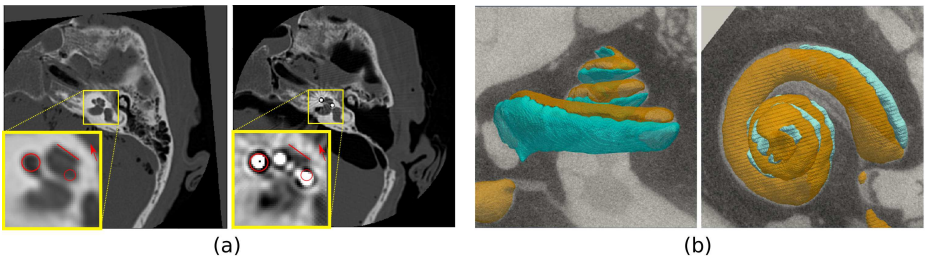

Fig. 1. (a) Pre-operative (Left) and post-operative (Right) CT imaging of the inner ear. Cochlear implant creates metallic artifacts. (b) μCT images of the cochlea including the scala tympani (green) and scala vestibuli (orange).

For CI metal artifacts reduction, conditional generative adversarial networks (cGAN) were proposed by Wang *et al.*[8] to generate artifact-free images. The cGAN operates on each 2D slice and was trained on 76 pairs of registered pre and post-operative images. The registration of those image pairs is very challenging precisely due to those artifacts and the 2D approach does not guarantee spatial consistency of the result from one slice to the next.

In this paper, we propose a 3D generative adversarial network for MAR on cochlear implants. Instead of training the GANs on pairs of pre and postoperative images, our approach relies on the physical simulation of artifacts in CT images from pre-operative CT images. The simulation involves the automatic segmentation of the scala tympani on preoperative CT, the automatic estimation of the positions of the electrode arrays and finally the simulation of beam hardening on fan-beam projections based on the Beer-Lambert law discretized on 5 different energies. This approach was applied on 1090 3D preoperative CT images to create 1090 3D images with simulated artifacts. The 1090 image pairs are then used to train an original 3D generative adversarial network (GANs) named MARGANs for which a generative loss derived from Retinex theory was introduced. This unsupervised approach was successfully tested on 10 post-operative images to generate artifact reduced images and its quantitative performance was favorably compared with two other MAR methods. Unlike [8], our method is unsupervised and does not require any registration of pre and

post operative images which is a difficult task. The use of a 3D GAN improves a lot the spatial consistency of the generated images. It is also generic to the type of artifacts without requiring pre and postoperative images which is not available in many surgery procedures.

2 Method

2.1 CI Metal Artifacts Simulation

CI Electrode Array Simulation The processing pipeline to generate the training set for the MARGANs is displayed in Fig. 2. Given a preoperative cochlea CT image, it is first automatically rigidly registered on a reference cochlea CT image where the modiolus axis is along the Z direction. The segmentation of the cochlea is then performed through a parametric cochlea shape model similarly to the approach proposed by Demarcy *et al.* [1]. As a result of the segmentation, the scala tympani (ST) is extracted and then a signed distance map of the ST is generated (Step 2). The positions of the simulated electrode array are estimated by thresholding the distance map to create a 3D tubular binary mask near the center of the ST (Step 3). The Hounsfield unit of simulated electrode array was then set to $3071HU$ (Step 4), which is the maximum observed value on CI metal artifacts.

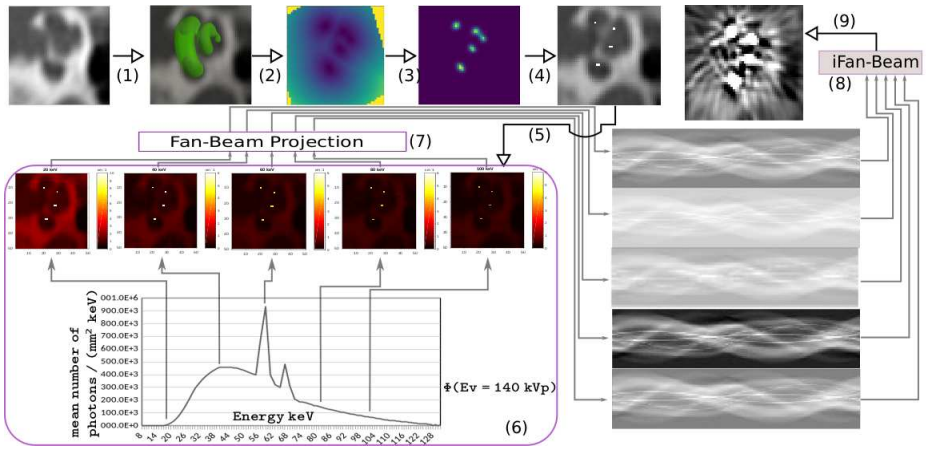

Fig. 2. CI metal artifacts simulation workflow starting from a pre-operative image and ending with the simulated post-operative image after 9 processing steps.

Metal Artifacts Simulation We rely on the by Beer-Lambert law, relating the number of photons $I(x, y, \delta z)$ for a monoenergetic X-ray source at point (x, y) through a slice δz to the the initial photon number I_0 and linear absorption

coefficient $\mu(x, y)$: $I(x, y, \delta z) = I_0 e^{-\mu(x,y)\delta z}$. For a polychromatic X-ray beam, the attenuation coefficient depends on the chromatic photon energy $\mu(E_v)$, and the energy spectrum $\phi(E_v)$ must be taken into account. The number of photons received in the entire detector surface is then:

$$I = \int_{E_0}^{E_n} (\phi(E_v) e^{-\iiint \mu(x,y,z,E_v) dx dy dz} + S(E_v)) dE_v \quad (1)$$

where E_0 and E_n are the minimum and maximum energies for a fixed tube peak voltage, and where $S(E_v)$ captures X-ray scattering. The energy spectrum $\phi(E_v)$ was downloaded from a CT manufacturer dedicated site [4] for a tungsten anode tube at 140 kVp.

The simulation starts by computing attenuation maps $\mu(x, y, z, E_{v_i})$ (Step 5) for five sample energies shown in Fig. 2(6). This computation is based on the Hounsfield unit formula and the water absorption coefficients as a function of energy[9]. We then perform fan-beam projection (Step 7) of the 5 attenuation maps to produce sinograms-like images representing absorbed energy on CT detectors. Then a weighted sum of the 5 sinograms (Step 8) is computed (including energy spectrum and scatter) as a discretization of Eq. 1. Finally inverse fan beam projection produces the output image with metallic artifacts (Step 9).

2.2 CI CT image Metal Artifacts Reduction based on 3D GANs

MARGANs The issue of removing metal artifact from images is clearly an ill-posed problem as it requires to restore information unseen in the input image. To tackle this problem, we propose a 3D generative adversarial approach for reducing metal artifacts inspired by the Super-Resolution (SR) GANs of Sánchez *et al.*[6] devised for brain MR images super-resolution. The idea of MARGANs is to combine two neural networks, the former generating the MAR images (generator) and the latter discriminating between images containing metal artifacts or not. Formally, given an input image $I^{metal}(\mathbf{x})$ with metal artifacts, the generator neural networks G_{w_g} with parameters w_g performs the mapping between I^{metal} and the reconstructed MAR image I^{MAR}. The discriminator neural network D_{w_d} evaluates the realism of the generated MAR I^{MAR} or any artifact-free image I. To train the G_{w_g} and D_{w_d} networks, the classical GANs optimize the following objective function as the sum of the discriminator and generator losses:

$$\min_{w_g} \max_{w_d} U = \mathbb{E}_{x \sim I}[\log D_{w_d}(x)] + \mathbb{E}_{y \sim I^{metal}}[\log (1 - D_{w_d}(G_{w_g}(y)))] \quad (2)$$

Network Architecture The generator network architecture vastly differs from [6] as it is similar to U-Net with convolution and deconvolution layers, and batch normalization layers to improve the training efficiency. Moreover, unlike [6] which is patch-based, the input of the network consist of full 3D images as it easily fits in GPU memory. The number of filters increases from 1 to N_f and

[4] https://www.oem-xray-components.siemens.com/x-ray-spectra-simulation

then decreases to 1, where N_f is the maximum number of feature maps, which depends on the GPU memory size. The discriminator network follows that of [6] with eight groups of convolution layers and batch normalization layers combined sequentially. Instead of using the original GANs formula, MARGANs is trained by minimizing both the discriminator loss and generator loss.

Discriminator Loss The discriminator network D_{w_d} loss should consider differently the artifact-free images I from images with metal artifacts I^{metal}. we proposed loss for the discriminator is:

$$L_{discriminator} = \log(|D_{w_d}(I) - 1|) + \log(D_{w_d}(G_{w_g}(I^{metal})) \tag{3}$$

Generator Loss We propose a new loss function based on Retinex theory [5] to measure the discrepancy in the reconstruction of the MARGANs. Retinex theory is mostly used to improve images seriously affected by environmental illumination. We propose a novel generator loss based on the Retinex theory consisting in two parts : the content and Retinex penalties. For content penalty, mean square error (MSE), $L_{mse} = \mathbb{E}(I - G_{w_g}(I^{MAR})$ were used to encourage the generator to generate voxels consistent with the artifact free images. But using only the MSE loss leads to blurred MAR images with a lack of high frequency image details. To avoid this excessive smoothing, we add the Retinex penalty written as follows to get numerically stable evaluations:

$$L_{retinex} = \frac{1}{|I^{metal}|} \sum_{\mathbf{x}} |G_{w_g}(I^{metal}) - e^{\log I^{MAR} - \log I^{MAR} * \mathcal{N}(0,\sigma)}| \tag{4}$$

The Retinex theory assumes that a given image can be expressed by the product of environmental brightness $L(x, y)$ and the object reflectance $R(x, y)$. To get the high frequency information of the object reflectance $R(c, y)$ we use Gaussian function to remove the low frequency part of the image as done in [10] (known as single-Scale Retinex algorithm) for ultrasound image enhancement, $\mathcal{N}(0, \sigma)$ is a Gaussian function of standard deviation σ and $*$ is the convolution operator. This loss enforces salient features in the image that would be attenuated otherwise. Putting it all together with the adversarial term $L_{adv} = \frac{1}{2}|D_{w_d}(G_{w_g}) - 1|^2$ as in [6] , the full loss function of the generator is:

$$L_{generator} = \alpha \cdot L_{retinex} + L_{mse} + L_{adv} \tag{5}$$

where α is a parameter controlling the influence of the Retinex loss.

3 Experiments and Evaluation

3.1 Dataset

Simulation Dataset CT scans of the temporal bones include 493 left and 597 right images collected from 597 patients by a GE LightSpeed CT scanner with

a standard protocol (without metal artifact reduction filters) at the Radiology Department of the Nice University Hospital. Imaging voxel size is $0.2 \times 0.2 \times 0.2\ mm^3$. All images were registered by a pyramidal blocking-matching algorithm to find the ROI and then cropped as $60 \times 50 \times 50$ volume images.

Evaluation Dataset The evaluation dataset includes 10 temporal bones images outside the simulation dataset but collected with the same conditions. It additionally includes pre-operative and post-operative pairs of images. Due to resected tissues and artifacts caused by implanted electrodes during surgery, all tested rigid registration algorithms fail to register the pre-operative with post-operative images. Therefore they were manually registered in 3D using landmarks and then both were cropped similarly to the simulation dataset.

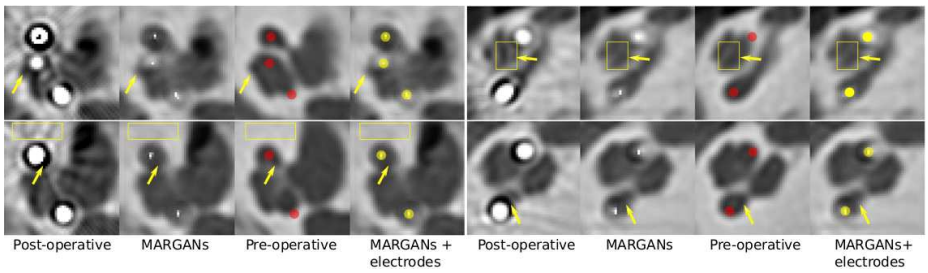

Fig. 3. Results on patients #4 (Left) and #6 (Right) for two middle slices (First and second rows). The 4 columns correspond respectively to : original post-operative images, output of MARGANs, registered pre-operative images with manually positioned electrodes in red and post-operative images with electrodes in yellow.

3.2 Implementation Details

The proposed GANs were implemented with TensorFlow following section 2. Convolution kernel size is set to $3 \times 3 \times 3$, the number of filters was $N_f = 512$ and the weight of Retinex loss was set as $\alpha = 0.00002$ experimentally. The generator and discriminator were trained by using RMSprop optimizer with learning rate $l_{rg} = 1\text{e}{-4}$ and $l_{rd} = 1\text{e}{-3}$ respectively. The simulation of metal artifacts took about 1 hour on a CPU cluster for each of the 1090 images. Training the MARGANs took about 23 hours on one NVIDIA 1080Ti GPU.

3.3 Performance on Clinical Data

Performance on Evaluation Dataset Fig. 3 shows the application of the proposed method on two middle slices of two patient images, including registered pre-operative images. The metal artifacts are significantly reduced in the MARGANs generated images without important geometry distortions. Furthermore,

some visible internal structures inside the cochlea are restored by the MARGANs as shown with added arrows. The electrodes positions in yellow and red were manually added to allow for the visual assessment of the electrodes with respect to the pre-operative images.

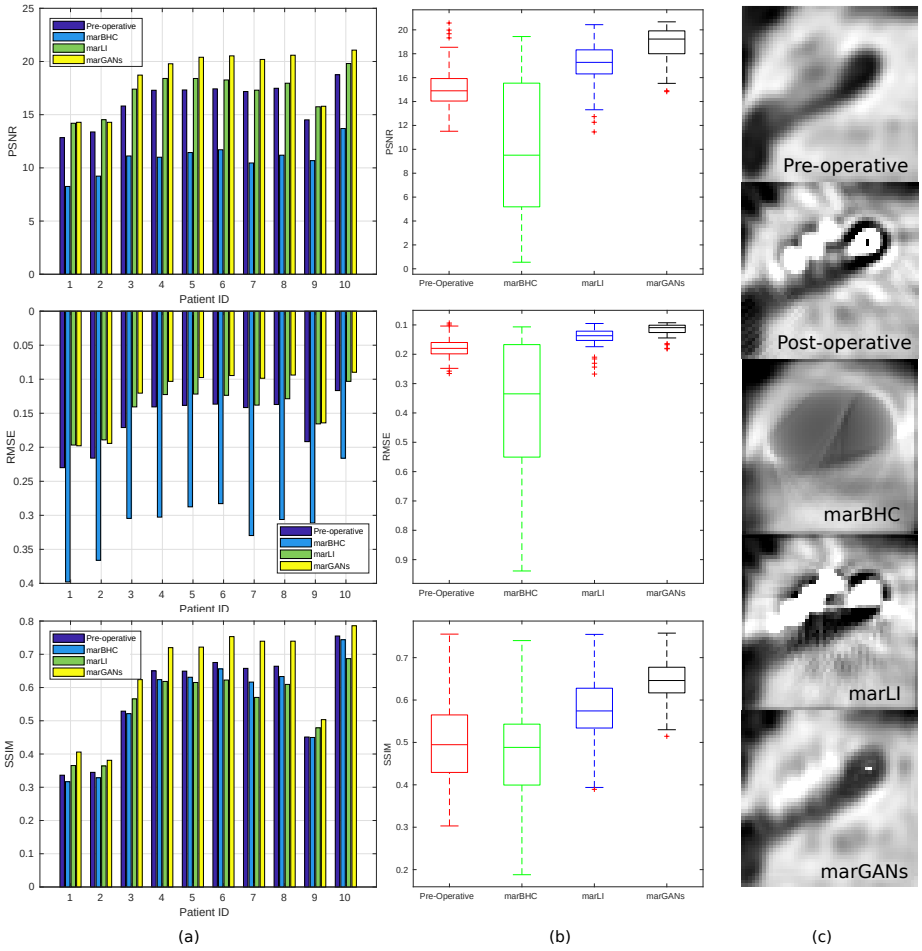

Fig. 4. Comparison based on three metrics (PSNR, RMSE, and SSIM) of the proposed MARGANs methods with two other MAR algorithms : marLI and marBHC. (a) overall performance on 10 clinical post-operative images, (b) the 3D consistency between slices for patient #4, (c) middle slice generated by the different methods.

Performance Comparison MARGANs is compared with 2 open-source 2D fast metallic artifacts reduction methods: projection linear interpolation replace-

ment [4] (marLI) and beam hardening correction [7] (marBHC). The 10 evaluation dataset images were processed by marLI, marBHC slice by slice. In Fig. 4(a), for each of the 10 images, we compare the Root Mean Square Error (RMSE), Structural Similarity Index (SSIM) and Peak Signal to Noisy Ratio (PSNR) computed between the pre-operative images and the MAR images generated by those two methods and our proposed approach. Those indices measure the preservation of visible structures, the errors and quality of the reconstructed images. We show that our proposed method outperforms those 2 MAR methods for all three metrics. In Fig. 4 column (b), the same indices are shown for all slices of patient #4 to evaluate the spatial consistency of the reconstruction. Clearly MARGANs exhibits the best performances with a lower mean value and much lower variance. This can be explained by the fact that it is the only MAR algorithm working directly on 3D images. The visualization of the 3 different methods on one middle slice from patient #4 can be found in Fig. 4(c). Clearly, the output of MARGANs where the metal artifacts are largely reduced is closer to the pre-operative image compared to other methods.

4 Conclusion

In this paper we presented an unsupervised framework for generating images with reduced metal artifacts from CI postoperative images. Our 3D approach is based on a GANs that was trained on a large number of preoperative and simulated images. The simulation was based on a segmentation algorithm of the cochlea and the simulation of beam hardening for the metal artifacts. Quantitative and qualitative comparisons with two fast 2D MAR methods show the superiority of the proposed approach. The performance of MARGANs may be improved in the future (i) by including more physically realistic metal artifacts in the simulation such as noisy detectors and exponential edge-gradient effects etc. (ii) by using supervised learning with annotated pairs of CT images.

Acknowledgements This work was partially funded by the regional council of Provence Alpes Côte d'Azur, by the French government through the UCA$^{\text{JEDI}}$ "Investments in the Future" project managed by the National Research Agency (ANR) with the reference number ANR-15-IDEX-01, and was supported by the grant AAP Santé 06 2017-260 DGA-DSH.

References

1. Demarcy, T., Vandersteen, C., Guevara, N., Raffaelli, C., Gnansia, D., Ayache, N., Delingette, H.: Automated analysis of human cochlea shape variability from segmented μCT images. Computerized Medical Imaging and Graphics **59**(July 2017), 1–12 (2017)
2. Gjesteby, L., Shan, H., Yang, Q., Xi, Y., Claus, B., Jin, Y., De Man, B., Wang, G.: Deep Neural Network for CT Metal Artifact Reduction with a Perceptual Loss Function. In: The fifth international conference on image formation in X-ray computed tomography. pp. 439–443. Salt Lake city (May 2018)

3. Huang, X., Wang, J., Tang, F., Zhong, T., Zhang, Y.: Metal artifact reduction on cervical CT images by deep residual learning. BioMedical Engineering OnLine **17**(175) (November 2018)
4. Kalender, W.A., Hebel, R., Ebersberger, J.: Reduction of CT artifacts caused by metallic implants. Radiology **164**(2), 576–577 (1987)
5. Land, E.H., McCann, J.J.: Lightness and Retinex theory. Journal of the Optical Society of America **61**(1), 1–11 (January 1971)
6. Sánchez, I., Vilaplana, V.: Brain MRI super-resolution using 3D generative adversarial networks. In: 1st Conference on Medical Imaging with Deep Learning (MIDL 2018). Amsterdam (2018)
7. Verburg, J.M., Seco, J.: CT metal artifact reduction method correcting for beam hardening and missing projections. Physics in Medicine and Biology **57**(9), 2803–2818 (April 2012)
8. Wang, J., Zhao, Y., Noble, J.H., Dawant, B.M.: Conditional generative adversarial networks for metal artifact reduction in CT images of the ear. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. pp. 3–11. Springer International Publishing, Granada, Spain (2018)
9. Wunderlich, A., Noo, F.: Image covariance and lesion detectability in direct fan-beam X-ray computed tomography. Physics in Medicine and Biology **53**(10), 2471–2493 (April 2008)
10. Zhang, R., Yali, H., Zhen, Z.: A ultrasound liver image enhancement algorithm based on multi-scale Retinex theory. In: 2011 5th International Conference on Bioinformatics and Biomedical Engineering. pp. 1–3. Wuhan, China (May 2011)
11. Zhang, Y., Yu, H.: Convolutional neural network based metal artifact reduction in X-ray computed tomography. IEEE Transactions on Medical Imaging **37**(6), 1370–1381 (June 2018)