Article Review : Pseudo-contrast cardiac CT angiography derived from non- contrast CT using conditional generative adversarial networks [5]

Simon Queric and Vincent Herfeld

June 13, 2024

1 Paper summary

Introduction Data-driven decision making is transforming healthcare. However, machine learning models often require significant amounts of data to be trained effectively and produce high-quality outputs. In healthcare, big data is not always available and it's a problem to design deep learning models that can generalise. Therefore, generating high-quality synthetic datasets could be an interesting idea for data augmentation for instance. Building a generative model that produces high quality data with low bias is quite challenging. One needs to be cautious and rigorous in the evaluation of the model ([2]).

Cardiac contrast-enhanced computed tomography angiography (CTA) is increasingly used to assess heart anatomy and function in various cardiac conditions like coronary artery disease or valvular heart disease. This acquisition process involves ionizing radiation and contrast agent administration, which can harm kidneys and cause allergic reactions. Non-contrast CT (NCCT) scans of the heart primarily evaluate calcification but have limited clinical use. That said they are acquisitions that are made in most of cardiac CT examinations and are safer than their CTA counterpart. During the COVID-19 pandemic, researchers proposed a method relying on a conditional generative adverserial network (cGAN) to generate synthetic chest CT scans [3]. Others [1] also proposed a method to synthesise pseudo chest CTA images from true NCCT scans. In this article the authors take the same ideas but this time apply them to cardiac NCCT images. The goal is to obtain images with the same clinical possibilities as real CTA images, while avoiding the invasive acquisition process. The use of this technology could assist radiologists in interpreting CT images, potentially resulting in improved diagnoses.

Proposed method Authors preprocessed the images which is a common practice in medical imaging, especially registering such that images are properly aligned. The authors rely on an Image-to-Image architecture developed in [7]. This architecture is a conditional GAN that learns a conditional generative model : knowing the original NCCT scan generate a pseudo-CTA scan. It follows the adversarial scheme where we train a generator to generate images that are as real as possible and a discriminator that has to learn to classify images as either real or pseudo. This alternating optimization causes the generator to generate pseudo scans that are very close to real scans. The obtained generator is a map of the form $G : (z, x) \mapsto y$, where z is usually Gaussian noise, x the input NCCT scan and y the output pseudo-CTA scan. The final model contains for the Generator a U-Net architecture, and for the discriminator a PatchGAN classifier developed in [6]. To train this cGAN the authors combine the standard GAN objective with an L1 loss on the generator to make sure it generates sharp and realistic images. For the GAN loss, PatchGAN takes the generator output image in small tiles and averages the loss over all tiles. For the L1 loss, it is split in two differently weighted parts to give more importance to the heart region (we understand that they have binary masks on the heart region).

Evaluation and results The metrics used to evaluate the quality of the synthetic data are Fréchet inception distance (FID). The principle au FID is to compare statistics from features of real and synthetic images computed with the model Inception V3 trained on ImageNet. Authors used ablation studies for model/hyperparameter selection and chose the model with the best FID score. They also analyze their results with peak-signal-to-noise ratio and mean absolute error. As we can see the visual results seem very interesting and realistic but the quality should be evaluate by experts and radiologists.

Proposition A first thought would be to think about other elements to condition the cGAN on. For instance additionally giving binary masks that focus on the heart region as input could help the model focus on the region of interest. This would rely on more human work or to have an efficient segmentation model at hand. After following the course session on Foundation Models, we can also propose a new alternative to the cGAN, which is to use state of the art conditional diffusion models such as Latent Diffusion Models (Stable Diffusion). These models are discussed in [8] and are thought to apply diffusion in a latent space to harness faster computation time and better statistical results. A drawback would be that in medical imaging, data is scarce and diffusion models do require a lot of data for training. That said, studies such as [4], have shown that state-of-the-art pretrained diffusion models such as GLIDE have captured a lot of medical knowledge that could be used for other tasks. So retraining a model specifically for medical data could be avoided.

2 Critical assessment

Paper strengths The proposed method is not new but the paper propose a novel application of conditional GANs for generating cardiac pseudo-CT Angiography derived from thin non-contrast CT slices. Visually, the model seems to generate well contrasted CT. It could help radiologists to better read these images.

Paper weaknesses The validation procedure is one of the paper's weaknesses. Authors performed model selection with ablation studies using the FID score which relies on a neural network, trained on ImageNet (a dataset of natural images), to extract features from real and synthetic CTA images. Hence, there is no guarantee that these features are relevent for comparing medical images. Even if it seems to be a classical metric, the evalutaion method should be specifically designed for medical images.

The authors did not mention conducting cross-validation or addressing the issue of overfitting or lack of representation and bias in synthetic data. Unfortunately, authors do not compare with other existing methods, but only compare the difference between NCCT and CTA images and their pseudo-CTA with the CTA images. Authors didn't work with doctors and radiologists to assess their method which could be interesting. Finally, authors did not give their source code or a github repository link. This implies that the reproducibility of the method cannot be evaluated.

Recommendations for improvement Here are some proposals to the authors of this article.

Reproduciblity Authors share few information on the exact architecture they used (other than the main pix2pix [7] structure they consider), and since they don't share any code we cannot conduct our own experiments, and make use of their work. Providing a clean github repository is an improvement we believe they should consider for future work.

Evaluation They should split their dataset into several validation sets, this will add trust in the generalization of their results. They could also use several evaluation metrics other than only FID (we also have discussed that this may not be the most suited metric for this study 2) and PSNR (we consider MSE as a proportional metric and thus is the same). Also they should apply their solution to real life scenarios (observer study), meaning approaching doctors and radiologists to conduct experiments to see if they are able to detect pathologies more efficiently. Is it relevant to share a method that isn't useful ?

Discussion Authors don't discuss a lot their work, they don't have much self critique (what are the strengths and weaknesses of their work). They also don't propose future advancements other than saying they are working in an active field. Doing so could act on the quality of their publications, and improve how their work is received and influences the community.

References

[1] Cho Y.J. Ha J.Y. et al. Choi, J.W. Generating synthetic contrast enhancement from non-contrast chest computed tomography using a generative adversarial network. *Nature Sci Rep 11*, 2021.

[2] Shung D.L. Giuffrè, M. Harnessing the power of synthetic data in healthcare: innovation, application, and privacy. *Nature*, 2023.

[3] Yifan Jiang, Han Chen, Murray Loew, and Hanseok Ko. Covid-19 ct image synthesis with a conditional generative adversarial network, 2020.

[4] Foersch S. et al. Kather J.N., Ghaffari Laleh N. Medical domain knowledge in domain-agnostic generative ai. *npj Digit. Med. 5, 90*, 2022.

[5] Aditya Killekar, Jacek Kwiecinski, Mariusz Kruk, Cezary Kepka, Aakash Shanbhag, Damini Dey, and Piotr Slomka. Pseudo-contrast cardiac CT angiography derived from non-contrast CT using conditional generative adversarial networks. In Olivier Colliot and Ivana Išgum, editors, *Medical Imaging 2023: Image Processing*, volume 12464, page 124643F. International Society for Optics and Photonics, SPIE, 2023.

[6] Chuan Li and Michael Wand. Precomputed real-time texture synthesis with markovian generative adversarial networks. *Computer Vision–ECCV 2016: 14th European Conference*, 2016.

[7] Jun-Yan Zhu et al. Phillip Isola. Image-to-image translation with conditional adversarial networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017.

[8] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models, 2022.