State-of-the-art of CT enhancement methods for RT planning: a review

Özgür Özer[1,2,3], Alexandre Huat[2], David Gibon[2], and Juliette Thariat[1,3]

[1]GORTEC, Tours, France
[2]AQUILAB by Coexya, Lille, France
[3]Laboratoire de Physique Corpusculaire (UMR 6534), University of Caen Normandie, Caen, France

Abstract Computerized tomography (CT) scan uses X-rays and is the primary material for defining ROIs delineated by radiation oncologists for radiotherapy (RT) planning and radiotherapy quality assurance (RTQA). Additionally, doses are calculated from the CT electronic densities of tissues, but CT quality varies. CTs often exhibit blur, noise, or artefacts stemming from metals, which can be remedied by super resolution (SR) and metal artifact reduction (MAR) models, respectively. A literature search on AI models that can exploit the latest computer vision algorithms and combine both tasks is presented. An interesting approach is based on a content consistent super-resolution (CCSR) model, which utilises Stable Diffusion. The two tasks can be conducted by this model by simulating both SR and MAR distortions in the training data.

1. Introduction

Computerized tomography (CT) scan uses X-rays and stands as one of the preeminent medical imaging techniques for screening and diagnosis in many diseases including cancer, as well as radiotherapy (RT) planning and radiotherapy quality assurance (RTQA). Low quality CT scans may obscure the borders of structures considered as regions of interest (ROI, such as organs at risk (OAR) and tumour), hindering the delineation of ROIs. CTs with high resolution and high signal-to-noise ratio hold the potential to augment the precision of radiological features. Imaging quality contributes to more accurate definition of ROI and their delineation by radiation oncologists in their part of the RT planning. Super resolution (SR) reconstruction from deep learning approaches is an auspicious method of SR to increase the visual quality of images.

Metals in the area imaged by the CT scanner can degrade the quality by introducing artefacts. Linear inverse problem algorithms, such as filtered back projection, induce artefacts in CT images [1]. Deep learning based methods have achieved respectable performances for metal artefact reduction (MAR) tasks. These are based on sinogram domain enhancement, image domain enhancement, or dual domain (joint sinogram and CT image) enhancement.

SR and MAR are addressed by different models. A model that removes noise, blur or metal artefacts, or improves the visuals based on the deficiency of the CT has not been studied. Such a model may improve the quality of ROIs so that radiation oncologists can better delineate them during RT planning, and better assess delineations during RTQA processes as part of clinical research.

2. Materials and Methods

We reviewed the literature on the academic research website Scinaps [2] using the search keywords *CT super resolution* and *metal artifact reduction CT*. Only the articles that used AI on CT scans that surpassed previous models across the scores they presented were selected. Also, the latest algorithms on computer vision were investigated on the following websites: Stability AI Research Blog [3], Metal AI Research on computer vision [4], Microsoft research on computer vision [5], NVIDIA publications [6], Google Publications on machine intelligence [7], and Papers With Code on super resolution [8] and metal artefact reduction [9]. In all cases, the results earlier than 2023 were filtered. The most suitable models elected based on their potential to be adapted to the geometry and the unit of CT scans (3D and HU).

3. Results

Among AI methods, KerSRGAN [10] introduced a feature extractor in Generative Adversarial Networks (GAN) [11] in an unsupervised manner. Residual feature attentional fusion network [12] incorporated a feature extraction block to elevate visual acuity. Multi-scale Attention Mechanism [13] uses convolutional layers with various kernel sizes to capture information with different sizes. These papers used bicubic resizing to reduce the resolution rather than simulating long slice thickness. A more recent method named Linking In-Plane and Through-Plane Transformers [14] simulates longer slice thickness to reduce the quality.

For MAR, most methods published in 2023 were deep learning based. Standard vision transformer was tested for MAR [15] in image domain. A generative AI was attempted using Res-U-Net GAN [16] in the image domain [17]. Polyner [18] took an unsupervised approach using implicit neural representation [19] on image domain. Plugging Diffusion Priors in dual domains [20] used patch diffusion [21] for unsupervised training.

Outside medical AI, there have been major developments in computer vision and SR. Stable Diffusion [22] generates realistic images based on text. It incrementally refines noise from the vector representation from which the images are decoded. Its derivatives prevailed in many applications such as segmentation [23], anomaly detection [24] etc. Other models based on Stable Diffusion, such as Content Consistent Super-Resolution (CCSR) [25], Pixel-Aware Stable Diffusion (PASD) [26] outperformed other metods for SR [25]. PASD surpasses CCSR on image quality that aligns better with human perception [25]. Still, it could be unsuitable for RT. Generative AI models give a different output at every run. PASD score varied more than CCSR between runs for the same image. Frechet inception distance (FID), the score inversely related to the fidelity of generated output to real data, was high [25] for PASD i.e. it created non-existent details to have higher quality. On the other hand, CCSR made use of ControlNet [27], to confine the outputs to a narrower distribution for score consistency and lower FID score [25] compared to PASD on RealSR [28]: 105.55 vs 123.08 [25].

4. Discussion

A network that can manage MAR and SR simultaneously should be complex enough to model both but reasonably fast for training and inference. PASD was did not only fail to give consistent results and low FID score but also its very complex network may have a very costly training and inference load. In addition, for medical context it would be necessary to replace CLIP [29] with medCLIP [30], CLIP for medical context. Nevertheless, medCLIP is not mature enough for practical use because it has not been trained with enough data [30]. A modified version of CCSR for SR and MAR on CTs could overperform previous models in modelling both tasks with high fidelity to actual radiological features and consistency in quality throughout executions. For both SR and MAR, training data creation start with simulation of noise, blur, or metal artefact. By creating instances of CT scans that can have any combination of these in the training data, a model may gain the ability to conduct SR or/and MAR depending on the issue. Nonetheless, in real practice, there may be distortions that were not taken into account during training data creation.

5. Conclusion

One of the latest algorithms in computer vision, Stable Diffusion, has not been implemented in RT planning. Adapting Stable Diffusion-based CCSR model to 3D CT data and training with SR and MAR distortions may conceive a model that implements both simultaneously and consistently contingent upon CT deficiency with superior performance. A satisfactory model would improve the visual distinction between tissues and remove the metal artefacts before a radiation oncologist works defining ROIs during RT or RTQA in clinical trials.

Bibliography

[1] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. Society for Industrial, Applied Mathematics, 2001. doi: 10.1137/1.9780898719277.

[2] “Scinapse.” [Online]. Available: https://www.scinapse.io/

[3] “Research Blog.” Accessed: Feb. 02, 2024. [Online]. Available: https://stability.ai/research

[4] “Meta AI Publications.” Accessed: Feb. 02, 2024. [Online]. Available: https://ai.meta.com/results/?page=1&content_types%5B0%5D=publication

[5] “Computer Vision.” Accessed: Feb. 02, 2024. [Online]. Available: https://www.microsoft.com/en-us/research/research-area/computer-vision/

[6] “Publications | Research.” Accessed: Feb. 02, 2024. [Online]. Available: https://research.nvidia.com/publications

[7] “Publications – Google Research.” Accessed: Feb. 02, 2024. [Online]. Available: http://research.google/pubs/

[8] “Papers with Code - Super-Resolution.” Accessed: Feb. 02, 2024. [Online]. Available: https://paperswithcode.com/task/super-resolution

[9] “Papers with Code - Metal Artifact Reduction.” Accessed: Feb. 02, 2024. [Online]. Available: https://paperswithcode.com/task/metal-artifact-reduction

[10] Y. Li, L. Chen, B. Li, and H. Zhao, “4× Super-resolution of unsupervised CT images based on GAN,” *IET Image Processing*, vol. 17, no. 8, pp. 2362–2374, 2023, doi: 10.1049/ipr2.12797.

[11] R. Gao, B. Lou, Z. Xu, D. Comaniciu, and A. Kamen, “Flexible-Cm GAN: Towards Precise 3D Dose Prediction in Radiotherapy,” in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023, pp. 715–725. doi: 10.1109/CVPR52729.2023.00076.

[12] K. Yang *et al.*, “Residual Feature Attentional Fusion Network for Lightweight Chest CT Image Super-Resolution,” *Computers, Materials & Continua*, vol. 75, no. 3, pp. 5159–5176, 2023, doi: 10.32604/cmc.2023.036401.

[13] J. Yin, S.-H. Xu, Y.-B. Du, and R.-S. Jia, “Super resolution reconstruction of CT images based on multi-scale attention mechanism,” *Multimedia Tools and Applications*, vol. 82, no. 15, pp. 22651–22667, 2023, doi: 10.1007/s11042-023-14436-8.

[14] Z. Chen, C. Niu, Q. Gao, G. Wang, and H. Shan, “LIT-Former: Linking In-plane and Through-plane Transformers for Simultaneous CT Image Denoising and Deblurring,” *IEEE Transactions on Medical Imaging*, p. 1–2, 2024, doi: 10.1109/TMI.2024.3351723.

[15] S. Zheng *et al.*, “Vision Transformer with Progressive Tokenization for CT Metal Artifact Reduction,” in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1–5. doi: 10.1109/ICASSP49357.2023.10095579.

[16] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition.” 2015.

[17] Z. Zhang, M. Yang, L. Xu, J. Yang, H. Guo, and J. Wang, “An Innovative Metal Artifact Reduction Algorithm based on Res-U-Net GANs,” *Current Medical Imaging*, vol. 19, no. 13, pp. 1549–1560, 2023, doi: 10.2174/1573405619666230217102534.

[18] Q. Wu *et al.*, “Unsupervised Polychromatic Neural Representation for CT Metal Artifact Reduction.” 2023.

[19] Y. Xie *et al.*, “Neural Fields in Visual Computing and Beyond.” 2022.

[20] X. Liu, Y. Xie, S. Diao, S. Tan, and X. Liang, “Unsupervised CT Metal Artifact Reduction by Plugging Diffusion Priors in Dual Domains.” 2024.

[21] T. Luhman and E. Luhman, “Improving Diffusion Model Efficiency Through Patching.” 2022.

[22] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, “High-Resolution Image Synthesis with Latent Diffusion Models.” 2022.

[23] J. Tian, L. Aggarwal, A. Colaco, Z. Kira, and M. Gonzalez-Franco, “Diffuse, Attend, and Segment: Unsupervised Zero-Shot Segmentation using Stable Diffusion.” 2023.

[24] H. He *et al.*, “DiAD: A Diffusion-based Framework for Multi-class Anomaly Detection.” 2023.

[25] L. Sun, R. Wu, Z. Zhang, H. Yong, and L. Zhang, “Improving the Stability of Diffusion Models for Content Consistent Super-Resolution.” 2023.

[26] T. Yang, P. Ren, X. Xie, and L. Zhang, “Pixel-Aware Stable Diffusion for Realistic Image Super-resolution and Personalized Stylization.” 2023.

[27] L. Zhang, A. Rao, and M. Agrawala, “Adding Conditional Control to Text-to-Image Diffusion Models.” 2023.

[28] J. Cai, H. Zeng, H. Yong, Z. Cao, and L. Zhang, “Toward Real-World Single Image Super-Resolution: A New Benchmark and A New Model.” 2019.

[29] A. Radford *et al.*, “Learning Transferable Visual Models From Natural Language Supervision.” 2021.

[30] Z. Wang, Z. Wu, D. Agarwal, and J. Sun, “MedCLIP: Contrastive Learning from Unpaired Medical Images and Text.” 2022.