Diffusion Model for Generating Synthetic Contrast Enhanced CT from Non-Enhanced Heart Axial CT Images

Victor Rogerio Sousa Ferreira[1], Anselmo Cardoso de Paiva[1], Aristofanes Correa Silva[1], João Dallyson Sousa de Almeida[1], Geraldo Braz Junior[1] and Francesco Renna[2]
[1]*Universidade Federal do Maranhão, Av. dos Portugueses, 1966, Bacanga, São Luís/MA, Núcleo de computação Aplicada, UFMA, São Luís, Brazil*
[2]*INESC TEC, Faculdade de Ciências da Universidade do Porto, Porto, Portugal*

Keywords: Diffusion Models, Image Translation, Adversarial Networks.

Abstract: This work proposes the use of a deep learning-based adversarial diffusion model to address the translation of contrast-enhanced from non-contrast-enhanced computed tomography (CT) images of the heart. The study overcomes challenges in medical image translation by combining concepts from generative adversarial networks (GANs) and diffusion models. Results were evaluated using the Peak signal to noise ratio (PSNR) and structural index similarity (SSIM) to demonstrate the model's effectiveness in generating contrast images while preserving quality and visual similarity. Despite successes, Root Mean Square Error (RMSE) analysis indicates persistent challenges, highlighting the need for continuous improvements. The intersection of GANs and diffusion models promises future advancements, significantly contributing to clinical practice. The table compares CyTran, CycleGAN, and Pix2Pix networks with the proposed model, indicating directions for improvement.

1 INTRODUCTION

Non-communicable diseases (NCDs) are medical conditions that cannot be spread directly from one person to another and are often caused by a confluence of behavioural, physiological, environmental, and genetic variables.
In (WHO, 2023), it is stated that NCDs are the world's greatest cause of mortality, accounting for 41 million deaths per year (74% of all deaths worldwide). The bulk of NCD-related fatalities (17.9 million/year) are attributable to cardiovascular diseases, which also have a significant role in premature death and disability globally (Dondi, 2021).

A timely and efficient way to improve population health overall is through diagnostic imaging. By early discovery, they can be utilised as a preventive approach to lessen cardiovascular issues.

A common imaging modality for diagnosing cardiovascular disorders is computed tomography (CT) (Corballis, 2023; Counselor, 2023)—an imaging modality with growing diagnostic utility.

Cardiac CT plays a crucial role in diagnosing and managing heart diseases. It is possible to obtain detailed three-dimensional images of the heart through cardiac CT, allowing for precise evaluation of cardiac anatomy, function, and circulation. This makes cardiac CT a valuable tool for diagnosing various heart conditions, including coronary artery disease, cardiomyopathies, and congenital heart defects

Cardiac CT is particularly useful in detecting coronary artery disease (CAD), one of the leading causes of morbidity and mortality worldwide.

Coronary artery disease (CAD) is a significant cardiovascular disease defined by a narrowing or blockage of the coronary arteries. With cardiac CT, the existence and severity of CAD may be evaluated non-invasively. Coronary calcium deposits, linked to an elevated risk of coronary artery disease, are often evaluated by non-contrast cardiac computed tomography (NCCT). On the other hand, contrast-enhanced cardiac computed tomography (CECT) is used when the goal is to quantify cardiovascular disease, evaluate blood flow dynamics, define the composition of plaque, and offer quantitative measurements of disease severity. This tomography technique injects contrast materials into the body to increase the contrast between certain organs, blood

Ferreira, V., Cardoso de Paiva, A., Silva, A., Sousa de Almeida, J., Braz Junior, G. and Renna, F.
Diffusion Model for Generating Synthetic Contrast Enhanced CT from Non-Enhanced Heart Axial CT Images.
DOI: 10.5220/0012724600003690

In *Proceedings of the 26th International Conference on Enterprise Information Systems (ICEIS 2024) - Volume 1*, pages 857-864
ISBN: 978-989-758-692-7; ISSN: 2184-4992

arteries, or tissues and the surrounding structures on CT images.

CECT improves patient outcomes by helping physicians detect and track many elements of cardiovascular disease. It does this by making cardiovascular structures and abnormalities more visible.

But in contrast to NECT, CECT is more costly, involves more radiation exposure, and may have unfavourable side effects, including headaches and vomiting. Furthermore, anyone with allergies or renal problems might be in danger when using CECT.

There has been great potential for using artificial intelligence in cardiac CT to improve diagnosis and prognosis. This exam has unique characteristics that make it even more attractive for the application of artificial intelligence, although the complexity of this application is increasing.

Another application is in the assessment of cardiac function from cardiac CT images. Algorithms can automatically analyse cardiac volumes, ejection fraction, and wall motion, providing precise measurements that aid in evaluating heart function. This automated analysis saves time and resources, allowing physicians to focus on more complex interpretations and personalised treatment plans for each patient.

Among these challenges, we highlight the need for quantitative evaluations, which generally involve quantitative evaluations such as ventricular volume, fraction of blood volume ejected out of the ventricle, volumetric evaluation of heart muscle tissue, amount of plaque present in the coronary arteries, area of stenosis, among others. In addition, the images are acquired with thinner slices, and the evaluation targets are smaller (e.g., coronary arteries).

Given the potential risks associated with CECT, generative AI-based techniques for cardiac imaging can assist professionals in assessing coronary artery disease without the drawbacks of CECT. More precisely, we can create a CECT image that matches the given NECT image using data-driven methods without requiring contrast substance injection.

The challenge of medical image synthesis may be approached through picture-to-image translation (Parmar, 2023) or style transfer (Jing, 2020). This topic poses extra complications in the context of cardiac CT images. Since the same patient's NECT and CECT images are frequently significantly out of alignment, direct monitoring for NECT to CECT mapping is rarely feasible.

In recent years, medical image translation has emerged as a powerful solution to overcome these challenges. This process involves synthesising images of the target modality based on the guidance of images acquired from the source modality. However, the inherent nonlinear variations in tissue signals between modalities make this problem complex and ill-conditioned.

Learning-based methods, especially Generative Adversarial Networks (GANs), have shown remarkable success in image translation tasks. GANs employ an adversarial mechanism in which a discriminator guides a generator to perform a one-time mapping to produce the target image. While GANs exhibit exceptional realism in image synthesis, they indirectly characterize the target modality distribution, potentially introducing biases and limiting the mapping process's reliability.

As an alternative approach, recent studies in computer vision have explored diffusion models based on explicit likelihood characterisation and a gradual sampling process to enhance sample fidelity. However, the potential of diffusion methods in medical image translation remains largely unexplored, partly due to computational challenges and difficulties in the non-paired training of regular diffusion models.

In this work, we propose a deep learning-enabled image-to-image translation model that can map contrast-free CT images of the heart to contrast-enhanced ones. To achieve this, we implemented an adversarial diffusion model, applying concepts from GANs to generate high-quality images. This method aims to provide an accurate model compared to other approaches.

2 RELATED WORKS

In (Azarfar, 2023), authors present several papers proposing deep learning architectures to reduce or eliminate administered contrast media to acquire clinically useful computer tomographies.

The introduction of GANs (Goodfellow, 2014) presented an innovative approach to generative models. GANs operate based on the principle of rivalry between two networks - the generator and the discriminator. The generator aims to produce synthetic data indistinguishable from real data, while the discriminator strives to differentiate between the two. Through adversarial training, GANs achieve Nash equilibrium, converging the generator's distribution to the training data.

In image translation, especially in the analysis of medical images, GANs are widely used for their ability to automatically learn patterns in input data so that the model can generate new examples (output) that could exist in the original dataset. When performing image generation, the simplest model maps from source to destination through a trained generator using adversarial loss (Goodfellow, 2014). Consequently, the GAN-based translation approach has been extensively adopted in various applications.

Conditional GANs excel in mapping a single source to a destination, improving sensitivity to high-frequency details in tissue structure compared to traditional pixel-to-pixel losses. Integrating adversarial loss terms has proven effective in enhancing spatial accuracy and realism in target images synthesised with GANs, surpassing conventional convolutional models.

Other studies, specifically using GANs, whose primary focus lies in the synthesis of contrast-enhanced computed tomography (CECT) images from non-contrast CT (NCT) scans, are (Chun, 2022) and (Seo, 2021). They employ a two-stage framework and sophisticated network architectures as generators, such as DenseNet and SPADE (Park, 2019). They successfully align NCT and CECT images, surpassing previous methods in accuracy and applicability.

Other research extends to artery-contrasted computed tomography (ACT), which is crucial for diagnosing conditions like aneurysms. To mitigate the risks of contrast agents, they introduce an aorta-aware deep learning approach that synthesises artery-contrasted CT volumes directly from non-contrast CT data (Hu, 2022). Utilising aGANs and innovative loss functions, their model demonstrates remarkable accuracy in estimating ACT slices, thus enhancing diagnostic precision while minimising patient risk.

The Pix2Pix (Zhu, 2017) model is one of the approaches designed for image-to-image translation tasks. It consists of a generator to create synthetic images and a discriminator to distinguish between real and generated images. Training involves an adversarial process, where the generator tries to deceive the discriminator, and the discriminator seeks to identify fake images.

Applications of the Pix2Pix architecture (Choi, 2021) utilise the fundamental structure of the original pix2pix model to generate synthetic contrast enhanced from non-contrast chest CT, with the distinction that the 2D convolutional layers are substituted by their 3D equivalents. This model comprises a generator and discriminator networks akin to a conventional GAN. The generator network is a U-Net convolutional neural network encoder-decoder with skip connections. The discriminator network is a PatchGAN that classifies each pixel patch as real or fake, and its convolutional module is identical to the encoder block of the generator.

A dissertation (Domingues, 2022) compares the performance of two GAN models, Pix2Pix-GAN and Cycle-GAN, in generating contrast-enhanced images from non-contrast CT scans. The study explores the trade-offs of using 2D, 2.5D, and 3D inputs, employing different types of generators and datasets. Evaluation metrics include Structural Similarity Index Measure (SSIM), Peak Signal-to-Noise Ratio (PSNR), Mean-Square Error (MSE), and Dice metric for high contrast region fidelity.

CyTran (Ristea, 2021) is a GAN-based model designed for working with CT images. This innovative approach focuses on bidirectional translation of contrast and non-contrast computed tomography (CT) scans, even when the images lack direct pairing. CyTran aims to address the challenge of generating contrast scans for patients who cannot receive contrast and to enhance the alignment between contrast and non-contrast CT scans. The method employs a cycle-consistent architecture based on generative adversarial transformers designed for transferring CT scans across different contrast phases. Inspired by the CycleGAN framework, CyTran comprises two discriminators and two generators, enabling training on unpaired images through a multi-level cycle consistency loss. In addition to ensuring image-level consistency, Cytran utilises additional losses between intermediate feature representations to enhance the model's performance further. This comprehensive strategy contributes to the model's effectiveness in translating CT images bi-directionally, offering valuable applications in medical imaging scenarios.

However, GANs present their challenges. Issues such as lower reliability in mapping a single sample, premature convergence of the discriminator, and poor representational diversity leading to mode collapse can compromise the quality and diversity of generated samples. Despite these challenges, GANs currently lead in image generation tasks, surpassing other models based on metrics such as Inception Score and Accuracy.

Lately, deep diffusion models have become an alternative to GANs in generative modelling tasks in computer vision (Yang, 2022). These models are inspired by non-equilibrium thermodynamics, defining a Markov chain of diffusion steps to add random noise to the data slowly and then learning to reverse the diffusion process to construct desired data samples from the noise. Noise removal is conducted by a neural network architecture trained to maximise the correlation between adjacent pixels.

This diffusion technique provides greater reliability in network mapping and improves the quality and diversity of generated samples. The two-step structured diffusion model starts with direct diffusion, where input data is gradually perturbed over multiple steps by adding Gaussian noise. In the reverse step, the model is trained to recover the original data, reversing the diffusion process step by step. This innovative method offers a robust and effective approach to generating realistic data in various computer vision contexts.

Moreover, diffusion models are easily adaptable, able to use different architectures such as Transformers (Peebles, 2023) and adversarial networks (Wang, 2023), achieving results that surpass the quality of previous diffusion models in metrics like peak signal-to-noise ratio (PSNR), ratio is used as a quality measurement between the original and a compressed image, and structural similarity index measure (SSIM) for measuring the similarity.

Table 1: Overview of the applied techniques cited in related works.

Related Works	Applied Techniques
(Ristea, 2021)	CycleGan structure with Pix2Pix + Transformers
(Seo, 2021)	GAN (SPADE +DCGAN)
(Chun, 2022)	GAN (FC-DenseNet + PatchGAN)
(Choi, 2021)	Pix2Pix
(Domingues, 2022)	CycleGan and Pix2Pix with SkipResidual Generator
(Park, 2019)	SPADE

3 METHODOLOGIES

This work proposes a methodology for synthesising NECT to CECT images, utilising a GAN-based approach with diffusion models. This methodology consists of the following steps: data acquisition, data pre-processing and proposed network architecture.

3.1 Data Acquisition

The dataset was obtained at the Orca Score in the Grand Challenge platform (Wolterink,2022). Images in this dataset were acquired on four different CT scanners from four different vendors in four different hospitals using standard parameters for calcium scoring in cardiac CT. For each patient, both a non-contrast-enhanced CT and a contrast-enhanced computed tomography angiography (CTA) image are provided. The training set consists of images of 32 patients. The test set consists of images of 40 patients.

From this dataset, 6209 images were extracted, divided into 2812 for testing and 3397 for training and validation. The entire set consists of images with and without contrast from the same patients.

3.2 Data Pre-Processing

For this study, it was essential to conduct specific preprocessing steps before utilising these CT images to enhance the overall quality of the model. The preprocessing involved segregating the slices of contrast and non-contrast CT images of each patient's heart, selecting those in the same position with a high similarity index. This approach ensured that only the most relevant and corresponding images were used to refine the model's analysis.

To achieve this, the images were correlated using the SSIM and Normalised Cross-Correlation (NCC) to assess the structural similarity.

Images with higher similarity indices were subsequently considered equivalent. Following this, the best images from each patient, meaning those with the same position and the highest similarity indices correlating contrast and non-contrast, were separated and allocated into training, testing, and validation sets. The number of retained images was as follows.

Table 2: Orca Dataset Paired Filtration Summary.

	Contrast	Non-Contrast
Train	200	200
Test	100	100
Validation	50	50

3.3 Proposed Network Architecture

Based on the SynDiff network (Özbey, 2023), a diffusion model was developed with a conditional origin adversarial projector for fast and accurate reverse diffusion sampling. Unlike conventional models that use a relatively large number of steps, this network employs fast-forward diffusion, adaptively adjusting noise variance to balance efficiency and precision in image generation.

The proposed network utilises a Cycle-GAN architecture consisting of diffusive generators and a non-diffusive discriminator (Figure 1). The diffuse

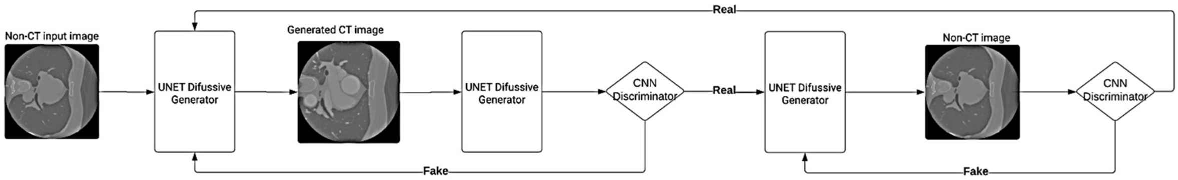

Figure 1: Architecture of proposed network.

generator translates images from the NECT to the CECT domain and vice versa. Conversely, discriminators aim to distinguish between real and generated images. In the diffusive module, generators employ a UNet backbone comprising six encoding and decoding blocks (Ho, 2020). Each block includes two residual subblocks followed by a convolutional layer. During encoding, the convolutional layer reduces the feature map resolution by half, while the channel dimensionality is doubled every other block. The convolutional layer doubles the resolution for decoding, while the channel dimensionality is halved every other block.

The proposed network utilises a Cycle-GAN architecture consisting of diffusive generators and a non-diffusive discriminator (Figure 1). The diffuse generator translates images from the NECT to the CECT domain and vice versa. Conversely, discriminators aim to distinguish between real and generated images. In the diffusive module, generators employ a UNet backbone comprising six encoding and decoding blocks (Ho, 2020). Each block includes two residual subblocks followed by a convolutional layer. During encoding, the convolutional layer reduces the feature map resolution by half, while the channel dimensionality is doubled every other block. The convolutional layer doubles the resolution for decoding, while the channel dimensionality is halved every other block.

The discriminator model is designed as a sequential neural network (Radford, 2015), tailored for input images of size 256 by 256 pixels. It consists of two convolutional layers with 64 and 128 filters, each with a kernel size of (5, 5) and a stride of (2, 2) for downsampling. Leaky ReLU activation functions introduce non-linearity after each convolutional layer. To prevent overfitting, dropout layers with a dropout rate of 0.3 are incorporated after each Leaky ReLU layer.

Given the larger input dimensions, the architecture is adapted to handle the increased spatial information. Following the convolutional layers, a flattening layer transforms the 2D feature maps into a 1D vector. Finally, a dense layer with one neuron is added, serving as the output layer for binary classification (discriminating between real and generated CT images).

During training, the proposed network enforces cycle consistency, a crucial property that ensures the translated images maintain semantic content and realism. By incorporating an additional loss function to quantify the disparity between the NECT image generated by the second generator and the original NECT image, as well as vice versa, the proposed network promotes cycle consistency. This regularisation technique guides the generator models in the creation of CECT images. The generators aim to minimise both the adversarial loss, which measures their ability to generate realistic images, and the cycle-consistency loss simultaneously. Meanwhile, discriminators are trained to improve their ability to distinguish between real and generated CT images, thereby providing feedback to the generators.

The training objective of the proposed network resembles the CycleGAN method, which utilises two main loss functions: adversarial loss and cycle-consistency loss. Adversarial loss incentivises the generators to produce images indistinguishable from real images, as perceived by the discriminators. On the other hand, cycle-consistency loss enforces the constraint that translating an image from one domain to another and then back should result in a reconstruction close to the original image

4 EXPERIMENT AND RESULTS

To evaluate the adequacy of the proposed architecture, we conducted an experiment using the Orca dataset and compared the results with other papers that employ GAN approaches to generate CECT images from NECT images..

The network hyperparameters were set as follows: 100 epochs, the Adam optimizer with beta1 = 0.5 and beta2 = 0.9, a learning rate of 10^-4 for the diffusion method and GAN, a batch size of 2, T = 1000, which represents the number of interactions in the noising and denoising process, a step size of k = 250, and T/k = 4 diffusion steps. The weight loss in diffusion and cycle models was set to $\lambda 1 \varphi = 0.5$

The metrics were obtained through the comparison between generated CECT images with real ones. The metrics presented in Table 3 are based on the averages of these results.

The results obtained, exemplified by Figure 2 and Table 3, showcase the remarkable performance of the proposed diffusion model. Notably, the competitiveness of the PSNR and SSIM indicators in generating contrast-enhanced heart images reflects the model's significant ability to preserve both quality and visual similarity.

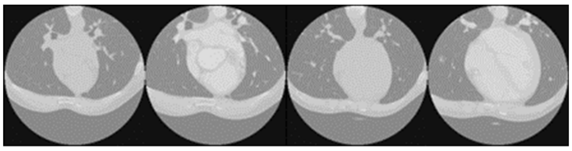

Figure 2: Images with contrast generated by the network from non-contrast images.

A noteworthy point is that although the images have a high degree of visual similarity, MAE and RMSE values are still much higher than expected. A good example is the two images below, which exhibit considerable visual resemblance but yield MAE and RMSE values as high as 0.6 and 0.7, respectively.

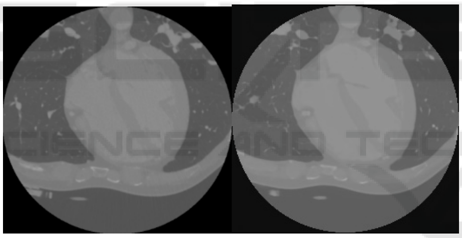

Figure 3: Real Contrast Images (Left) and the Generated One (Right).

However, in other images, the MAE and RMSE values reached 0.11 and 0.15, respectively, demonstrating that depending on the image, the network can generate a more accurate version closer to the real one.

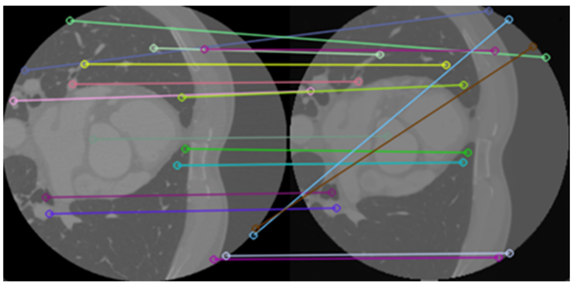

Figure 4: A Feature Matching of Real Contrast Images (Left) and the Generated One (Right).

However, upon analysing the Root Mean Square Error (RMSE) values, it is observed that, despite the visual resemblance of the generated images, the model predictions deviate significantly from the actual values, as seen in Table 3.

Upon closer examination of the images, a subtle yet discernible variance in the absolute pixel values between the original and the generated samples becomes apparent. Indeed, a slight disparity exists between the generated pixel values and the corresponding ideal pixel values, as shown in Figure 2, with the generated pixel values exhibiting a marginally higher magnitude. While these minor discrepancies may seem inconspicuous individually, their cumulative effect in the summation process significantly contributes to the observed dissimilarity reflected in the RMSE.

It is important to note that the interpretation of RMSE depends on the specific domain of the problem and the units of the variable being predicted. In some cases, a high RMSE may be acceptable if it aligns with the natural variations in the data or is justified by the nature of the problem being addressed. These indicators suggest the presence of substantial variations that require a deeper understanding of the generated images. Studying these discrepancies can provide valuable insights further to enhance the effectiveness of the image generation process.

Table 3: Comparison of results among CyTran, Cycle-GAN, Pix2Pix-GAN networks, Cycle-GAN-2D, and Cycle-GAN-2D with SkipResidual Generator against the proposed model.

Model	MAE	RMSE	SSIM	PSRN
CyTran	0.061	0.144	0.745	29.66
Cycle-GAN	0.066	0.150	0.724	29.22
Pix2Pix-GAN	0.070	0.165	0.729	29.51
Cycle-GAN-2D	0,030	-	0,433	15,569
Pix2Pix-GAN-2D	0,025	-	0,492	16,375
Proposed Diff-Model	0.061	0.200	0.701	32,85

5 CONCLUSIONS

This paper addresses the translation of contrast and non-contrast cardiac computed tomography (CT) images using a deep learning-based adversarial diffusion model. By overcoming challenges associated with medical image translation, we explore an approach that combines concepts from generative adversarial networks (GANs) and diffusion models. The obtained results, evaluated through metrics such as PSRN and SSIM, showcase the remarkable capability of the model in generating contrast-enhanced cardiac images while preserving quality and visual similarity. However, the analysis of RMSE indicates persistent challenges, suggesting the presence of variations that require a deeper understanding to enhance the consistency and fidelity of the generated images.

In conclusion, the developed model delivers notable results, but the study acknowledges the need for continuous improvements to address variations in the generated images. The intersection of GANs and diffusion models proves promising, pointing towards future research and developments in medical image translation and significantly contributing to advancing this crucial area in clinical practice.

ACKNOWLEDGEMENTS

The authors acknowledge the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil - Finance Code 001, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil, and Fundação de Amparo à Pesquisa Desenvolvimento Científico e Tecnológico do Maranhão (FAPEMA) (Brazil), Empresa Brasileira de Serviços Hospitalares (Ebserh) Brazil (Grant number 409593/2021-4), and the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia, within project UIDB/50014/2020.DOI.10.54499/UIDB/50014/2020 | https://doi.org/10.54499/uidb/50014/2020 for the financial support.

REFERENCES

Azarfar, G., Ko, SB., Adams, S.J. et al. (2023) Applications of deep learning to reduce the need for iodinated contrast media for CT imaging: a systematic review. Int J CARS 18, 1903–1914. https://doi.org/10.1007/s11548-023-02862-w

Choi, J.W., Cho, Y.J., Ha, J.Y. et al. (2021) Generating synthetic contrast enhancement from non-contrast chest computed tomography using a generative adversarial network. Sci Rep 11, 20403. https://doi.org/10.1038/s41598-021-00058-3

Chun, J., Chang, J. S., Oh, C., Park, I., Choi, M. S., Hong, C. S., ... & Kim, J. S. (2022). Synthetic contrast-enhanced computed tomography generation using a deep convolutional neural network for cardiac substructure delineation in breast cancer radiation therapy: a feasibility study. Radiation Oncology, 17(1), 1-9.

Corballis, N.; Tsampasian, V.; Merinopoulis, I.; Gunawardena, T.; Bhalraam, U.; Eccleshall, S.; Dweck, M.R.; Vassiliou, V. CT.(2023) angiography compared to invasive angiography for stable coronary disease as predictors of major adverse cardiovascular events—A systematic review and meta-analysis. Heart Lung, 57, 207–213.

Counseller, Q. and Aboelkassem, Y. (2023) Recent technologies in cardiac imaging, Frontiers in Medical Technology, Vol. 4, DOI 10.3389/fmedt.2022.984492,

Croitoru, F. A., Hondru, V., Ionescu, R. T., & Shah, M. (2023). Diffusion Models in Vision: A Survey, in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 9, pp. 10850-10869.

Dondi M, Paez D, Raggi P, Shaw LJ, Vannan M.(2021).Integrated non-invasive cardiovascular imaging: a guide for the practitioner. *International Atomic Energy Agency*.

Domingues, R. A. D. (2022). Automatic contrast generation from contrastless CTs, Master Thesis, Universidade do Porto, FCUP - Faculdade de Ciências.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139-144.

H., & Kwak, S. (2021). Neural contrast enhancement of CT image. *In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 3973-3982).

Henry, J., Natalie, T., & Madsen, D. (2021). Pix2Pix GAN for Image-to-Image Translation. Research *Gate Publication*, 1-5.

Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in neural information processing systems. Editors: H. Larochelle and M. Ranzato and R. Hadsell and M.F. Balcan and H. Lin}, Vol. 33, pages 6840--685}, Curran Associates, Inc.}, https://proceedings.neurips.cc/paper_files/paper/2020/file/4c5bcfec8584af0d967f1ab10179ca4b-Paper.pdf*

Hu, T., Oda, M., Hayashi, Y., Lu, Z., Kumamaru, K. K., Akashi, T., ... & Mori, K. (2022). Aorta-aware GAN for non-contrast to artery contrasted CT translation and its application to abdominal aortic aneurysm detection. *International Journal of Computer Assisted Radiology and Surgery*, 1-9. *https://doi.org/10.1007/s11548-021-02492-0*

Jing, Y., Yang, Y., Feng, Z., Ye, J., Yu, Y., & Song, M. (2019). Neural style transfer: A review. *IEEE transactions on visualization and computer graphic*s, 26(11), 3365-3385.

Karras, T., Aittala, M., Aila, T., & Laine, S. (2022). Elucidating the design space of diffusion-based generative models. Advances in Neural Information Processing Systems, 35, 26565-26577.

Park, T., Liu, M. Y., Wang, T. C., & Zhu, J. Y. (2019). Semantic image synthesis with spatially-adaptive normalization. *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 2337-2346).

Parmar, G., et all. (2023). Zero-shot Image-to-Image Translation. In ACM SIGGRAPH 2023 Conference Proceedings (SIGGRAPH '23). Association for Computing Machinery, New York, NY, USA, Article 11, 1–11. https://doi.org/10.1145/3588432.3591513

Peebles, W. and Xie, S. (2023) Scalable Diffusion Models with Transformers," IEEE/CVF International Conference on Computer Vision (ICCV), Paris, France, 2023, pp. 4172-4182, doi: 10.1109/ICCV51070.2023.00387

Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conferenc*e, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18 (pp. 234-241). Springer International Publishing.

Ristea, N. C., Miron, A. I., Savencu, O., Georgescu, M. I., Verga, N., Khan, F. S., & Ionescu, R. T. (2021). Cytran: Cycle-consistent transformers for non-contrast to contrast ct translation. arXiv preprint arXiv:2110.06400.

Seo, M. et al., "Neural Contrast Enhancement of CT Image", 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 2021, pp. 3972-3981, doi: 10.1109/WACV48630.2021.00402.

Özbey, M., Dalmaz, O., Dar, S. U., Bedel, H. A., Özturk, Ş., Güngör, A., & Çukur, T. (2023). Unsupervised medical image translation with adversarial diffusion models. *IEEE Transactions on Medical Imaging.*

Wang, Z., Zheng, H., He, P., Chen, W., & Zhou, M. (2022). Diffusion-gan: Training gans with diffusion. arXiv preprint arXiv:2206.02262.

Wolterink, J., Vo, B.D., Leine, T. , Viergever, M. A., Išgum, I.(2022) Orca score. https://orcascore.grand-challenge.org

WHO World Health Organization. Noncommunicable diseases. (2023) Available online: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases.

Zhu, J. Y., Isola, P .Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks*. In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1125-1134).