Computed Tomography Slice Reconstruction using Deep Learning

Margarida Pereira[1]
up202204379@edu.fe.up.pt
Rúben Silva[1]
up202103374@edu.fe.up.pt
Fábio Nunes[3,4]
fsn@med.up.pt
Jennifer Mâncio[3]
jenni@med.up.pt
Ricardo Fontes-Carvalho[3,4]
ricardo@med.up.pt
João Pedrosa[1,2]
joao.m.pedrosa@inesctec.pt

[1] Faculty of Engineering of the University of Porto (FEUP), Portugal
[2] Institute for Systems and Computer Engineering, Technology and Science (INESC TEC), Porto, Portugal
[3] Faculty of Medicine of the University of Porto (FMUP), Portugal
[4] Centro Hospitalar de Vila Nova de Gaia e Espinho, Portugal

Abstract

Coronary Artery Disease (CAD) is a prevalent and life-threatening condition, necessitating accurate diagnosis due to its substantial impact on morbidity and mortality. This study explores the potential of an autoencoder architecture for cardiac CT image reconstruction, a critical aspect of CAD diagnosis and treatment planning.
A dataset of 20 patients' CT scans was utilized, partitioned into training, validation, and test sets. The model, combining U-Net and autoencoder principles, demonstrated effective learning during training. However, some challenges in generalization were observed, with higher errors in unseen data, possibly due to overfitting. Future work should focus on enhancing the model's adaptability to new data. This research lays the groundwork for advancing cardiac CT image reconstruction, promising improved cardiac disease diagnosis.

Introduction

Coronary Artery Disease (CAD) is a prevalent global health concern, marked by narrowed arteries and restricted blood flow to the heart. Its diagnosis is pivotal due to its impact on morbidity and mortality [1]. Combining initial CT scans with high-resolution computed tomography coronary angiography (CTCA) exams has emerged as a promising approach [2]. This dual-exam strategy provides comprehensive insights into coronary health, improving accuracy and guiding interventions.
Effective CAD diagnosis hinges on accurate CT image reconstruction. Recent advancements in deep learning, particularly Convolutional Neural Networks (CNNs) and autoencoder architectures like U-Net and CycleGAN, offer potential to enhance image quality [3-6]. In this research, the focus is on testing the effectiveness of an autoencoder architecture specifically in the reconstruction of cardiac CT images, a critical component of CAD diagnosis and treatment planning.

Methods

Dataset

This study utilized a dataset comprising CT images of patients who underwent CAD screening. The original DICOM data was provided by the Centro Hospitalar de Vila Nova de Gaia e Espinho, which was then pre-processed and converted into PNG images. The dataset includes 20 CT scans from 20 different patients, with each scan being composed of between 40 to 50 images.
The dataset was partitioned into three distinct subsets: a 60% training set, a 20% validation set, and a remaining 20% for testing. This division facilitated comprehensive model training, rigorous validation, and the ultimate assessment of model performance.

Image Reconstruction Model

The proposed model architecture is a U-Net framework, aiming to produce an output slice that corresponds to the reconstruction of the input slice. The model leverages an encoder-decoder design (Figure 1), wherein the encoder captures low-level image features, and the decoder performs detailed image reconstruction [7]. Notably, skip connections facilitate the direct transfer of low-level information, enhancing reconstruction accuracy.

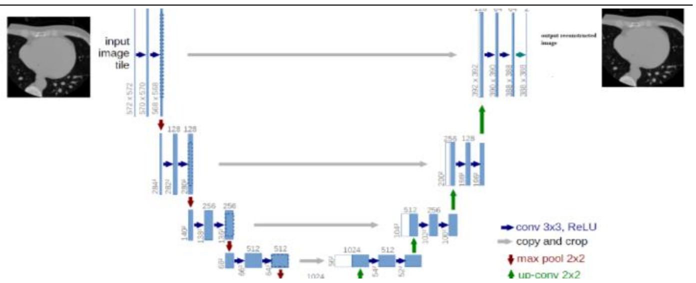

Figure 1- U-Net architecture for CT image reconstruction (adapted from [7]).

Incorporating autoencoder principles, the model's encoder employs convolution and max-pooling layers to detect image patterns and create a compact latent representation [6]. This latent representation is then processed by the decoder, employing transposed convolution layers and skip connections inspired by U-Net. Additional convolution blocks, optionally with Batch Normalization, enhance network performance.
The decoder culminates in a convolution layer with linear activation, generating the reconstructed output as a continuous representation of the input image. This model, with 124,361,025 trainable parameters, is trained for image reconstruction.
The input images, with dimensions of 256x256 and a batch size of 12, underwent a single preprocessing step involving normalization, which rescaled pixel values from the range of 0 to 255 to 0 to 1, relative to the Hounsfield Unit (HU) values in the original DICOM images. This normalization is a common step in data preprocessing for Deep Learning tasks as it facilitates algorithm convergence.
The training process incorporated the Mean Squared Error (MSE) loss function and the adaptive moment estimation (ADAM) optimizer. MSE quantifies the squared discrepancy between model predictions and ground truth, averaged across the entire dataset.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \widehat{y_i})^2 \qquad (1)$$

where y_i is the original image, $\widehat{y_i}$ is the model prediction, the reconstructed output image, and N the number of samples.
Training spanned 105 epochs with a learning rate of 0.0001. The number of steps per epoch was determined by dividing the training image count by the batch size. It was implemented an early stopping method, which monitored validation loss and halted training if it ceased to improve over 20 consecutive epochs. An adaptive learning rate schedule was also employed, progressively reducing the learning rate throughout training to enhance convergence and model performance. Specifically, the learning rate was set at 0.0001 for epochs under 20, reduced to 0.00001 between epochs 15 and 40, and held at 0.000001 for the remaining epochs.
Ultimately, the model with the best performance during training was automatically saved, allowing the model to cease training at the optimal point, thereby conserving time and computational resources.

Results and Discussion

Concerning the training process itself, the loss curves shown in Figure 2 were obtained for the training and validation sets. The training loss starts at a high value, decreasing rapidly as the model adjusts to the data, while the validation loss begins at lower values and decreases, stabilizing at a slightly higher level. This indicates that the model is effectively learning the patterns in the training dataset and generalizing well to different data.

The convergence of the loss curves suggests a high level of optimization achieved by the model.

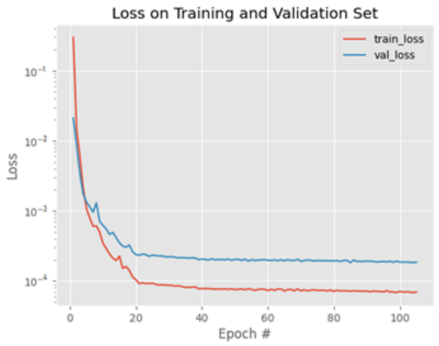

Figure 2- Loss curve on training and validation datasets, on a logarithmic y-axis scale.

During the testing phase, the model was evaluated on unseen data, and MSE results were calculated for the training, validation, and test sets. These MSE values are detailed in Table 1.

Table 1- Results of the mean values of MSE, with the highest standard deviations for each subset.

Metric / Dataset	MSE
Training Set	$0{,}0004 \pm 0{,}229$
Validation Set	$0{,}0007 \pm 0{,}241$
Test Set	$0{,}0098 \pm 0{,}463$

The MSE values for the training set remained consistent and relatively low, indicative of accurate image reconstruction. Similarly, the validation set exhibited slightly higher but still comparable MSE values, signifying a robust generalization capability. Conversely, the test set displayed higher MSE values, suggesting some limitations in the model's capacity to generalize to entirely new and unseen data. This disparity may be attributed to potential overfitting, where the model excels in memorizing training data but struggles to extend its performance to diverse, unobserved samples.

During the testing phase, the Structural Similarity Index (SSIM) was employed to assess the quality of the model's reconstructions [8]. The SSIM metric is valuable in evaluating the performance of image reconstruction models because it considers not only pixel-level differences but also the perceptual quality of the reconstructed image, considering factors such as brightness, contrast, and structural similarity between images, providing a more comprehensive measure of the perceptual quality of the reconstructions. The average SSIM value obtained was high ($0{,}968 \pm 0{,}002$), indicating a high structural similarity between the reconstructed and original images.

Figures 3 and 4 present some visual examples of the model's predictions, contrasting input test images with their reconstructed counterparts.

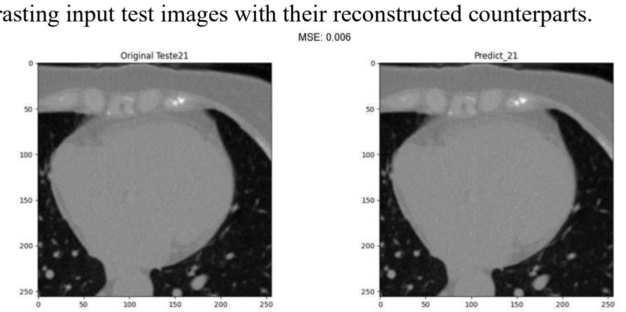

Figure 3-Visual example of the reconstructed image with a good model prediction result, compared to the input image of the test set.

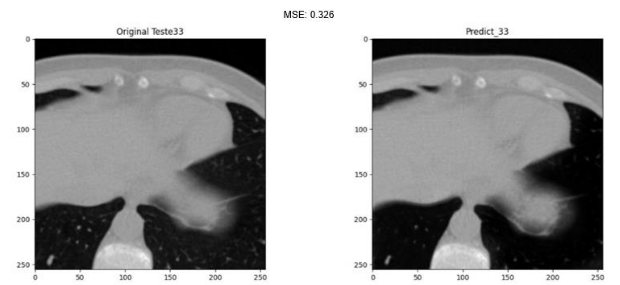

Figure 4- Visual example of the reconstructed image with a poor model prediction result, compared to the input image of the test set.

Figure 3 demonstrates a successful reconstruction, where the model accurately identifies and reproduces the primary features of the original image. The obtained image closely resembles the input test image, serving as a compelling instance of the model's ability to capture the fundamental characteristics of cardiac tissue.

Conversely, Figure 4 shows the opposite scenario, with deficiencies in the model's reconstruction. Although the contours and overall aspects of the image are preserved, there are discrepancies and imperfections throughout, such as missing details, in this case, vessels and bronchi. This highlights the model's limitation in fully capturing the intricacies and finer details of cardiac tissue in such instances. It's noteworthy that in this example, the original image features a transition region between different tissues with a blurred appearance, posing a challenge for the precise extraction of essential characteristics.

Conclusions

In summary, this study successfully applied an autoencoder architecture to cardiac CT image reconstruction, achieving promising results in preserving essential image features with low losses and high structural similarity, critical for CAD diagnosis. However, the study revealed certain limitations, particularly the model's challenges in generalizing to new data, leading to higher errors in such cases. Future efforts should concentrate on enhancing the model's adaptability to novel data through robustness improvements and data augmentation techniques. This research serves as a foundational step in advancing cardiac CT image reconstruction techniques, offering the potential to significantly enhance the diagnosis and treatment planning for individuals with cardiac conditions.

Acknowledgments

This work is financed by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia, **within projects UIDB/50014/2020 (uxCADCT) and LA/P/0063/2020**.

João Pedrosa is supported by Fundação para a Ciência e Tecnologia (FCT) through application reference 2022.06138.CEECIND.

References

[1] R. S. Udaya Ralapanawa, "Epidemiology and the Magnitude of Coronary Artery Disease and Acute Coronary Syndrome: A Narrative Review," em *Journal of Epidemiology and Global Health*, Sri Lanka, Atlantis Press, June 2021, pp. Volume 11, Issue 2, pp. 169 - 177.

[2] Bauersachs, R., Zeymer, U., Brière, J. B., Marre, C., Bowrin, K., & Huelsebeck, M. (2019). Burden of Coronary Artery Disease and Peripheral Artery Disease: A Literature Review. Cardiovascular therapeutics, 2019, 8295054. https://doi.org/10.1155/2019/8295054

[3] Arndt, C., Güttler, F., Heinrich, A., Bürckenmeyer, F., Diamantis, I., & Teichgräber, U. (2021). Deep Learning CT Image Reconstruction in Clinical Practice. CT-Bildrekonstruktion mit Deep Learning in der klinischen Praxis. RoFo: Fortschritte auf dem Gebiete der Rontgenstrahlen und der Nuklearmedizin, 193(3), 252–261.

[4] Wu, D., Kim, K., & Li, Q. (2019). Computationally efficient deep neural network for computed tomography image reconstruction. Medical physics, 46(11), 4763–4776. https://doi.org/10.1002/mp.13627

[5] Choi, K., Kim, S.W., & Lim, J.S. (2018). Real-time image reconstruction for low-dose CT using deep convolutional generative adversarial networks (GANs). Medical Imaging.

[6] A. Majumdar, "Magnetic resonance imaging," An autoencoder based formulation for compressed sensing reconstruction., pp. 62-68, 2018.

[7] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer International Publishing, 2015.

[8] Wang, Zhou; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. (2004-04-01). "Image quality assessment: from error visibility to structural similarity". IEEE Transactions on Image Processing. 13 (4): 600–612.