EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
2025, Vol. 18, Issue 4, Article Number 6300
ISSN 1307-5543 – ejpam.com
Published by New York Business Global

Coronary Artery Segmentation in CTA Images: Evaluating Automated Segmentation of Coronary Arteries Using U-Net Variants and Vesselness Enhancement

Omar Alirr

College of Engineering and Technology, American University of the Middle East, Kuwait

Abstract. Accurate segmentation of the coronary arteries from computed tomography (CTA) images is crucial to the diagnosis and treatment of cardiovascular diseases, especially with the increasing prevalence of coronary artery disease (CAD), a leading cause of death worldwide. This paper introduces a comprehensive pipeline that combines vessel enhancement, heart region of interest (ROI) extraction, and advanced deep learning techniques, specifically the ResUNet model, to efficiently and accurately extract vessels from the coronary artery. A comparative study of several U-Net-based networks, including U-Net, ResU-Net, Attention U-Net, U-Net++, and TransU-Net, was performed using a five-fold cross-validation approach. Performance metrics such as the Dice Similarity Coefficient (DSC), Jaccard Index (JI), Recall, and Precision were used to assess the models. The results showed that the proposed method achieved a DSC of 0.867, recall of 0.881, and precision of 0.892, outperforming other state-of-the-art methods. Specifically, the ResU-Net model achieved a DSC of 0.865 and a JI of 0.764, demonstrating superior segmentation accuracy. These results highlight the potential of automated segmentation techniques to reduce cardiologists' workload, minimize human error, and improve clinical decision making in CAD management. The proposed method not only improves the accuracy and efficiency of coronary artery segmentation, but also plays a crucial role in the timely diagnosis and effective treatment of cardiovascular diseases.

2020 Mathematics Subject Classifications: 92C55, 68U10, 65D18

Key Words and Phrases: Coronary arteries, computed tomography angiography (CTA), u-net, segmentation, resu-net

1. Introduction

Coronary artery disease is a leading cause of mortality worldwide, and accurate diagnosis and monitoring of coronary artery plaques is crucial for effective treatment. Computed Tomography Angiography (CTA) has become a widely used noninvasive imaging modality

DOI: https://doi.org/10.29020/nybg.ejpam.v18i4.6300

Email address: `omar.alirr@aum.edu.kw` (O. Alirr)

to visualize the coronary arteries and detect plaque build-up. However, manually segmenting coronary arteries and plaques from CTA images is a time-consuming and error-prone task, making it impractical for large-scale clinical applications [1–4].

To address this challenge, researchers have explored the use of deep learning methods for automated coronary artery and plaque segmentation. One of the most widely used deep learning architectures in this field is the U-Net, a convolution neural network designed for biomedical image segmentation. U-Net and its variants have shown promising results due to their ability to capture local and global characteristics of vascular structures. For example, a study by Zhang et al. proposed a two-step approach using U-Net for coronary artery extraction followed by a fully convolution network for plaque segmentation, achieving a dice coefficient of 0.8954 for artery segmentation and an AUC score of 0.9202 for plaque segmentation [5–7]. The Transformer architecture, which employs self-attention to identify long-range dependencies in data, was first presented by Vaswani et al. Lei et al. used this idea for medical image segmentation and used attention-based deep learning networks to create an automated technique for coronary artery segmentation [7–9]. Additionally, Song et al. proposed an efficient 3D-UNet model for automatic coronary artery segmentation in CTA images, incorporating feature fusion and rectification modules. This method achieved a Dice similarity coefficient (DSC) of 0.8795, demonstrating its effectiveness in enhancing segmentation accuracy through the use of a 3D-UNet architecture with these advanced features [10]. Authors in [11], examined the segmentation of the coronary artery lumen using 3D U-Net convolutional neural networks and tested its usefulness on several datasets, demonstrating that the suggested deep learning method performed better than alternative approaches assessed by Dice coefficients.

In addition to the standard U-Net architecture, researchers have explored several variations and extensions to improve the performance of deep learning methods for coronary artery segmentation from CTA images. The aim of this paper is to show and compare the segmentation results from these different deep learning methods, including U-Net++, TransUNet, Attention U-Net, ResUNet, and Attention U-Net [7, 12–14]. U-Net++ is a modification of the original U-Net that aims to bridge the semantic gap between the contracting path and the expansive path by introducing nested and dense skip connections, leading to better segmentation accuracy. TransUNet combines the strengths of convolutional neural networks and transformer architectures to capture both local and global features, demonstrating promising results in various medical image segmentation tasks, including coronary artery segmentation from CTA. Attention U-Net introduces attention gates within the standard U-Net architecture to selectively focus on relevant features during the encoding and decoding stages, helping the model concentrate on informative regions of the input image and resulting in more accurate segmentation of the coronary arteries [15–19].

Other deep learning architectures, such as ResUNet and Attention U-Net, have also been investigated for coronary artery segmentation. ResUNet incorporates residual connections to address the vanishing gradient problem and improve the flow of information through the network, potentially enhancing the segmentation performance. Attention U-Net+ combines the attention mechanism with the U-Net++ architecture, further refining the

segmentation by selectively emphasizing relevant features at multiple scales [6, 7, 12].
This study's main contributions include applying UNet, U-Net++, TransUNet, Attention U-Net and ResUNet for automatic coronary arteries segmentation from CTA, comparing model performance for segmentation. Comparing models performance on different coronary artery datasets can help cardiologists select the best model and make informed decisions. The study evaluated models performance using various metrics, including sensitivity, precision,and Dice Score. By comparing the segmentation results from these different deep learning methods, this paper aims to provide insights into the strengths and limitations of each approach, ultimately contributing to the development of more accurate and efficient tools for clinical diagnosis and monitoring of coronary artery disease. The study will evaluate the performance of these deep learning architectures on divers CTA dataset [9, 20, 21].
The findings of this study will have important implications for the field of coronary artery disease diagnosis and management. By identifying the most effective deep learning-based segmentation techniques, clinicians and researchers can leverage these tools to improve the accuracy and efficiency of coronary artery and plaque detection, leading to earlier intervention and better patient outcomes. Furthermore, the insights gained from this comparative analysis can guide the development of future deep learning architectures and segmentation algorithms, driving continuous advancements in this critical area of medical imaging and cardiovascular research.

2. Methods and Materials

The proposed study introduces a comprehensive approach to address the segmentation of coronary arteries from computed tomography angiography (CTA) images. The methodology employs a structured pipeline consisting of two main stages; image preprocessing and deep learning segmentation. That employ several key steps forming the proposed framework as shown by the flowchart in figure 1 below.
This method uses advanced approaches, such as deep learning models, vesselness filtering, and heart area of interest (ROI) extraction, to efficiently and precisely extract coronary artery vessels, which is essential for the diagnosis and treatment of cardiovascular disorders. Using filters like the Frangi filter, which enhance vessel-like features while lowering background noise, vesselness filtering is essential for improving the appearance of coronary arteries in CTA images. The cardiac region is then identified via heart ROI extraction, which improves the accuracy of segmentation algorithms that target coronary arteries.
An evaluation of the five deep learning models undergoes, U-Net, ResU-Net, Attention U-Net, U-Net++, and TransU-Net, for the segmentation of coronary arteries from the extracted ROI from the CTA images. The performance of these segmentation models is rigorously assessed using a 5-fold cross-validation strategy, ensuring that the results are robust and generalizable across different subsets of the dataset.
For the segmentation task, the input images go through a pre-processing step to standardize their size and enhance the visibility of the arterial structures. The models are trained on images of size 512 x 512 pixels, which allows them to effectively utilize spatial infor-

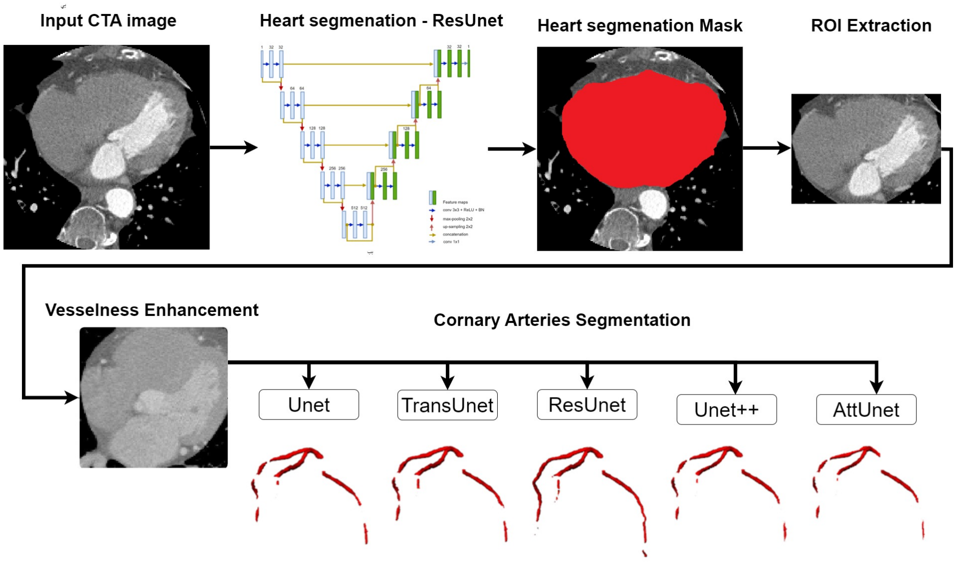

Figure 1: Methodology Flowchart.

mation while maintaining a balance between computational efficiency and segmentation accuracy. By systematically comparing the performance of these segmentation models, the study aims to identify the most effective approaches for automated coronary artery segmentation. This focus on deep learning techniques has the potential to significantly reduce the workload of radiologists and improve the efficiency of diagnosis and treatment planning for coronary artery disease, ultimately contributing to better patient outcomes.

2.1. Clinical Datasets

Two well-known resources in the field of medical image analysis, namely for the segmentation of cardiac structures and coronary arteries, are the MM-WHS and CCTA-ASOCA databases.

The CCTA-ASOCA dataset is a publicly available dataset provided by the ASOCA (Automatic Segmentation of Coronary Arteries) challenge. It consists of 40 Coronary Computed Tomography Angiography (CCTA) images, with a consistent CT slice resolution of 512 x 512 pixels with number of slices per volume ranges from 168 to 224 and a slice thickness of 0.625 mm. The dataset is divided into 20 healthy cases and 20 cases with confirmed coronary artery disease. It includes annotations produced by three expert annotators for the training set, ensuring the reliability and consistency of the ground truth. This multi-expert annotation enables fair comparison and reproducibility of results across different research groups. The CCTA-ASOCA dataset is notable for its large scale, with 40 CCTA images, allowing for more comprehensive training and evaluation of deep learning models that typically require substantial amounts of data to achieve optimal performance.

The MM-WHS (Multi-Modality Whole Heart Segmentation) 2017 dataset provides a total of 60 cardiac CT images, comprising 20 labeled and 40 unlabeled CT volumes. The dataset includes CT images with a slice thickness ranging from 0.45 mm to 0.625 mm and a slice resolution of 512 x 512 pixels. Additionally, the number of slices per volume ranges from 177 to 363, reflecting variability in scan coverage. The inclusion of both labeled and unlabeled data enhances the dataset's utility by enabling the exploration of semi-supervised and unsupervised learning approaches, which are particularly beneficial when dealing with limited labeled data. The dataset's structure reflects real-world clinical scenarios where a combination of labeled and unlabeled data can be used to improve model performance. By focusing on CT imaging, the dataset allows researchers to develop and refine techniques specifically for analyzing and interpreting cardiac CT data, which is crucial for accurate assessment and diagnosis of cardiovascular conditions. Both datasets are used in structured challenges to foster research and development in their respective domains. The ASOCA challenge for coronary artery segmentation provides the CCTA-ASOCA dataset. The Multi-Modality Whole Heart Segmentation Challenge 2017 uses the MM-WHS 2017 dataset. These challenges allow diverse research groups to test and validate their methods in an open and equitable manner, allowing results from different studies to be compared and reproduced. The CCTA-ASOCA and MM-WHS datasets offer unique specifications tailored to the segmentation of coronary arteries and cardiac structures, respectively. The CCTA-ASOCA dataset provides a large-scale, annotated resource for coronary artery segmentation, while the MM-WHS 2017 dataset incorporates multi-modality cardiac images and labeled/unlabeled data for comprehensive whole heart segmentation research. Both datasets support structured challenges that drive progress in their respective fields.

2.2. Image Pre-processing

Preprocessing is a crucial step in analyzing coronary computed tomography angiography (CTA) images, as it enhances image quality and standardizes the data for accurate deep learning-based segmentation of coronary arteries. This process includes image resampling, noise reduction, contrast enhancement, and normalization. Image resampling improves segmentation accuracy by ensuring that the images are consistently aligned and spaced. The CTA input image was resampled to a uniform resolution of 1mm for all images. This interpolation generates new voxels with consistent 1mm spacing, facilitating segmentation and ensuring that the segmented arteries are properly aligned and spaced. Another crucial step in prepping CTA images for coronary artery segmentation is normalization. Normalization attempts to standardize the picture data by bringing the intensity values to a similar scale, as CTA scans may have different intensity ranges. We applied Min-max normalization, a linear method that modifies the CTA image's intensity values according to its minimum and maximum values. This technique lessens the impact of picture capture and reconstruction artifacts while successfully standardizing the image data.

2.3. Region of Interest Extraction

The segmentation of coronary arteries is significantly impacted by the extraction of the region of interest (ROI) by heart segmentation from CTA images. Accurately identifying and separating the heart region from surrounding tissues is a crucial stage in the vital ROI extraction process, which is necessary to separate the coronary arteries inside the heart from other tissues and structures. Heart segmentation enhances the efficiency of the segmentation process by narrowing the focus to the heart region. This allows deep learning models to process less data, resulting in faster segmentation times and reduced use of computational resources. This improvement is especially valuable in medical imaging, where datasets are often large, and images are typically high-resolution [22]. Furthermore, ROI segmentation can enhance the accuracy of the segmentation process by addressing the imbalance between foreground and background. Concentrating on the heart region reduces the prevalence of background pixels, allowing deep learning models to achieve more accurate segmentation [23].

This study utilizes a ResUnet-based model to accurately capture the complex structures and patterns in medical images, enabling precise segmentation of the heart region from CTA images. By achieving accurate heart segmentation, subsequent models can focus on segmenting the coronary arteries within the heart, improving both the accuracy and efficiency of the segmentation process.

2.4. Coronary Arteries Appearance Enhancement

Enhancing the visibility of coronary arteries in coronary CT angiography (CTA) images is essential for accurate diagnosis and effective treatment in cardiovascular care. Increased intensity of coronary arteries compared to surrounding tissues helps deep learning systems more effectively recognize features. The Frangi vesselness filter is a sophisticated technique designed to address the low contrast-to-noise ratio common in coronary CTA datasets. It leverages the Hessian matrix to evaluate local curvature in the image, applying 3D Gaussian filters at multiple scales to detect vessels of varying sizes. By analyzing the eigenvalues of the Hessian matrix, the filter assigns a vesselness score to each voxel, reflecting its likelihood of being part of a tubular structure. Fine-tuning the Frangi filter's parameters, such as the Gaussian filter scale and sensitivity of the discrimination function, is key to optimizing coronary artery visibility. Adjustments enhance artery contrast while suppressing noise, with smaller scales emphasizing fine vessels and larger scales highlighting broader ones, ensuring detailed visualization. This improved clarity significantly impacts clinical evaluations, enabling precise assessment of coronary artery stenosis, detection of abnormalities, and planning for interventional procedures. Enhanced visualization also supports coronary artery disease (CAD) research and aids in developing innovative diagnostic and therapeutic strategies.

2.5. Segmentation Methods

In medical image segmentation, various advanced neural network architectures have been designed to improve accuracy and performance. U-Net, developed by Ronneberger et al., is a pioneering model featuring an encoder-decoder architecture with skip connections that retain spatial information. Numerous U-Net-based approaches have since been proposed, extending the original U-Net's capabilities to tackle specific challenges in medical image segmentation.

2.5.1. U-Net

U-Net is introduced by Ronneberger et al. in 2015, is a pioneering architecture specifically designed for biomedical image segmentation. The architecture is characterized by its symmetric encoder-decoder structure and the use of skip connections, which contribute to its effectiveness in tasks requiring high spatial precision. The encoder section of U-Net consists of a series of convolutional layers followed by max pooling operations. Each convolutional block typically includes two 3x3 convolutions followed by a ReLU activation function, which extracts and refines features from the input image. Pooling operations reduce the spatial dimensions of the feature maps while increasing the depth of the feature channels. This process allows the network to capture hierarchical features at different scales. The decoder section nm provides intermediate outputs at various levels during the training process. These intermediate outputs are used to guide the learning, helping the model to acquire more robust features and improving the final segmentation accuracy. Deep supervision stabilizes the training by providing additional gradient signals, encouraging the network to learn meaningful features at each level of the hierarchy. These enhancements to the base U-Net architecture make U-Net++ a more powerful and versatile model, particularly for complex segmentation tasks in medical imaging. The ability to capture multi-scale features and the guidance provided by deep supervision allow U-Net++ to achieve superior segmentation performance compared to the original U-Net. This is crucial in medical applications where precise and accurate segmentation is of utmost importance, such as the segmentation of coronary arteries from Computed Tomography Angiography (CTA) images. Overall, U-Net++ builds upon the strong foundation of U-Net, incorporating innovative techniques to address the challenges of medical image segmentation and push the boundaries of what is possible in this field.

U-Net++ is an advanced variant of U-Net that introduces several enhancements aimed at improving the segmentation performance further. The key innovations in U-Net++ include nested skip pathways and deep supervision. The architecture incorporates additional skip connections between intermediate layers in the encoder and decoder, creating nested pathways. These nested skip connections facilitate more refined feature aggregation and enhance the flow of information through the network. The presence of multiple skip pathways allows the network to better capture and combine features at different levels of abstraction, leading to more accurate and detailed segmentation results. Furthermore, U-Net++ employs deep supervision by providing intermediate outputs at various levels of the network [18].

2.5.2. Attention U-Net

Attention U-Net, introduced by Oktay et al. in 2018, enhances the original U-Net architecture by integrating attention mechanisms to tackle the challenge of irrelevant or noisy features in image segmentation. This innovative approach centers on the use of attention gates within the skip connections that link the encoder and decoder sections of the network. These gates dynamically evaluate the significance of feature maps from the encoder before merging them with those in the decoder, enabling the network to emphasize more relevant and informative features while suppressing less useful or noisy ones. The inclusion of attention mechanisms significantly improves the network's ability to focus on critical details within the image, thereby enhancing the overall quality of segmentation results. In practical terms, attention gates are implemented through a lightweight neural network that computes attention coefficients based on feature maps from both the encoder and decoder. These coefficients are then used to scale the encoder's features before concatenation with the decoder's features, ensuring that the network prioritizes significant features and minimizes irrelevant ones. This dynamic weighting process allows Attention U-Net to effectively handle challenges such as variations in object size, shape, and appearance, making it particularly valuable in medical imaging where precision and accuracy are paramount. Medical images often contain complex structures and subtle variations that require careful attention to detail for accurate segmentation. Attention U-Net's ability to adapt its focus based on feature importance enhances its versatility and effectiveness in medical applications. By selectively prioritizing critical regions within an image, the network ensures that essential features are highlighted and accurately segmented. This adaptability is crucial for medical imaging, where different regions may contain varying levels of detail and importance. In summary, Attention U-Net's integration of attention mechanisms represents a significant advancement in image segmentation. By dynamically weighting feature maps and focusing on relevant details, it addresses the limitations of traditional U-Net architectures, offering improved performance for complex segmentation tasks. This makes Attention U-Net particularly useful for medical imaging, where precise segmentation is essential for diagnosis and treatment planning. The innovative use of attention gates within the network underscores the ongoing evolution of neural network architectures, enhancing their capability to manage complex and nuanced image segmentation challenges[24].

2.5.3. ResU-Net

ResU-Net, proposed by Zhu et al. in 2018, enhances the U-Net architecture by integrating residual connections to improve training stability and performance. The primary innovation of ResU-Net is the incorporation of residual blocks designed to address the vanishing gradient problem and facilitate the training of deeper networks. In this architecture, traditional convolutional blocks are replaced with residual blocks that include shortcut connections bypassing one or more convolutional layers. These residual connections help maintain the flow of gradients during backpropagation, allowing for more effective network training. The encoder and decoder sections of ResU-Net retain the general structure of U-Net, but with the addition of residual blocks that enhance feature extraction and refinement. By incorporating residual connections, ResU-Net can capture more complex features, thereby improving segmentation accuracy. Furthermore, ResU-Net retains the skip connections from the original U-Net, combining them with the residual blocks to further refine feature maps and enhance overall segmentation performance. This combination of residual learning and skip connections results in a more powerful and stable architecture for image segmentation. The residual connections ensure that gradients can flow more smoothly through the network, reducing the likelihood of gradient vanishing and making it possible to train deeper models effectively. This allows ResU-Net to maintain high levels of detail and precision in its segmentation outputs, even in complex scenarios. By improving the flow of information within the network, ResU-Net enhances its ability to learn intricate patterns and features from medical images, leading to better segmentation results. The addition of residual blocks also improves the robustness of the network, making it less susceptible to training difficulties that often arise in deeper networks. This robustness is particularly important in medical imaging, where accurate segmentation can significantly impact diagnosis and treatment planning. By ensuring stable and effective training, ResU-Net can consistently produce high-quality segmentation results, even with challenging datasets. In summary, ResU-Net's integration of residual connections into the U-Net framework represents a significant advancement in image segmentation technology. By addressing the vanishing gradient problem and enhancing feature extraction, ResU-Net provides a more stable and powerful solution for medical image segmentation. This makes it an invaluable tool in medical imaging, where precise and reliable segmentation is essential for effective diagnosis and treatment planning. The combination of residual learning and traditional U-Net skip connections underscores the ongoing evolution of neural network architectures, highlighting the continual improvements being made to enhance performance and stability in complex image segmentation tasks [17].

2.5.4. TransUNet

TransUNet, introduced by Chen et al. in 2021, represents a groundbreaking approach that combines the strengths of convolutional neural networks (CNNs) and transformers to enhance image segmentation. TransUNet integrates transformer blocks into the traditional U-Net architecture, effectively leveraging the benefits of both convolution and transformer-based methods. In this hybrid architecture, the initial layers consist of convolutional operations designed to capture local features from the input image. These convolution layers are adept at identifying detailed and fine-grained patterns in the data, providing a strong foundation for subsequent processing. Following the convolution layers, TransUNet incorporates transformer blocks, which are crucial for capturing global context and long-range dependencies within the image. The transformer blocks utilize self-attention mechanisms to model relationships between different parts of the image, enabling the network to understand complex patterns and structures that span larger regions. This dual approach allows TransUNet to combine the precise, localized feature extraction capabilities of CNNs with the broader, context-aware processing of transformers. The overall architecture of TransUNet retains the encoder-decoder structure characteristic of U-Net but is significantly enhanced by the integration of transformer modules. The encoder captures and downsamples image features, while the decoder upsamples and reconstructs the segmented output. The transformer blocks within this framework add a layer of sophistication by enhancing the network's ability to process and integrate information from disparate regions of the

image. This leads to improved segmentation performance, particularly in challenging scenarios where variability and complexity are high.
TransUNet's hybrid approach results in a more robust and effective segmentation model, demonstrating significant improvements in accuracy and robustness compared to traditional models. By capturing both local details and global context, TransUNet is particularly well-suited for medical image segmentation tasks, where precision and the ability to discern subtle patterns are critical. The ability to model complex relationships within the image allows TransUNet to handle various medical imaging challenges, from identifying small lesions to segmenting large and complex anatomical structures. TransUNet's innovative combination of CNNs and transformers represents a significant advancement in the field of image segmentation. By integrating transformer blocks into the U-Net architecture, TransUNet effectively captures both detailed local features and comprehensive global context, enhancing segmentation accuracy and robustness. This makes TransUNet a powerful tool for medical imaging, providing reliable and precise segmentation essential for accurate diagnosis and treatment planning. The introduction of transformers into this domain highlights the ongoing evolution of neural network architectures, paving the way for future innovations that further enhance the capabilities of image segmentation technologies [14].

2.6. Models Implementation and Experimental Setup

All models were trained on 2D labeled slices, using 256 $\times$ 256 patches extracted from the target CTA scans. Each CTA slice underwent patch-wise normalization with zero mean and unit variance. The model architecture was developed using Keras with a TensorFlow backend. The 2D slices used for training represented the entire axial slice area in the CTA scan.
To address the issue of class imbalance in vessel regions, multiple strategies were adopted to improve the network's performance. A hybrid loss function was utilized during training, combining binary cross-entropy (BCE) and soft dice loss (DSC) with equal weighting to evaluate how well the predicted vessel regions aligned with the ground truth. The training process specifically targeted the heart region of interest (ROI) to help the network accurately distinguish vessels from the surrounding heart tissue. Training patches were derived by isolating the heart region from CTA slices, as depicted in Figure 2. Only patches with corresponding annotation masks were included in the dataset, while those without masks were excluded to ensure relevant and precise training data.

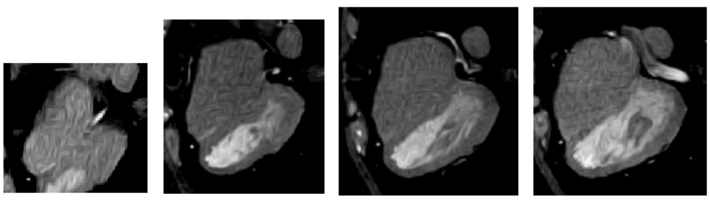

Figure 2: Training patches (Heart region).

A robust experimental framework was established to thoroughly assess the performance of deep learning models for coronary artery segmentation from CTA images. The dataset was systematically divided into distinct training, validation, and test subsets, ensuring complete separation to prevent overlap. An 80/20 split was employed for training and validation, with model selection and hyperparameter tuning driven by performance on the validation set. The training process utilized the Adam optimizer, starting with an initial learning rate of 0.0001. To optimize efficiency, adaptive monitoring mechanisms were incorporated: the learning rate was halved if the validation Dice similarity coefficient (DSC) did not improve over three consecutive epochs, and training was terminated if no further improvement was observed after five consecutive epochs. The models were trained with a batch size of 16 for up to 500 epochs, using gradient clipping with a maximum norm of 1.0 to ensure stable training dynamics.

Model performance was evaluated with a focus on coronary artery segmentation, using metrics such as the Dice similarity coefficient (DSC), Jaccard Index (JI), Sensitivity, and Precision. These metrics were derived by comparing the predicted segmentations against the ground truth annotations. This rigorous evaluation approach encompassed all critical elements, including dataset organization, preprocessing steps, architectural design, training parameters, and performance assessment metrics. The detailed findings presented in subsequent sections highlight the reliability and potential clinical significance of these methodologies in advancing coronary artery segmentation and improving cardiovascular care.

2.7. Performance Measures

In the realm of medical image segmentation, it is essential to evaluate deep learning model performance to gauge their accuracy and reliability. This study focuses on three common metrics: Dice Similarity Coefficient (Dice), Precision, and Recall. The Dice Similarity Coefficient (Dice) measures the overlap between predicted and ground truth segmentation. It is calculated using the following formula:

$$DSC = \frac{2|A \bigcap B|}{|A| + |B|} \tag{1}$$

In this context, A and B represent the number of pixels in the predicted segmentation and the ground truth segmentation, respectively, while intersection denotes the count of overlapping pixels between them. The Dice score ranges from 0 to 1, where a value of 1 signifies perfect overlap. The Jaccard Index, also called the Jaccard Similarity Coefficient or Intersection over Union (IoU), quantifies the similarity between the predicted and ground truth segmentation. It is calculated using the formula:

$$JI = \frac{|A \bigcap B|}{|A \bigcup B|} \tag{2}$$

Precision and Recall are complementary metrics that offer valuable insights into a model's performance. Precision evaluates the ratio of true positive predictions (TP) to the total number of positive predictions and is calculated as:

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

Recall measures the proportion of true positive predictions among all actual positive instances, calculated as:

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

True Positives (TP) are defined as the pixels accurately identified as belonging to the intended structure, whereas False Positives (FP) refer to pixels that the model incorrectly labels as part of the target structure. Conversely, False Negatives (FN) are pixels that genuinely belong to the target structure but are erroneously omitted by the model. Precision is a critical metric that gauges the accuracy of the model's predictions by measuring the proportion of correctly identified pixels among all pixels predicted as part of the target structure. Recall, on the other hand, assesses the model's ability to detect a significant portion of the actual target structure, emphasizing its sensitivity in identifying relevant features. Together, these metrics provide a comprehensive view of the model's segmentation performance.

In this study, metrics such as Dice, Precision, and Recall play a pivotal role in evaluating the model's capacity to accurately identify and segment coronary arteries from CTA images. These metrics provide critical insights into the model's performance, which are essential for informed clinical decision-making and patient care. High Dice scores highlight the model's ability to precisely locate and delineate coronary arteries, which is vital for diagnosing and managing cardiovascular diseases. Similarly, high Precision reflects reliable predictions, and high Recall ensures that a significant proportion of the target structures are successfully identified, further reinforcing the model's utility in medical applications.

3. Results and Discussion

3.1. Coronary Arteries Segmentation

The qualitative results of various U-Net based models and the impact of incorporating vesselness enhancement as a preprocessing step are illustrated in Figure 3. This figure specifically highlights the segmentation of the Right Coronary Artery (RCA) from computed tomography angiography (CTA) scans, clearly visible from the coronal view. The first column presents the RCA segmentation prior to applying vesselness enhancement, while the right column displays the results following its application. The effectiveness of vesselness enhancement in improving RCA segmentation is evident, as it significantly enhances the model's ability to capture the intricate features of the vessels within the CTA images. By emphasizing vessel structures, this preprocessing step allows for more precise delineation of coronary arteries, thereby facilitating better performance in subsequent segmentation tasks. Overall, these findings underscore the importance of vesselness enhancement in optimizing deep learning models for accurate coronary artery segmentation in clinical applications. On the other hand, figure 4 highlights the comparisons

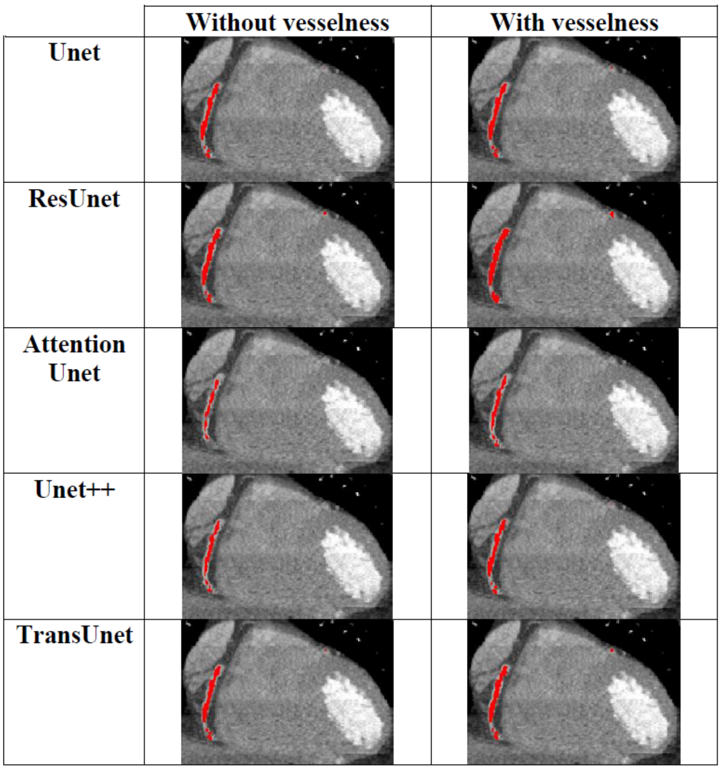

Figure 3: Effect of vesselness enhancement on the segmentation output of the Right Coronary Artery (RCA) across different models (coronal view). The left column displays the segmentation results without vesselness enhancement, while the right column shows the results with vesselness enhancement.

between the outputs of the different U-Net based models and the ground truth annotations. These comparisons provide valuable insights into the accuracy and precision of each model's segmentation, highlighting areas where they excel or fall short in capturing the true extent of the coronary arteries. These comparisons are made for both coronal slices and 3D visualizations of the coronary artery tree, allowing for a thorough assessment of the models' performance. The first column in figure 4 shows the Right Coronary Artery (RCA) segmentation in the coronal slice, the ground truth annotations serve as the reference standard for arterial boundaries, enabling a direct evaluation of how well each model's segmentation aligns with the true extent of the arteries. This visual comparison highlights the strengths and weaknesses of each model in terms of their ability to accurately delineate the arterial structures and avoid over or under segmentation. Similarly, the second column shows the 3D visualizations of the coronary artery tree provide a comprehensive view of how the models capture the complex geometry and connectivity of the arterial network. By comparing these 3D renderings with the ground truth, it can be assessed the models' performance in preserving the continuity of arterial branches, avoiding discontinuities or leakage, and maintaining the overall topology of the coronary artery tree. The

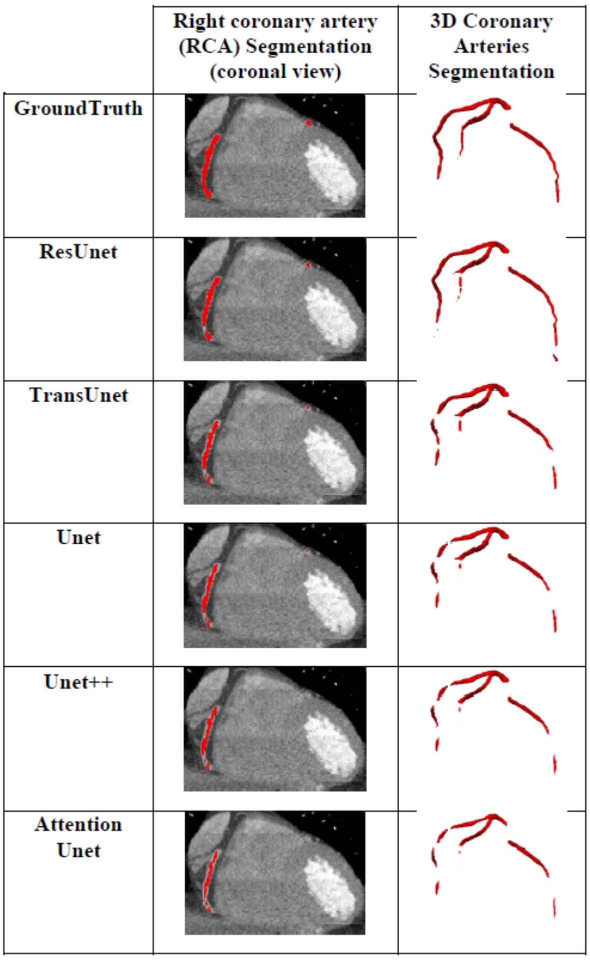

Figure 4: Different Unet variants Models segmentation output compared to the ground-truth. (Left column) Right coronary artery (RCA) Segmentation (coronal view). (Right column) 3D coronary arteries segmentation.

visual inspection of the ResUnet presented by the second row in figure 4, clearly demonstrate that the ResU-Net model outperforms the other U-Net based models in segmenting coronary arteries, both in coronal slice views and 3D visualizations of the arterial tree. In the coronal slice images, the ResU-Net model consistently provides the most accurate and precise segmentation of the right coronary artery (RCA). The arterial boundaries are well-defined, with minimal leakage into surrounding tissues or structures. In contrast, the other models, such as U-Net, Attention U-Net, U-Net++, and TransU-Net, exhibit varying degrees of over or under segmentation in the coronal slices. While some models may capture the majority of the arterial pixels, they struggle to accurately delineate the boundaries, leading to the inclusion of extraneous regions or missing parts of the RCA. These visual discrepancies are reflected in their lower quantitative scores compared to the ResU-Net.

The superiority of the ResU-Net model becomes even more apparent in the 3D visualizations of the entire coronary artery tree. The arterial network is clearly and accurately

depicted, with well-defined branches and smooth transitions between segments. The 3D renderings generated by the ResU-Net closely resemble the ground truth annotations, demonstrating its ability to capture the complex geometry and connectivity of the coronary arteries.
The quantitative performance of the various U-Net based models in segmenting coronary arteries from computed tomography angiography (CTA) images is clearly summarized in the below tables (1-5). The tables provide a comprehensive summary of the 5-fold cross-validation results for the various U-Net based models in segmenting coronary arteries from computed tomography angiography (CTA) images.
The original U-Net model achieved a DSC of 0.809 and a JI of 0.708. These values indicate a robust ability to segment coronary arteries, reflecting a good overlap between the predicted and ground truth regions. The Recall of 0.772 suggests that the model successfully identifies a substantial proportion of the actual arterial pixels, while the Precision of 0.876 indicates a relatively low rate of false positives, demonstrating that most of the predicted pixels correspond to the arteries.

Table 1: U-Net results for the k-fold (k=5) cross-validation.

Fold	DSC	JI	Recall	Precision
Fold 1	0.810	0.710	0.770	0.870
Fold 1	0.800	0.705	0.765	0.880
Fold 1	0.820	0.715	0.780	0.875
Fold 1	0.810	0.710	0.770	0.870
Fold 1	0.805	0.700	0.775	0.885
Avg	0.809	0.0670	0.772	0.876

The ResU-Net model showed significant improvements over the original U-Net, achieving a DSC of 0.865 and a JI of 0.764. The increased Recall of 0.881 indicates that this model is particularly effective at detecting arterial structures, capturing a higher percentage of true positive pixels. Furthermore, the Precision of 0.892 remains high, suggesting that the model maintains accuracy in its predictions, making it a strong candidate for clinical applications.

Table 2: ResU-Net results for the k-fold (k=5) cross-validation.

Fold	DSC	JI	Recall	Precision
Fold 1	0.867	0.765	0.881	0.892
Fold 1	0.861	0.760	0.885	0.890
Fold 1	0.864	0.768	0.878	0.895
Fold 1	0.869	0.762	0.883	0.889
Fold 1	0.865	0.766	0.879	0.893
Avg	0.865	0.764	0.881	0.892

The Attention U-Net model achieved a DSC of 0.775 and a JI of 0.677, indicating moderate segmentation performance. Notably, it recorded the highest Recall of 0.914, suggesting that the attention mechanism effectively enhances the model's ability to identify arterial pixels. However, this model's Precision of 0.678 is the lowest among the evaluated networks, indicating a higher rate of false positives, which may limit its utility in precise clinical settings.

Table 3: Attention U-Net results for the k-fold (k=5) cross-validation.

Fold	DSC	JI	Recall	Precision
Fold 1	0.774	0.645	0.910	0.691
Fold 1	0.780	0.690	0.920	0.680
Fold 1	0.790	0.700	0.930	0.690
Fold 1	0.770	0.680	0.910	0.670
Fold 1	0.760	0.670	0.900	0.660
Avg	0.775	0.677	0.914	0.678

The U-Net++ model achieved a DSC of 0.790 and a JI of 0.676, demonstrating solid segmentation capabilities. Like the Attention U-Net, it also recorded a high Recall of 0.914, indicating its effectiveness in capturing arterial structures. However, the Precision of 0.708 suggests that while it is good at identifying true positives, it also presents a risk of false positives, albeit to a lesser extent than the Attention U-Net.

Table 4: U-Net++ results for the k-fold (k=5) cross-validation.

Fold	DSC	JI	Recall	Precision
Fold 1	0.790	0.667	0.913	0.709
Fold 1	0.780	0.667	0.910	0.700
Fold 1	0.795	0.680	0.915	0.715
Fold 1	0.785	0.675	0.910	0.705
Fold 1	0.800	0.690	0.920	0.710
Avg	0.790	0.676	0.910	0.708

The TransU-Net model achieved a DSC of 0.812 and a JI of 0.676, indicating strong segmentation performance. Its Recall of 0.917 is the highest among the models, suggesting excellent sensitivity in detecting arterial structures. The Precision of 0.719 reflects a reasonable balance between true positives and false positives, making it a viable option for segmentation tasks.

3.2. Discussion

First, to emphasize the significance of the vesselness enhancement step within the proposed framework, all models were evaluated and tested both with and without this preprocessing step. This comparative analysis demonstrates the substantial impact that vesselness enhancement can have on the accuracy and efficiency of the segmentation process. A series of experiments were conducted to assess the performance of each model

Table 5: TransU-Net results for the k-fold (k=5) cross-validation.

Fold	DSC	JI	Recall	Precision
Fold 1	0.810	0.667	0.912	0.709
Fold 1	0.805	0.675	0.915	0.715
Fold 1	0.815	0.680	0.920	0.725
Fold 1	0.810	0.675	0.910	0.720
Fold 1	0.820	0.685	0.925	0.730
Avg	0.812	0.676	0.917	0.719

under these two conditions, yielding valuable insights into their effectiveness. The results of this comparative study are summarized in Table 6.

Table 6: Comparison of Segmentation results for the different models (without vesselness step).

Model	DSC	JI	Recall	Precision
U-Net	0.780	0.639	0.751	0.869
ResU-Net	0.852	0.742	0.858	0.871
Attention U-Net	0.752	0.607	0.902	0.655
U-Net++	0.773	0.627	0.898	0.693
TransU-Net	0.801	0.668	0.908	0.711

The impact of incorporating the vesselness enhancement step as a preprocessing technique prior to deep learning segmentation has been analyzed across various U-Net based models, as illustrated in Table 7. A comparison with Table 6 clearly demonstrates that vesselness enhancement significantly improves the segmentation capabilities of the models. Specifically, Table 7 presents a detailed comparison of segmentation performance among different U-Net variants. The ResU-Net model achieved the highest Dice Similarity Coefficient (DSC) of 0.865 and a Jaccard Index (JI) of 0.764, indicating its superior performance. Its Recall value of 0.881 suggests effective identification of true positive arterial pixels, while a Precision of 0.892 reflects a low false-positive rate. The enhanced performance of ResU-Net highlights the advantages of integrating residual connections within the U-Net framework, which facilitates improved gradient flow and feature propagation. This capability enables ResU-Net to capture complex features more effectively, resulting in enhanced segmentation accuracy. Overall, these findings underscore the importance of vesselness enhancement and the ResU-Net architecture in advancing coronary artery segmentation for clinical applications.

While the Attention U-Net and U-Net++ models exhibited the highest Recall of 0.914, their Precision scores were lower compared to the ResU-Net. This suggests that these models are effective in detecting arterial structures but may also include more false positives in their segmentation outputs. The attention mechanism in the Attention U-Net and the dense skip connections in U-Net++ seem to enhance the models' ability to capture arterial features, but they may also introduce more false positives.

The TransU-Net model strikes a balance between Recall and Precision, achieving a Recall

Table 7: Comparison of Segmentation results for the different models (with vesselness step).

Model	DSC	JI	Recall	Precision
U-Net	0.809	0.708	0.772	0.876
ResU-Net	0.865	0.764	0.881	0.892
Attention U-Net	0.775	0.677	0.914	0.678
U-Net++	0.790	0.676	0.910	0.708
TransU-Net	0.812	0.676	0.917	0.719

of 0.917 and a Precision of 0.719 as depicted in table 5. This suggests that the TransU-Net can detect most of the arterial pixels while maintaining a reasonable false-positive rate. The combination of the U-Net architecture with the Transformer module in the TransU-Net may help the model better distinguish between arterial and non-arterial regions, leading to a more balanced segmentation performance.

The choice of the best model for coronary artery segmentation from CTA images depends on the specific requirements of the application. If maximizing the detection of arterial pixels is the primary goal, the Attention U-Net or U-Net++ models may be preferred. However, if overall segmentation accuracy is more important, the ResU-Net or TransU-Net models are better suited. The superior performance of the ResU-Net model in both coronal slice and 3D segmentation has significant implications for its clinical applicability. Its ability to accurately delineate coronary arteries can aid cardiologists in various tasks, such as stenosis detection, plaque characterization, and treatment planning. The clear and consistent visualization of the arterial tree provided by the ResU-Net can enhance diagnostic accuracy and facilitate better decision-making in patient care.

4. Conclusion

This paper introduces a well-organized pipeline that integrates vesselness enhancement, heart region of interest (ROI) extraction, and the ResUNet deep learning architecture to achieve precise segmentation of coronary arteries in computed tomography angiography (CTA) images.The results demonstrate the method's exceptional performance metrics, particularly highlighting the importance of vesselness enhancement in improving the visibility of coronary arteries, which enables the deep learning model to capture intricate details and boundaries more effectively. The heart ROI extraction step further enhances efficiency by allowing the model to focus on relevant regions, thereby reducing computational complexity. Additionally, a comprehensive comparative analysis of five U-Net based models—U-Net, ResU-Net, Attention U-Net, U-Net++, and TransU-Net—reveals that the ResU-Net model outperforms others with a Dice Similarity Coefficient (DSC) of 0.865 and a low false-positive rate, making it suitable for clinical applications. While Attention U-Net and U-Net++ show high recall for detecting arterial structures, they struggle with precision, potentially increasing false positives. TransU-Net offers a balanced approach with competitive recall and precision values. The findings emphasize the need to choose models based on specific clinical requirements and suggest exploring hybrid approaches that combine strengths from multiple models to enhance the clinical utility of deep learning in cardiovascular imaging, ultimately improving patient outcomes in coronary artery disease management.

Funding
The APC was funded by American university of Middle East in Kuwait.
Conflict of interest
The authors declare no conflict of interest

References

[1] Brown JC, Gerhardt TE, and Kwon E. *Risk Factors for Coronary Artery Disease. Risk Factors in Coronary Artery Disease.* CRC Press, Boca Raton, 2006.

[2] Trigka M and Dritsas E. Long-term coronary artery disease risk prediction with machine learning models. *Sensors*, 23, 2023.

[3] Alirr OI and Rahni AAAbd. Survey on liver tumour resection planning system: Steps, techniques, and parameters. *J Digit Imaging*, 33:304–323, 2019.

[4] Omar Alirr, Razan Alshatti, Shahad Altemeemi, Shahad Alsaad, and Alzahraa Alshatti. Automatic brain tumor segmentation from mri scans using u-net deep learning. In *2023 5th International Conference on Bio-engineering for Smart Technologies (BioSMART)*, pages 1–5, Paris, France, 2023. IEEE).

[5] Fernando Pascaner A, Rosato A, Fantazzini A, Vincenzi E, Curzio Basso, Secchi F, and et al. Automatic 3D Segmentation and Identification of Anomalous Aortic Origin of the Coronary Arteries Combining Multi-view 2D Convolutional Neural Networks. *Journal of Imaging Informatics in Medicine*, 37:884–891, 2024.

[6] Pan LS, Li CW, Su SF, Tay SY, Tran QV, and Chan WP. Coronary artery segmentation under class imbalance using a u-net based architecture on computed tomography angiography images. *Scientific Reports*, 11:1–7, 2021.

[7] Alirr OI. Automatic deep learning system for covid-19 infection quantification in chest ct. *Multimed Tools Appl*, 81:527–41, 2022.

[8] Gao Z, Liu X, Qi S, Wu W, Hau WK, and Zhang H. Automatic segmentation of coronary tree in ct angiography images. *Int J Adapt Control Signal Process*, 33:1239–47, 2019.

[9] Alirr OI, Al-Absi HRH, Ashtaiwi A, and Khalifa T. Efficient extraction of coronary artery vessels from computed tomography angiography images using resunet and vesselness. *Bioengineering*, 11, 2024.

[10] Song A, Xu L, Wang L, Wang B, Yang X, Xu B, and et al. Automatic coronary artery segmentation of ccta images with an efficient feature-fusion-and-rectification 3d-unet. *IEEE J Biomed Health Inform*, 26:4044–55, 2022.

[11] Serrano-Anton B, Otero-Cacho A, Lopez-Otero D, Diaz-Fernandez B, Bastos-Fernandez M, Perez-Munuzuri V, and et al. Coronary artery segmentation based on transfer learning and unet architecture on computed tomography coronary angiography images. *IEEE Access*, 11:75484–96, 2023.

[12] Kaba Ş, Haci H, Isin A, Ilhan A, and Conkbayir C. The application of deep learning for the segmentation and classification of coronary arteries. *Diagnostics*, 13, 2023.

[13] Wang Q, Xu L, Wang L, Yang X, Sun Y, Yang B, and et al. Automatic coronary artery segmentation of ccta images using unet with a local contextual transformer. *Front Physiol*, 14, 2023.

[14] Chen J, Lu Y, Yu Q, Luo X, Adeli E, Wang Y, and et al. Transunet: Transformers make strong encoders for medical image segmentation. *ArxivOrgJ*, 2021.

[15] Du Z Meng Y, Zhao C, Dong M, Pienta D, Xu Za, and et al. Automatic extraction of coronary arteries using deep learning in invasive coronary angiograms. *Technology and Health Care*, 31:2303–2317, 2023.

[16] Tao X, Dang H, Zhou X, Xu X, and Xiong D. A lightweight network for accurate coronary artery segmentation using x-ray angiograms. *Front Public Health*, 13, 2022.

[17] Diakogiannis FI, Waldner F, Caccetta P, and Wu C. Transunet: Transformers make strong encoders for medical image segmentation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 162:94–114, 2020.

[18] Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, and Liang J. Unet++: A nested u-net architecture for medical image segmentation. In *4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018.*, pages 3–11, Granada, Spain,, 2018. Springer, Cham.

[19] Lei Y, Guo BJ, Fu Y, Wang T, Liu T, Curran W, and et al. Automated coronary artery segmentation in coronary computed tomography angiography (ccta) using deep learning neural networks. In *Proc. SPIE 11318, Medical Imaging 2020: Imaging Informatics for Healthcare, Research, and Applications*, Houston, Texas, United States, 2020. SPIE Medical Imaging.

[20] Alirr OI and Aizzuddin A. Rahni A. Hepatic vessels segmentation using deep learning and preprocessing enhancement. *J Appl Clin Med Phys*, 24, 2023.

[21] Kim D. Hepatic vessel segmentation on contrast enhanced ct image sequence for liver transplantation planning. *J Biomed Sci Eng*, 6:498–503., 2013.

[22] Irr OIA and Rahni AAA. Automatic volumetric localization of the liver in abdominal ct scans using low level processing and shape priors. In *2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, Kualampur, Malaysia, 2015. ICSIPA).

[23] Galea RR, Diosan L, Andreica A, Popa L, Manole S, and Bálint Z. Region-of-interest-based cardiac image segmentation with deep learning. *Applied Sciences*, 11, 2021.

[24] Oktay O, Schlemper J, Folgoc L Le, Lee M, Heinrich M, Misawa K, and et al. Attention u-net: Learning where to look for the pancreas. *arXiv*, 2018.