Pakistan Journal of Scientific Research, PJOSR
ISSN (p): 0552-9050
Volume: 3, Number: 1, Pages: 124- 129, Year: 2023

Automated Segmentation of Coronary Arteries using Attention-Gated UNet for Precise Diagnosis

Shabir Hussain[1,*], Junaid Abdul Wahid[2], Muhammad Ayoub[3,*] , Huan Tong[1,] and [4]Rukhshanda Rehman

[1]School of Architecture, Harbin Institute of Technology, Shenzhen, China
[2]School of Computer and Artificial Intelligence, Zhengzhou University, Zhengzhou,450001, Henan, China
[3]School of Computer Science and Engineering, Central South University, Changsha, 336017, Hunan, China
[4]Department of Zoology, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan

Corresponding authors: (e-mail: shabir.nicaas@gmail.com; ayoubncbae@gmail.com).

Received: 15/04/2023, Revised: 05/06/2023, Accepted: 08/06/2023

***Abstract*- Computed Tomography Angiography (CTA) has revolutionized coronary artery disease diagnosis and treatment with its high-resolution and non-invasive advantages. Precision in coronary artery segmentation is critical for accurate diagnosis and effective treatment. In this study, we introduce an innovative two-stage approach utilizing fully convolutional neural networks (CNNs) for coronary artery segmentation. Our model combines coarse segmentation with fine segmentation, significantly enhancing accuracy. This dual-stage strategy outperforms conventional single-stage methods, as validated through empirical evaluations. Our approach demonstrates a substantial performance improvement, with a MEAN Jaccard Similarity of 0.8217 and a MEAN Dice Similarity Coefficient of 0.9005, affirming its potential in advancing medical imaging diagnostics and improving coronary artery segmentation.**

***Index Terms*-- Coronary Artery, Attention Gated UNet, Two-Stage Segmentation, healthcare, biomedical**

I. INTRODUCTION

Cardiovascular disease (CVD) looms as a formidable global health challenge, encompassing a spectrum of conditions impacting the heart and blood vessels. The severity of this health crisis is emphasized by the "China Cardiovascular Health and Disease Report 2023"[1]. The report highlights that a combination of factors, including evolving national lifestyles, an aging population, the rapid urbanization of rural areas, and the prevalence of unhealthy habits, has significantly contributed to the increasing prevalence of cardiovascular diseases. [2]. The implications of these conditions are far-reaching, as they progressively impact human well-being and escalate health risks on an unprecedented level. The stark reality reflected in statistics, with 46.66% of deaths transpiring in rural enclaves and 43.81% in urban centers, highlights the distressing omnipresence of cardiovascular afflictions that have catapulted to the forefront of health priorities. Particularly, coronary heart disease stands out due to its widespread occurrence and severity, demanding special consideration as a quintessential representation of cardiovascular diseases (CVD). The intricate interaction of various factors, primarily coronary artery stenosis, leads to the narrowing of blood vessels, which, in turn, jeopardizes the supply of oxygen to crucial myocardial cells [3]. The outcome is myocardial infarction, a catastrophic cardiovascular event. In the pursuit of precise diagnoses and effective treatments, computed tomography angiography (CTA) has emerged as a vital tool in the medical field, renowned for its non-invasive approach and high-resolution imaging capabilities. Within this medical context, the pursuit of accurate coronary artery segmentation holds a pivotal role, serving as a fundamental element supporting both diagnostics and therapeutic interventions [4]. Though treatment plays a crucial role, the current scenario is marred by the prevailing manual segmentation approach, which requires significant efforts from experts, consumes time, and results in inefficiencies [5]. This evident gap emphasizes the urgent requirement for innovative solutions that leverage advanced technologies to simplify and improve the precision of coronary artery segmentation. In this paper, we present the following key contributions:

- We proposed a two-stage segmentation approach, combining coarse and fine segmentation for enhanced coronary artery segmentation precision.
- Utilization of advanced deep learning models, such as 3D U-Net and Attention Gate U-Net, showcasing their effectiveness in medical imaging diagnostics.
- Significant performance enhancement, which improves the accuracy of coronary artery segmentation, benefiting coronary artery disease diagnosis and treatment.

II. RELATED WORK

As computer vision has advanced, numerous techniques have been employed for the segmentation of coronary arteries in CTA images. These techniques can be broadly categorized into traditional image segmentation methods and deep learning-based image segmentation methods[6]. Within the realm of traditional image processing, several algorithms based on edges and regions have been applied to the segmentation of coronary arteries in CTA images. Cai, Renhui, et al. [7] introduced the utilization of the watershed algorithm for delineating coronary artery sub-regions within preprocessed CTA images. Subsequently, they applied the region-growing technique to amalgamate these delineated regions, ultimately achieving a comprehensive three-dimensional segmentation of the coronary arteries. Zeng, An, et al. [8] employed the level set method for the identification and segmentation of coronary arteries within CT images, ultimately producing a three-dimensional model.

Zhu, Hongyan, et al. [9] initiated the process by partitioning the image into multiple regions through the application of the threshold-based region growing algorithm. They further processed the regions irrelevant to the target, culminating in the utilization of graph cut theory for the segmentation of a network graph and the acquisition of coronary artery segmentation images. Zhou, Long, et al. [10] initially employed a threshold-based method to perform a preliminary segmentation of the three-dimensional dual-source CT image. Subsequently, they utilized an interactive approach to segment the sections where the left and right coronary arteries connect to the aorta. Finally, the segmentation of the three-dimensional coronary arteries was executed based on the anatomical position of the coronary artery origins, involving the use of morphological techniques and accounting for the relationship between adjacent layers in the three-dimensional tomographic image.

In recent years, deep learning methodologies have experienced significant advancements. Notably, Convolutional Neural Networks (CNNs), as described by [11-13], have demonstrated impressive performance across a wide range of image processing tasks.

In the realm of image segmentation, Fully Convolutional Neural Networks (FCNs), as discussed in [14], find common usage in segmentation tasks. This is attributed to their end-to-end input-output structure, automatic extraction of effective segmentation features, and their capability to generate consistent-sized segmentation maps. They often yield improved results. However, when dealing with high-resolution 3D CTA images, the 3D CNN structure and the accompanying intermediate features demand substantial computational resources. In the context of deep learning-based coronary artery segmentation, a common approach is to reduce the resolution and perform direct coronary artery segmentation. Alternatively, this can be achieved by partitioning the image into smaller three-dimensional slices (patches) or two-dimensional slices, allowing for coronary artery segmentation at a high resolution. Gupta et al. [15] conducted a comprehensive process in their research. They initiated the extraction of the outer contour of the heart by leveraging grayscale information. This information was then fused with prior knowledge. To achieve pre-segmentation of the CTA image slices, they employed a three-dimensional fully convolutional neural network. Ultimately, they employed the conditional random field approach to accomplish the precise segmentation of the coronary arteries, achieving a complete and accurate segmentation. Zhao, Guangzhe, et al. [16] applied the U-Net network with an attention mechanism to segment CTA images at the slice level, achieving notable results. Yan, Qingsen, et al. [17] employed a 3D FCN network, complemented by an Attention Gate (AG), for down-sampling the CTA image and subsequent segmentation. They further utilized the level set algorithm for post-processing. Additionally, they employed a 3D U-Net to carry out the segmentation of coronary arteries. This involved segmenting the original image into multiple smaller slices [18], followed by segmentation of these slices into smaller regions, ultimately reconstructing them into a three-dimensional representation of the coronary arteries. In addition to using CNN for direct segmentation [19], some studies construct data structures based on the structure of coronary arteries to assist segmentation, such as tree structures. In response to these problems, this study proposes a two-stage fully convolutional neural network coronary artery segmentation method, which effectively reduces the computational resources required for high-resolution 3D CTA image segmentation. Compared with other segmentation schemes at the original resolution, such as block segmentation and slice segmentation, the fine segmentation guided by coarse segmentation obtains global information and effectively improves the accuracy of coronary artery segmentation.

III. TWO-STAGE BASED CORONARY ARTERY SEGMENTATION MODEL

This paper proposes a two-stage coronary artery segmentation method, which is mainly divided into two stages: coarse segmentation and block segmentation. Its main flow chart is shown in Figure 1.

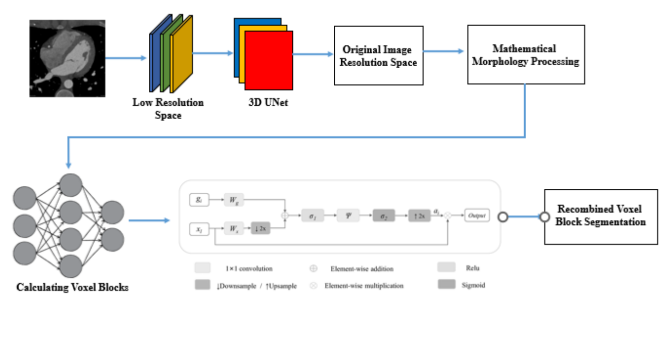

Fig. 1. The framework of two-stage fully convolutional neural networks

The method's workflow is broadly divided into two key stages: a preliminary segmentation stage and a subsequent block-based segmentation stage.

Coarse Segmentation: In this stage, the entire image is initially segmented as a whole. It's important to note that due to the high image resolution at this stage, the segmentation results might be constrained by computational resources, leading to limitations in overall segmentation accuracy.

Block-Based Segmentation: Building upon the results of the preliminary segmentation, the second stage focuses on dividing the image into smaller, manageable blocks. These blocks are achieved through morphological processing and are known for their simpler structural characteristics. Block segmentation excels at capturing local details, and for this part of the process, the use of U-Net with an attention gate has proven particularly effective.

A. Coarse Segmentation

In the coarse segmentation stage, this paper adopts the commonly used 3D U-Net structure as the segmentation network for this part. Since 3D CTA is a high-resolution image, if it is directly input into 3DCNN to obtain high-latitude intermediate features, it will take up a lot of computing resources, and the interpolated image is usually used as input to the neural network [20]. However, the down-sampled image will lose part of the image information, which is quite unfavorable for segmenting small coronary tubular objects. At this stage, a compromise solution is used to scale the image to a low-resolution space while reducing the parameters of the neural network [13] to achieve coarse segmentation and reduce the consumption of computing resources. The structure of the segmentation network adopted at this stage is shown in Figure 2. U-Net is in the shape of "U" as a whole and is mainly divided into two parts. The first part is the encoder part that locally reduces the resolution of the feature image.

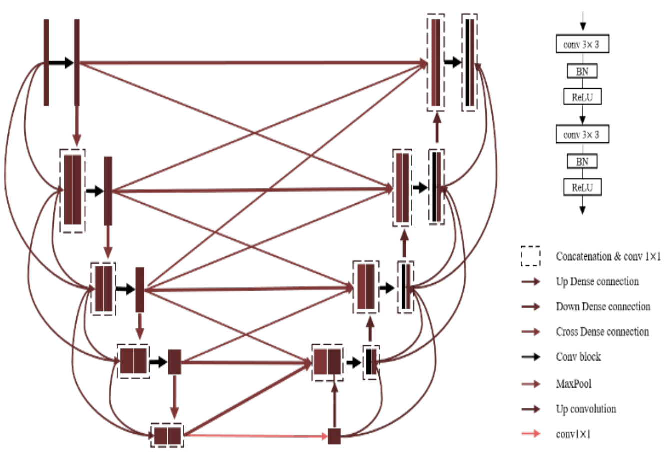

Fig. 2. Network architecture of U-Net used in this study

At each resolution scale, two Conv + ReLU + BN modules are passed, and then the Max Pooling layer is used to reduce the spatial size of the feature map and increase its channel number [21]. The second part is the decoder part that restores the image resolution, through the up-sampling operation to expand the spatial characteristics of the image while reducing the number of channels. In the decoding process, the decoder fuses the features from each scale in the encoding process through the concatenation operation to restore the details of the image. The predicted probability map is fed into the last convolutional layer using a convolutional layer with a kernel size of $1 \times 1 \times 1$ and a sigmoid layer. Except for the convolution layer of the last layer, the size of the convolution kernel of other convolution layers is $3 \times 3 \times 3$, and the step size of all convolution operations is 1. The network input is an interpolated image, and the output is a segmentation map at the same resolution.

B. 3D Slice Segmentation

1) Making slice

Making suitable slices that contain the segmentation target is an important part of performing slice segmentation, which is related to the integrity of the results in slice segmentation [22]. Since the coronary arteries are small tubular structures, the details lost in the coarse segmentation stage need to be improved by the block segmentation at the original resolution [14]. In order to reduce the generation of slices that do not contain segmentation objects, the process of making slices uses the result of coarse segmentation as a priori region. In the first stage, a coarse segmentation of the coronary arteries is obtained, and the skeletonized morphological skeletonization processing to refine the coarsely segmented labels into skeleton lines. The skeleton line represents the central area of the tubular object, that is, the center area of the tubular section. Using the skeleton line point as the center of the slice to obtain the slice, the set of slices will contain the coronary artery region in the coarse segmentation; however, it will produce a large number of overlapping regions[23]. This article uses a "digging" cut block collection. Specifically, firstly, a certain skeleton line point is used as the center of the cut block, and the rest of the skeleton line points included in the cut block will not be used as the center to obtain the cut block, that is, all the skeleton line points in the cut block are "removed". Through the loop, all the skeleton points are excavated, and the cutting of a single sample is completed. This method can effectively ensure that all regions from the coarsely segmented coronary arteries are included in the block set. Use this method to obtain multiple cuts for each coarse segmentation result to form a set of cuts.

2) Attention gate mechanism

In recent years, the attention mechanism has been widely used in machine vision [24]. The attention mechanism is similar to human beings focusing on certain information. It is used to focus on key parts of the input image and ignore information that is irrelevant to the task. Some attention mechanism modules are added to the convolutional neural network to enable the network to focus on key spatial object information and improve the accuracy of the results. The attention gate module (Attention Gate) was proposed and combined into the fully convolutional neural network [25]. The attention gate uses a combination mechanism of deep features and shallow features, and integrates the two features to provide an attention mechanism for the current feature. As shown in Figure 3, note that the gate module has two input features X and G, where X is the input feature and G is the gating feature, where X and G have the same size. The specific calculation method of attention is shown in Equation 1.

$$X_{att} = \sigma_2\left(\psi \times \sigma_1\left(W_g \times G + W_x \times X + b_g\right) + b_\psi\right) \cdot X \qquad (1)$$

Where W_g, W_x, W_ψ are convolution operations with a convolution kernel size of 1 × 1 × 1, and g, b_ψ are biases. σ_1, σ_2 are ReLU and Sigmoid activation functions respectively and $\times$ is the convolution operation, which is the corresponding multiplication of each element of the matrix.

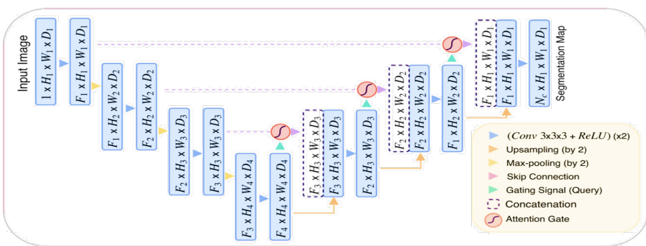

Fig. 3. Architecture of Attention gate used in our study

3) Splitting the Network

In the dicing segmentation, this paper uses the 3D U-Net model with the attention gate module (Attention Gate). For block segmentation, since the input block is much smaller than the entire original image, it becomes feasible to use a complex model with more parameters. As shown in Figure 4, the overall framework of the network is similar to the network framework in Figure 3. The attention gate is added to the jump operation of U-Net to enhance the jump feature, and the deep feature is used as the gating feature input to the Attention gate. At the same time, in terms of parameters, compared with the coarse segmentation, except for the last convolutional layer, the number of convolutional kernels in each convolutional layer is twice the coarse segmentation, the network's fitting ability is stronger. Input the slices in 2.2.1 into the network to obtain slice-level segmentation, and reconstruct the slices according to the position of the slices in the original image when making them. In the process of reconstruction, there are parts that are repeatedly cut and repeated. For the repeated parts, the maximum value of the generation probability among multiple cut blocks in the prediction is used as the generation probability

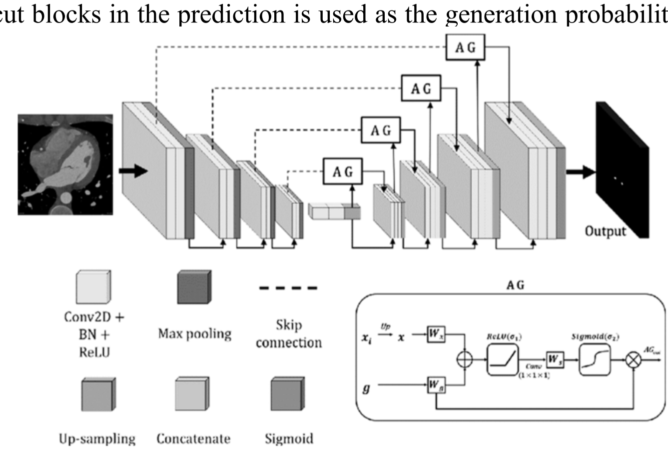

Fig. 4. Network architecture of attention gate U-Net

IV. Experimental Design And Analysis

a. Dataset

The CTA image data is provided by a hospital, with a total of 200 cases. The image size is 512 × 512 × (206 − 275), the planar resolution is 0.29~0.43 mm2, and the spacing of scanning layers is 0.25~0.45 mm. Two radiologists independently labeled the left and right coronary arteries in each image, and the results of their manual segmentation were determined by cross-validation. As shown in Figure 6, an example of the image is given, a and b are the 3D display of the 3DCTA image and label, c and d are the truncated display of the image and the label, and the display software is the open-source software 3D Slicer.

Figure 5 illustrates the segmentation results visualization. Each row in this figure represents a different stage of coronary CTA data; row (a) represents the original image, row (b), represents the segmentation results from the proposed framework, row (c), represents the ground truth, and row (purple) represents the human-labeled coronary arteries. Overall, the best visualization performance is achieved by the joint architecture that combines 3D FCN, attention gate, and level set function.

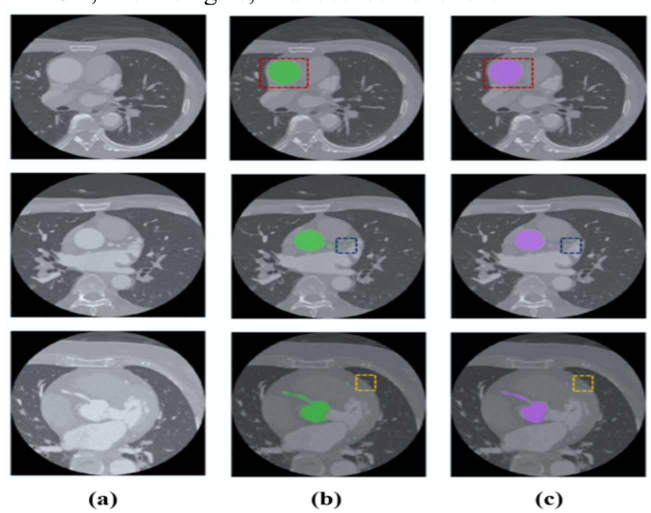

Fig. 5. Segmentation results of different methods for three stages of coronary CTA data

b. Experimental Parameter Settings

The entire dataset, comprising both training and testing images, is used for model development, with a training-to-testing ratio of 3:1, resulting in 160 cases for the training set and 40 cases for the test set. All experiments were conducted on an RTX 3090 GPU platform. Given the data's volume, the training process for all networks was set to 50 epochs. The optimizer employed was Adam, with a learning rate of 0.002. In the coarse segmentation stage, network training used the cross-entropy loss function, and the input image size was configured at 256 × 256 × 128, with specific calculations detailed in Equation 2.

$$L_1(\hat{Y}, Y) = [Y \log \hat{Y} + (1-Y)\log(1-\hat{Y})] \quad (2)$$

In the block segmentation stage, the input block size is 32 × 32 × 32, and the loss function is the Dice loss function. The specific calculation is shown in Equation 3.

$$L_2(\hat{Y},Y) = 1 - \frac{2|\hat{Y} \cap Y|+1}{|\hat{Y}| + |Y|+1} \quad (3)$$

where $\hat{Y}$ is the predicted probability output by the network, and Y is the real label. Finally, the segmentation results are evaluated using the Dice coefficient. 4.3. Analysis of experimental results In order to verify the effectiveness of the model, this paper designs a conventional segmentation method as a comparative experiment.

1) Slice Segmentation: The 3D CTA image is sliced along the Z-axis with a size of 512 × 512 as the input of the segmentation network 2D U-Net, and the segmentation is reorganized into a 3D image.

2) Direct segmentation: Similar to coarse segmentation, the 3DCTA image is scaled to a lower resolution space and segmented using 3D U-Net. 3) Segmentation: The 3D CTA image is neatly cut into several 32 × 32 × 32 small segments, which are input into the segmentation network to obtain the segmentation results and then reorganized to obtain the final segmentation results.

The networks of all comparative experiments use the same scale of U-Net, and the loss function uses the Dice loss function.

The networks of all comparative experiments use the same scale of U-Net, and the loss function uses the Dice loss function.

TABLE 1: COMPARATIVE EXPERIMENTS USE U-NET OF THE SAME SIZE, AND THE LOSS FUNCTION USES THE DICE LOSS FUNCTION

Method	Dice (%)
Direct split	75.33
slice division	71.56
Cut into pieces	68.72

Experimental results show that our two-stage coarse-fine segmentation is effective in segmenting coronary arteries in 3D CTA images. At the same time, using the two-stage segmentation method has a certain improvement compared to other conventional segmentation methods.

TABLE 2: COMPARATIVE PERFORMANCE EVALUATION OF PROPOSED FRAMEWORK WITH OTHER MODELS.

Method	MEAN JS	MEAN DSC
3D PCN	0.7961	0.8842
3D PCN+LS	0.8035	0.8886
Proposed Framework	0.8217	0.9005

Table 2, represent the performance of coronary arteries segmentation is evaluated by the Jaccard index (JI) and Dice similarity coefficient (DSC) scores as shown in Equations 4 and 5.

$$JI = \frac{|Y_+ \cap \hat{Y}_+|}{|Y_+| \cup |\hat{Y}_+|} \quad (4)$$

$$DSC = \frac{2|Y_+ \cap \hat{Y}_+|}{|Y_+| + |\hat{Y}_+|} \quad (5)$$

Although the boundary is not smooth, the deep learning-based method gives a good initial skeleton of coronary arteries. As a result, we employ the level set approach to improve the outcomes, and the final Mean JI and Mean DSC of our framework are, respectively, 0.8217 and 0.9005. In order to intuitively assess the segmentation results and their consistency, we recreate the final segmentation results into three dimensions using the marching cube algorithm. Figure 6 depicts the viewing of the 3D reconstruction. The 3D reconstruction of the coronary arteries can help the doctor locate artery stenosis and plaque, which is essential to determining the presence of coronary heart disease.

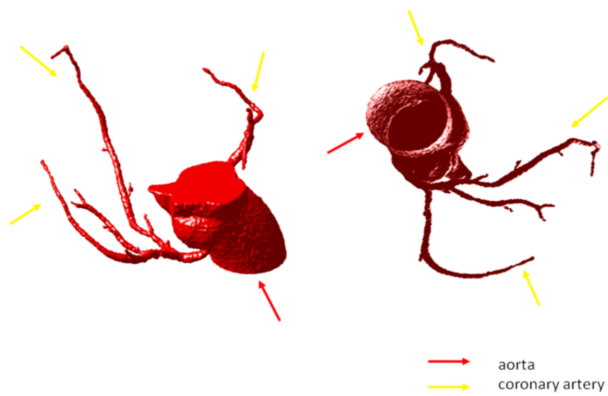

Fig. 6. 3D visualization of Coronary segmentation

V. CONCLUSION

In this paper, we present a novel two-stage fully convolutional neural network framework for the precise segmentation of coronary arteries. Our approach first acquires the global position of coronary arteries through a 3D U-Net-based coarse segmentation. Subsequently, it refines the segmentation by generating three-dimensional slices based on the initial results and applies a three-dimensional Attention Gate U-Net for the final segmentation. Comparative analysis with conventional direct segmentation models highlights the superior performance of our two-stage approach. Notably, the proposed framework achieves a MEAN Jaccard Similarity of 0.8217 and a MEAN Dice Similarity Coefficient of 0.9005, underscoring its efficacy in coronary artery segmentation.

REFERENCES

[1] Sheng-Shou, H., *Report on cardiovascular health and diseases in china 2021: an updated summary.* Journal of Geriatric Cardiology, 2023. **20**: p. 1-32.

[2] Matsushita, K., et al., Epidemiology and risk of cardiovascular disease in populations with chronic kidney disease. Nature Reviews Nephrology, 2022. 18(11): p. 696-707.

[3] Yan, Q., et al., Targeting oxidative stress as a preventive and therapeutic approach for cardiovascular disease. Journal of Translational Medicine, 2023. 21(1): p. 1-35.

[4] You, J.-R., et al., Exploring Possible Diagnostic Precancerous Biomarkers for Oral Submucous Fibrosis: A Narrative Review. Cancers, 2023. 15(19): p. 4812.

[5] Liu, Z., X. He, and Y. Lu, Combining UNet 3+ and transformer for left ventricle segmentation via signed distance and focal loss. Applied Sciences, 2022. 12(18): p. 9208.

[6] Bukhari, N., et al., *Deep learning based framework for emotion recognition using facial expression.* Pakistan Journal of Engineering and Technology, 2022. **5**(3): p. 51-57.

[7] Cai, R., et al., *CTA analysis of 482 cases of coronary artery fistula: A large-scale imaging study.* Journal of Cardiac Surgery, 2022. **37**(7): p. 2172-2181.

[8] Zeng, A., et al. *ImageALCAPA: A 3D Computed Tomography Image Dataset for Automatic Segmentation of Anomalous Left Coronary Artery from Pulmonary Artery*. in *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. 2022. IEEE.

[9] Zhu, H., et al., *Segmentation of coronary arteries images using spatio-temporal feature fusion network with combo loss.* Cardiovascular Engineering and Technology, 2022: p. 1-12.

[10] Zhou, L., et al., *3D slicer combined with neuroendoscope in treatment of a distal segment aneurysm of the anterior choroidal artery complicated intraventricular hemorrhage: A case report and literature review.* Heliyon, 2023. **9**(6).

[11] Hussain, S., et al., *IoT and deep learning based approach for rapid screening and face mask detection for infection spread control of COVID-19.* Applied Sciences, 2021. **11**(8): p. 3495.

[12] Ayoub, M., et al., *A predictive machine learning and deep learning approach on agriculture datasets for new moringa oleifera varieties prediction.* Pakistan Journal of Engineering and Technology, 2022. **5**(1): p. 68-77.

[13] Sahar, A., et al., *Transfer learning-based framework for sentiment classification of cosmetics products reviews.* Pakistan Journal of Engineering and Technology, 2022. **5**(3): p. 38-43.

[14] Ayoub, M., et al., *End to end vision transformer architecture for brain stroke assessment based on multi-slice classification and localization using computed tomography.* Computerized Medical Imaging and Graphics, 2023. **109**: p. 102294.

[15] Gupta, S., P. Gupta, and V.S. Verma, *Study on anatomical and functional medical image registration methods.* Neurocomputing, 2021. **452**: p. 534-548.

[16] Zhao, G., et al., *Bilateral U-Net semantic segmentation with spatial attention mechanism.* CAAI Transactions on Intelligence Technology, 2023. **8**(2): p. 297-307.

[17] Yan, Q., et al., *Attention-guided deep neural network with multi-scale feature fusion for liver vessel segmentation.* IEEE Journal of Biomedical and Health Informatics, 2020. **25**(7): p. 2629-2642.

[18] Hong, P., et al., *A U-Shaped Network Based on Multi-level Feature and Dual-Attention Coordination Mechanism for Coronary Artery Segmentation of CCTA Images.* Cardiovascular Engineering and Technology, 2023: p. 1-13.

[19] Zhang, X., et al., *Beyond vision: A multimodal recurrent attention convolutional neural network for unified image aesthetic prediction tasks.* IEEE Transactions on Multimedia, 2020. **23**: p. 611-623.

[20] Hussain, S., et al., *Ensemble Deep Learning Framework for Situational Aspects-Based Annotation and Classification of International Student's Tweets during COVID-19.* Computers, Materials & Continua, 2023. **75**(3).

[21] Yu, Y., et al., *Pedestrian counting based on piezoelectric vibration sensor.* Applied Sciences, 2022. **12**(4): p. 1920.

[22] Wahid, J.A., et al., *Identifying and characterizing the propagation scale of COVID-19 situational information on Twitter: A hybrid text analytic approach.* Applied Sciences, 2021. **11**(14): p. 6526.

[23] Irfan, L., et al., *A Comparative Analysis of Social Communication Applications using Aspect Based Sentiment Analysis.* Pakistan Journal of Engineering and Technology, 2022. **5**(3): p. 44-50.

[24] An, Y., et al., *High-risk prediction of cardiovascular diseases via attention-based deep neural networks.* IEEE/ACM transactions on computational biology and bioinformatics, 2019. **18**(3): p. 1093-1105.

[25] Wahid, J.A., et al. *Aspect oriented Sentiment classification of COVID-19 twitter data; an enhanced LDA based text analytic approach*. in *2021 International Conference on Computer Engineering and Artificial Intelligence (ICCEAI)*. 2021. IEEE.