Reconstruction of Cardiovascular System in AR/VR Environment using CT scans

Rakshith R
Electronics and Communication Engineering
PES University
Bangalore, India
rakshith2202@gmail.com

Rakshitha Rao HS
Electronics and Communication Engineering
PES University
Bangalore, India
rakshitharaohs@gmail.com

Rahul PG
Electronics and Communication Engineering
PES University
Bangalore, India
rahulpg1682@gmail.com

Samhitha Dwarakanath
Electronics and Communication Engineering
PES University
Bangalore, India
samdwark@gmail.com

HR Vanamala
Electronics and Communication Engineering
PES University
Bangalore, India
vanamalahr@pes.edu

Abstract— **Congenital heart diseases, with their critical and immediate diagnostic needs, pose a significant challenge for medical practitioners. The limitations of traditional 2D visualization methods hinder effective disease detection and comprehension. In response, this paper presents a novel approach to visualize the three-dimensional (3D) model of the heart within an augmented reality/virtual reality (AR/VR) environment, enhancing the understanding of congenital heart diseases. The Image Congenital Heart Disease (CHD) dataset serves as the foundation for heart segmentation, employing the UNET architecture with Resnet34 and Inceptionv3 as backbones. Our analysis demonstrates the superior performance of the UNET architecture with Resnet34 over its counterpart with Inceptionv3. The construction of the 3D heart volume is achieved through the application of Marching Cubes and Ray Casting Algorithms. In the experimental phase, the reconstructed 3D heart volume is visualized using the Oculus Meta Quest 2 headset. This immersive experience allows for a comparative analysis between a healthy heart and one afflicted by disease, providing valuable insights into the intricate details of cardiovascular anomalies. This innovative approach not only addresses the challenges of congenital heart disease visualization but also opens avenues for advanced diagnostic and educational applications in the medical field.**

I. INTRODUCTION

Congenital heart diseases pose a significant threat to individuals from the moment of birth, representing a formidable challenge in the realm of medical conditions. Their inherent presence since birth makes them particularly life-threatening, emphasizing the critical need for early intervention and treatment. Understanding the intricacies of these congenital heart diseases is a complex task for medical practitioners and students alike. The conventional 2D visualization often falls short in capturing the comprehensive nature of these conditions within the heart.

In response to this challenge, our study takes a groundbreaking approach, introducing a three-dimensional (3D) volumetric representation of the heart, coupled with advanced augmented reality (AR) and virtual reality (VR) technologies. By harnessing the capabilities of AR/VR, we aim to transcend the limitations of traditional 2D perspectives, providing a more immersive and detailed visualization of the heart. This innovation holds particular promise for medical practitioners, enabling them to detect congenital heart diseases at their early stages.

The utilization of AR/VR not only enhances the diagnostic capabilities for medical professionals but also serves as a valuable educational tool for students in the field. By offering an immersive experience, our approach enables students to visualize and comprehend the complexities of the heart in ways that surpass the limitations of 2D visuals. This transformative shift in visualization techniques has the potential to revolutionize the understanding and early detection of congenital heart diseases, thereby significantly contributing to improved patient outcomes and advancing medical education.

[1] presents an AR-based interactive tool designed for learning the anatomy of the knee and foot using 3D reconstructions of CT images. The methodology involves extracting the regions of interest from the CT images through automatic segmentation, reconstructing them in 3D, and displaying the resulting models in real-time using AR technology. [2] presents a novel method for reconstructing 3D images of the heart from cardiac MRI images using convolutional neural networks (CNNs). The methodology section describes in detail the proposed approach for 3D reconstruction, which involves pre-processing of the input MRI images, training a CNN model on a large dataset of cardiac MRI images, and using the trained model to generate 3D images of the heart. [3] discusses a technique for interactive volume rendering of large medical datasets in virtual reality, which allows users to explore and manipulate 3D visualizations of medical data. The methodology of this paper involves using a GPU-based algorithm to enable real-time rendering of volumetric data and integrating the rendering technique with a virtual reality headset to create an immersive visualization experience. [4] provides a comprehensive review of 3D visualization techniques for medical images. The review covers a wide range of visualization techniques, including volume rendering, surface rendering, direct volume rendering, and hybrid rendering. [5] presents a feature reconstruction approach for 3D medical image processing. The authors propose a framework for extracting and reconstructing the features of 3D medical

images, which is used to enhance the accuracy of image segmentation, registration, and analysis. The methodology of this study involves the extraction of multi-scale features from 3D medical images using a novel adaptive feature extraction algorithm. The extracted features are then reconstructed using a feature reconstruction algorithm that incorporates prior knowledge of the target features. [6] presents a research work where the authors start by introducing the problems associated with the classical U-Net model, which is typically used for image segmentation .The problem was related to the information loss during downsampling process in the U-Net architecture. The authors made two specific improvements to the U-Net model , first was to change the position of the skip connection and connect the encoder and decoder in a way that retains more image information, second was to add a concatenation operation to the downsampling part of the model to better fuse high and low-level image information. We have used some of the concepts and techniques that have been mentioned in this paper and incorporated in our project to improve our accuracy and enhance segmentation performance. [7] introduces a new method for liver segmentation called DU-Net. This architecture is based on a cascade U-Net , utilizing deep separable convolution and deconvolution techniques. The authors employed the Dice loss function for training, which helps in optimizing the model for achieving accurate segmentation. [8] discusses the difficulties of handling and working with small datasets as they can limit the efficiency of neural networks, and data augmentation becomes crucial in such cases. Twenty 3D CT scans are used which have been taken from the MICCAI 2017 Multi-Modality Whole Heart Segmentation Challenge dataset to assess the suggested method discussed in the paper. For testing purposes, the dataset has been divided into 15 training photos and 5 validation images. The Dice coefficient, which is a measurement of how closely the segmentation matches the real heart structure, is used to compare the segmentation findings with the ground truth. [9] in this paper, deep convolutional networks, especially fully convolutional networks like U-Net have been discussed for demonstrating higher performance. The authors have compared U-Net with different models and submitted a comparison report on the same. A large variety of backbones have also been analyzed, and various conclusions have been drawn as well. [10] compares Neural network architectures of ResNet class include ResNet Depth, ResNet18 and ResNet34. ResNet34 and ResNet18. In terms of evaluating quality, the ResNet34 model is discovered to be superior. ResNet18 uses more RAM, even though it trains more quickly. This demonstrates how memory utilization and training speed are traded off. [11] focuses on difficulty of effectively viewing very big 3D datasets is discussed in this paper, with a focus on biomedical image processing. The size of the data exceeds the capacity of the computer, making traditional techniques of presenting extremely enormous datasets directly in the computer's memory wasteful. [12] explores volume rendering algorithms, which are methods for visualizing three-dimensional data in a variety of areas, including geography, simulation, and medicine. In the paper,several techniques have been used in volume rendering, each with unique choices and complexity. This paper also examines a number of rendering methods, such as splatting, shear-wrapping, and Ray Casting. The authors evaluate the benefits and drawbacks of each and every technique in relation to image quality and algorithm complexity. [13] the analysis of patient data acquired using several contemporary imaging modalities, such as ultrasound (US), positron emission tomography (PET), magnetic resonance tomography (MRT), and computerized tomography (CT), is discussed. The importance of computed tomography (CT), one of the latest imaging methods in medical image processing, is highlighted in the research. [14] mainly talks about the congenital heart diseases .The research paper improves on a method of modeling created especially for using MR images for evaluating biventricular (left and right ventricular) function in patients with congenital heart disease. It also addresses the challenges involved in identifying congenital heart disease, which affects 75 out of 1,000 newborns and is the most common birth defect. Over ninety percent of adults with congenital heart disease survive into adulthood thanks to advances in medical treatments, and they still need regular imaging for diagnosis. Because of the heart's complicated geometry and variety of shapes, analyzing these images is challenging. [15] provides an automated method for effectively evaluating the huge amount of image data generated by computed tomography (CT) scanners during cardiac exams. The goal is to automatically segment 3D CT pictures into the heart's four chambers, myocardium, and major vessels. The research paper also presents a model-based method for heart segmentation. By slowly increasing the number of degrees of freedom for allowed deformations, the technique improves segmentation accuracy as well as resolution. [16] evaluates and compares four volume rendering algorithms that have become popular for rendering datasets described on uniform rectilinear grids: raycasting, splatting, shear-warp, and hardware-assisted 3D texture-mapping. [17] provides a quick and efficient technique for recognizing the whole heart region from dynamic CT scans is given in this work. The proposed approach integrates several approaches to produce precise segmentation. To attempt to decrease connectivity in dynamic CT images between the heart and its surroundings, the research presents a structural repetitive disintegration method. [18] discusses the use of an improved three-dimensional U-net convolutional neural network for segmenting coronary arteries in medical imaging. It utilizes deep neural networks for feature extraction and regression modeling. The research emphasizes how important accurate algorithms are for the diagnosis and management of coronary artery diseases. Multimodal optimization techniques, neural networks, U-net network topology, and deep learning are some of the technologies that are employed.

With the literature survey performed, we gained a lot of insight from different research papers. After careful examination, we have come up with a tool which will help the medical practitioners and students pursuing medicine understand the intricacies of the heart in an immersive environment. This section meets the introduction and literature available till now. The remaining of the document is organized as follows: Section 2 delves into the preprocessing of the dataset used in the research. Section 3 is dedicated to segmentation of heart using UNET architecture and analyzing the results of different backbones used along with the UNET architecture. Section 4 dives into the 3D volume reconstruction of the heart using Ray Casting and Marching Cubes Algorithms along with the visualization of the 3D reconstructed model in AR. Finally, section 5 concludes our work.

II. PREPROCESSING OF DATASET

In this study, the Image CHD dataset served as a pivotal resource, encompassing 110 3D Computed Tomography (CT) scans that comprehensively cover various types of Congenital Heart Diseases (CHD). This dataset, notably larger than many existing medical imaging datasets, consists of CT scans, each comprising a volume and a corresponding mask. The dataset encompasses both healthy and diseased hearts. The preprocessing of the dataset involved extracting 2D slices as PNG files from the 3D CT scans, resulting in a total of 17,364 samples. Each sample consists of a 512x512 volume image and an associated mask image. During the review of mask images, it was identified that there are seven labels corresponding to distinct parts of the heart. The labels and their corresponding intensity values are as follows: Pulmonary Artery (1), Aorta (2), Myocardium (3), Right Atrium (4), Left Atrium (5), Left Ventricle (6), and Right Ventricle (7). The background is represented by an intensity value of 0. To streamline the dataset, additional intensity values (9, 10, 11, 12, 13, and 14) representing airways were set to 0, ensuring that only the seven specified intensity values remained, aligning with the seven different structures of the heart. The model was trained on 12427 samples, validated on 3115 samples and tested on 1822 samples

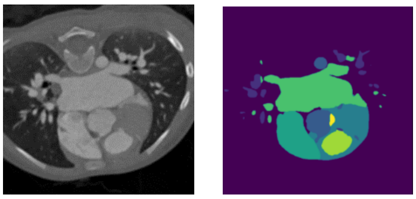


Fig. 1. 2D volume image (left) and its corresponding mask image (right)

III. UNET- BASED HEART SEGMENTATION

This section details of various backbones used along with the UNET architecture for heart segmentation. InceptionV3 and Resnet34 were used as backbones to perform 8-class heart segmentation. The results for each backbone are also included in this section

A. *Training results for Resnet34 as the backbone*

The UNET architecture, employing a Resnet34 backbone, underwent training on 12,427 samples with validation conducted on 3,115 samples. The training process utilized a learning rate of 0.001, and employed the softmax function as the activation function. The model's evaluation centered on key metrics, namely the Intersection over Union score (IoU) and F1-score, providing a comprehensive assessment of segmentation performance. Adams optimizer and categorical cross entropy are used during the training of the model.

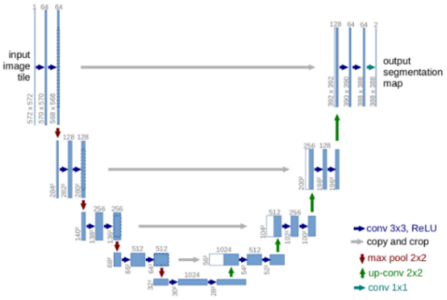


Fig. 2. UNET Architecture(example for 32x32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The White boxes represent copied feature maps. The arrows denote the different operations

The equation to calculate the Softmax function is depicted in equation (1)

$$S(y)_i = \frac{\exp(y_i)}{\sum_{j=1}^{n} \exp(y_j)}$$

Intersection over union is the main metric used to evaluate the accuracy of heart segmentation. IoU measures how well the predicted image overlaps with the ground truth image. IoU ranges from 0 to 1. Fig 3 represents the mathematical represent of IoU score

$$Intersection\ over\ Union\ \left(IoU\right) = \frac{|\mathrm{A} \cap \mathrm{B}|}{|\mathrm{A} \cup \mathrm{B}|}$$

Fig. 3. In the above formula, |A ∩ B| *represents the Area of overlap and* |A U B| *represents the Area of union.*

The model was run for 50 epochs with a batch size of 8. Callback functions such as Early Stopping were also used. The corresponding training and validation IoU scores and training and validation F1 scores are depicted in figures below.

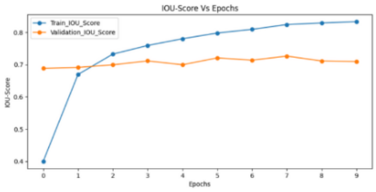


Fig. 4. Varying Training IoU score is represented by the blue line and varying validation IoU score is represented by the orange line for Resnet34 as the backbone.

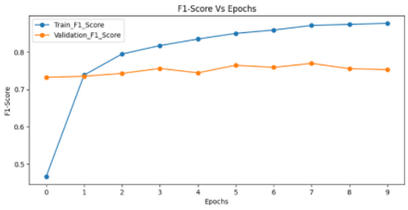


Fig. 5. Varying Training F1 score is represented by the blue line and varying validation F1 score is represented by the orange line for Resnet34 as the backbone

B. Training results for Inceptionv3 as the backbone

The UNET architecture, incorporating Inceptionv3 as the backbone, was trained on a dataset comprising 12,427 samples, with an additional 3,115 samples used for validation. Employing a learning rate of 0.001, the model underwent a 50-epoch training regimen, utilizing the softmax function as the activation function. Evaluation metrics, specifically the Intersection over Union score (IoU) and F1-score, were employed to gauge the effectiveness of the segmentation achieved by the model. Adams optimizer and categorical cross entropy are used during the training of the model.

The model was run for 50 epochs with a batch size of 8. Callback functions such as Early Stopping were also used. The corresponding training and validation IoU scores and training and validation F1 scores are depicted in figures below.

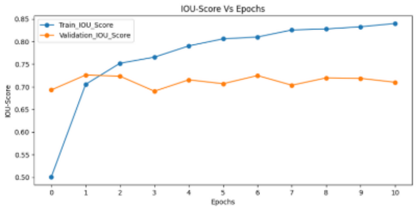


Fig. 6. *Varying Training IoU is represented by the blue line and varying validation IoU is represented by the orange line for Inceptionv3 as the backbone*

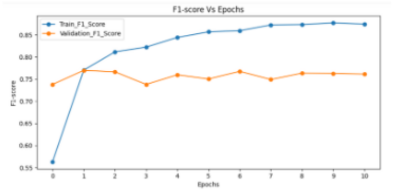


Fig. 7. Varying Training IoU is represented by the blue line and varying validation IoU is represented by the orange line for Inceptionv3 as the backbone

C. Segmentation Results for Resnet34 as the backbone

The model was initially set to run for 50 epochs, but due to early stopping, the model stopped after running for 10 epochs. Out of these 10 epochs, the best weights were chosen to run on the test dataset. The test dataset comprises of 1822 samples. A selection of sample prediction results are shown in Fig. 8 and Fig. 9.

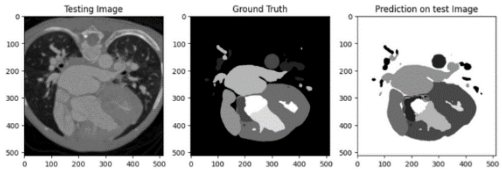


Fig. 8. The figure shows sample 1 prediction using Resnet34 as the backbone. This figure has 3 components, testing image (leftmost), ground truth (middle) and prediction on the test image (rightmost)

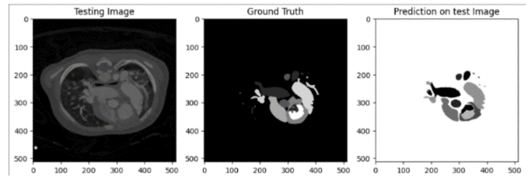


Fig. 9. The figure shows sample 1 prediction using Resnet34 as the backbone. This figure has 3 components, testing image (leftmost), ground truth (middle) and prediction on the test image (rightmost)

D. Segmentation Results for Inceptionv3 as the backbone

Initially configured for 50 epochs, the model underwent early stopping, concluding after only 10 epochs. The best weights, derived from these 10 epochs, were selected for testing on a dataset containing 1,822 samples. A representative display of sample predictions is showcased in Fig. 10 and Fig. 11

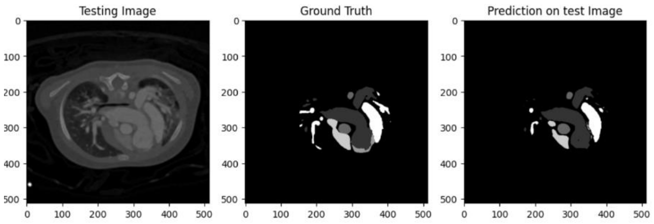


Fig. 10. The figure shows sample 1 prediction using Inceptionv3 as the backbone. This figure has 3 components, testing image (leftmost), ground truth (middle) and prediction on the test image (rightmost)

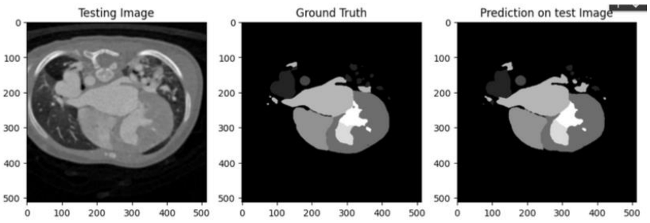


Fig. 11. The figure shows sample 2 prediction using Inceptionv3 as the backbone. This figure has 3 components, testing image (leftmost), ground truth (middle) and prediction on the test image (rightmost)

E. Models Comparison

After obtaining the best weights for each backbone, the model was testing a dataset comprising of 1822 samples. The Table 1 shows the models comparison based on their Testing IoU score and Testing F1 score.

Table 1. Comparative analysis of the two backbones used

Backbone	Testing IOU-Score	Testing F1-Score
Resnet34	0.72545	0.77234
Inceptionv3	0.72268	0.76988

From the following analysis, we concluded that, U-Net architecture with Resnet34 backbone is the best performing model with a Testing IOU-Score of 0.72545 and Testing F1-Score 0.77234.

IV. VOLUME RENDERING AND VISUALIZATION OF THE RECONSTRUCTED HEART IN AR/VR

A. Volume Rendering

After choosing the best backbone for the UNET architecture, We now predicted the mask image for every volume image present in each CT scan. A 3D volume of the heart was constructed using Ray casting algorithm as shown in Fig. 14. This was implemented using the VTK library. Once the 3D volume of the heart was generated, it was exported onto a live render window. RGB channels were added to the 3D volume, which made the volume visually appealing. Fig. 12 And Fig. 13 show the faces of the heart rendered in 3D using VTK library.

Fig. 12. Shows the frontal view of the heart rendered in 3D

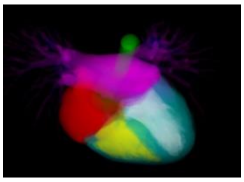
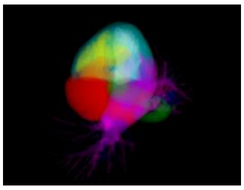


Fig. 13. Shows one of the faces Of the heart rendered in 3D



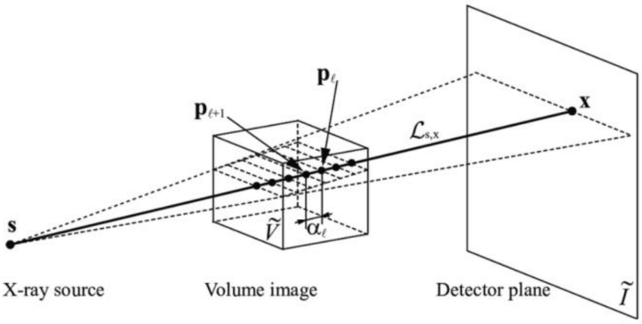


Fig. 14. This figure provides an overview on Ray casting algorithm to compute the simulated projection.

The render window can only be viewed and cannot be exported to any platform in any format as it is provides a real time view. To overcome this, we converted the 2D slices of each CT to a nifty (nii) file. We then used Marching cubes algorithm to create a 3D volume of the heart. The 3D volume of the heart was exported as a Standard triangle language (STL) format. Marching Cubes is an algorithm used in computer graphics to create a polygonal surface mesh from a three-dimensional scalar field. The workflow of the Marching Cubes algorithm is shown in Fig. 15.

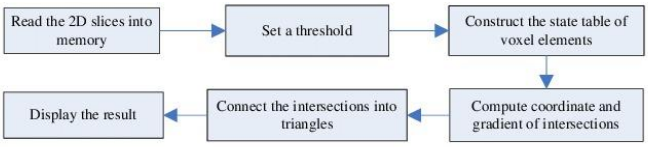


Fig. 15. The figures shows the working procedure of Marching Cubes algorithm

B. Visualization of the Reconstructed heart in AR/VR

The STL file consisting of the reconstructed heart, was imported to a platform called Sketchfab. Sketchfab is a open-source platform to visualize 3D models. Meta Oculus is used as the headset, which provides a complete immersive experience. Fig. 16 and Fig. 17 show one of the faces of the heart visualized in VR

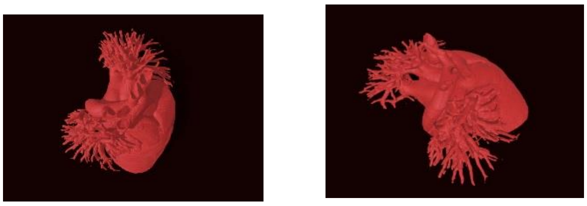


Fig. 16 and Fig. 17 show the different faces of the heart visualized in VR

Fig. 18 consists of 3D volume of the heart which has a CHD by the name Double Outlet Right Ventricle (DORV), which can be compared with Fig. 19 which displays a healthy heart

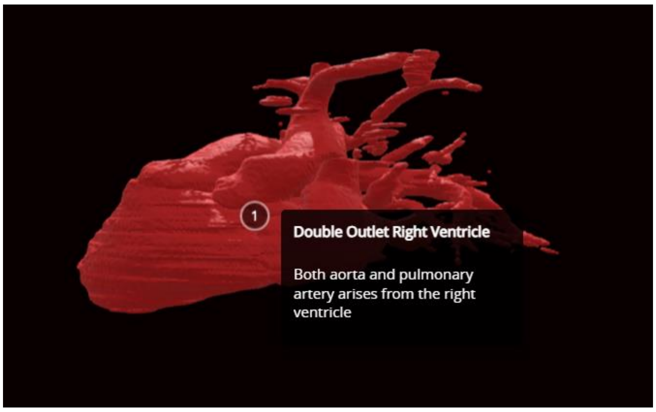


Fig. 18. This figure shows a heart with a very fatal CHD by the name Double Outlet Right Ventricle. DORV is condition when both Aorta and pulmonary artery arises from the right ventricle

Fig. 19. The figure shows a healthy heart.

V. CONCLUSION AND FUTURE SCOPE

We have developed a tool that employs CT scans for reconstructing and displaying the cardiovascular system within AR/VR environment. Heart segmentation was executed using a T4 GPU on Google Colab Pro for efficient computational processing. This offers a comprehensive and immersive experience tailored for both medical professionals and students in detecting and identifying different congenital diseases present in the heart. Our research project contributes to a deeper understanding of the heart, aiding in the identification of diseased areas and enabling a clear distinction between a healthy and an affected heart. In the future, we aspire to assist medical practitioners in meticulous preoperative planning for heart surgeries, enhancing surgical precision and overall patient care. We see our technology as a key component of careful preoperative preparation, giving surgeons comprehensive knowledge about the cardiovascular architecture of their patients. Our research project currently works at surface level and there are several fatal diseases that lie in the interiors of the cardiovascular system which need to be analyzed as well. We also aspire to facilitate the examination of cross-sectional views of the heart for the detection and observation of internal ailments and congenital diseases. This advancement is intended to provide greater benefits to medical students and practitioners compared to what our current research offers

References

[1] L. S. Lee, S. A. Z. S. Aluwee, G. C. Meng, P. Palanisamy and R. Subramaniam, "Interactive Tool Using Augmented Reality (AR) for Learning Knee and Foot Anatomy Based on CT Images 3D Reconstruction," 2020 International Conference on Computational Intelligence (ICCI), Bandar Seri Iskandar, Malaysia, 2020, pp. 281-286, doi: 10.1109/ICCI51257.2020.9247820.

[2] Wang, L., et al. "3D Reconstruction of Cardiac MRI Images Using Convolutional Neural Networks." IEEE Transactions on Medical Imaging, vol. 39, no. 2, 2020, pp. 421-431. doi: 10.1109/TMI.2019.2928084.

[3] Khan, M. N., et al. "Real-Time Interactive 3D Visualization of Medical Images Using OpenVR." IEEE Transactions on Visualization and Computer Graphics, vol. 26, no. 1, 2020, pp. 669-679. doi: 10.1109/TVCG.2019.2934171.

[4] Yao, T., Huang, H., Zhang, W., & Zhang, Y. (2015). "A review of 3D visualization techniques for medical images". IEEE Reviews in Biomedical Engineering, 8, 4-16. doi: 10.1109/RBME.2014.2359546.

[5] C. -S. Wang, M. -C. Lin, C. -C. Wang, C. -F. Chen and J. -C. Hsieh, "Feature reconstruction for 3D medical images processing," 2013 6th International Conference on Biomedical Engineering and Informatics, Hangzhou, China, 2013, pp. 69-74, doi: 10.1109/BMEI.2013.6746909.

[6] J. Di, S. Ma, J. Lian and G. Wang, "A U-Net Network Model for Medical Image Segmentation Based on Improved Skip Connections," 2022 14th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), Changsha, China, 2022, pp. 298-302, doi: 10.1109/ICMTMA54903.2022.00064.

[7]C. -L. D. Ye and Y. Li, "An Improved U-Net Method with Deep Feature for Liver Segmentation in CT Images," 2021 4th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE), Changsha, China, 2021, pp. 759-763, doi: 10.1109/AEMCSE51986.2021.00157.

[8]M. Habijan, H. Leventić, I. Galić and D. Babin, "Whole Heart Segmentation from CT images Using 3D U-Net architecture," 2019 International Conference on Systems, Signals and Image Processing (IWSSIP), Osijek, Croatia, 2019, pp. 121-126, doi: 10.1109/IWSSIP.2019.8787253.

[9]D. G. Kumar and S. Chaudhari, "Comparison of Deep Learning Backbone Frameworks for Remote Sensing Image Classification," IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 2022, pp. 7763-7766, doi: 10.1109/IGARSS46834.2022.9883153.

[10]X. Meng, L. Han, C. Zuo and P. Tao, "Resnet-based quality evaluation of image mosaic and training data balance," 2021 3rd International Conference on Advances in Computer Technology, Information Science and Communication (CTISC), Shanghai, China, 2021, pp. 333-340, doi: 10.1109/CTISC52352.2021.00067.

[11]Jian Xue, K. Lü and Jie Tian, "An efficient out-of-core volume rendering method based on ray casting and GPU acceleration," 2009 IEEE Youth Conference on Information, Computing and Telecommunication, Beijing, China, 2009, pp. 130-133, doi: 10.1109/YCICT.2009.5382411.

[12]R. Dubey, S. Jain and R. S. Jadon, "Volume Rendering: A Compelling Approach to Enhance the Rendering Methodology," 2016 Second International Conference on Computational Intelligence & Communication Technology (CICT), Ghaziabad, India, 2016, pp. 712-717, doi: 10.1109/CICT.2016.147.

[13] F. Ritter et al., "Medical Image Analysis," in IEEE Pulse, vol. 2, no. 6, pp. 60-70, Nov.-Dec. 2011, doi: 10.1109/MPUL.2011.942929.

[14] K. Gilbert et al., "Improving assessment of congenital heart disease through rapid patient specific modeling," 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 2016, pp. 1228-1231, doi: 10.1109/EMBC.2016.7590927.

[15] O. Ecabert et al., "Automatic Model-Based Segmentation of the Heart in CT Images," in IEEE Transactions on Medical Imaging, vol. 27, no. 9, pp. 1189-1201, Sept. 2008, doi: 10.1109/TMI.2008.918330.

[16] J. Huang, K. Mueller, R. Crawfis, D. Bartz and M. Meissner, "A Practical Evaluation of Popular Volume Rendering Algorithms," 2000 IEEE Symposium on Volume Visualization (VV 2000), Salt Lake City, UT, USA, 2000, pp. 81-90, doi: 10.1109/VV.2000.10009.

[17] Lixu Gu, "Hybrid 3D heart segmentation from dynamic CT images," 2004 2nd IEEE/EMBS International Summer School on Medical Devices and Biosensors, Hong Kong, China, 2004, pp. 67-70, doi: 10.1109/ISSMD.2004.1689562.

[18] Lorensen, W. E., & Cline, H. E. (1987). Marching cubes: A high resolution 3D surface construction algorithm. Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH ’87. doi:10.1145/37401.37422