



Structure-Aware V-Net Framework for 3D Multi-Class Segmentation of Pediatric Cardiovascular Structures

Irem Yilmaz
Biomedical Engineering
Yildiz Technical University
Istanbul, Turkey
irem.yilmaz@std.yildiz.edu.tr

Omer Faruk Sahin
Bioengineering
Marmara University
Istanbul, Turkey
umarfaruksahin@outlook.com

Mohamad Humam Alzaeim
Biomedical Engineering &
Bioinformatics
Istanbul Medipol University
Istanbul, Turkey
mohamad.Alzaeim@std.medipol.edu.tr

Beyza Nur Yazici
Computer Engineering
Istinye University
Istanbul, Turkey
beyzanur.yazici@stu.istinye.edu.tr

Doruk Gunveren
Cemberlitas Anatolian High School
Istanbul, Turkey
gunverendoruk@gmail.com

Sakip Onder
Biomedical Engineering
Yildiz Technical University
Istanbul Turkey
sonder@yildiz.edu.tr

Kevser Banu Kose
Biomedical Engineering
Istanbul Medipol University
Istanbul Turkey
kbkose@medipol.edu.tr

Fatma Zeynep Gungoren
Radiology Department
The Anadolu Medical
Center
Istanbul, Turkey
fzgungoren@yahoo.com

Musab Ahmed Khan Umair
Artificial Intelligence
Istinye University
Istanbul, Turkey
musabkhan2003@gmail.com

Senol Piskin*
Mechanical Engineering
Istinye University
Istanbul, Turkey
*Corresponding author:
senol.Piskin@istinye.edu.tr

JS Femilda Josephin
Computer Engineering
Istinye University
Istanbul, Turkey
femilda.bai@istinye.edu.tr

Ruqaiyah Mirza
Software Engineering
Istinye University
Istanbul, Turkey
ruqaiyah.mirza@stu.istinye.edu.tr

Fatmanur Aydogdi
Zuhtu Kurtulmus Anatolian High
School
Istanbul, Turkey
nurraydogdi21@gmail.com

Abstract—This research shows a structure-aware deep learning model. The model segments 3D multi-class pediatric cardiovascular anatomy, focusing on small complex structures, like the pulmonary arteries, also it is based upon a modified V-Net architecture. The multi-center dataset includes expert-annotated pediatric Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) volumes. The model underwent training through utilization of this specific dataset. For resolving class disequilibrium in conjunction with morphological variance, investigators incorporated a supplementary decoder division for specialization within pulmonary artery segmentation. The pipeline represents standardized preprocessing measures like resampling, normalization, also cropping, furthermore it represents advanced augmentation techniques. Evaluation was performed using Dice Similarity Coefficient (DSC) and Intersection over Union (IoU) metrics. This assessment revealed that the partitioning of minute anatomy progressed noticeably. The proposed architecture shows support is offered for preoperative planning in congenital heart disease (CHD) cases because structure-specific modifications are able to improve segmentation accuracy plus consistency in pediatric cardiovascular imaging.

Keywords—V-Net; pediatric cardiovascular segmentation; 3D medical imaging; pulmonary artery; deep learning; class imbalance

I. INTRODUCTION

Congenital heart diseases (CHDs) exist among the most typical birth anomalies, so pediatric morbidity as well as mortality are largely triggered [1]. Although CT and MRI provide detailed views of cardiovascular anatomy, manual 3D segmentation is time-consuming and subject to inter-observer variability [2]. Deep learning models,

notably V-Net, are frequently utilized for segmenting volumetric tasks [3][4]. However, conventional architecture struggles to segment small, detailed regions such as the pulmonary arteries due to class imbalance and thin anatomical structure [5].

The proposed study puts forth an adjusted V-Net architecture for pulmonary artery segmentation using a supplementary decoder. The model performs multi-class segmentation and structure-specific segmentation. It was trained on a collected, multi-center dataset including pediatric CT and MRI scans. An effective preprocessing pipeline was utilized, including resampling, normalization, cropping, and TorchIO-centered augmentation [6]. Assessment using Dice and IoU metrics affirmed performance in augmented reality. This became improved in a way that was most obvious in smaller anatomical regions. The structure-aware design improved segmentation accuracy. For CHD diagnosis along with surgical planning, benefit could come from it as well.

II. BACKGROUND

A. Challenges in Pediatric Cardiovascular Imaging

Pediatric cardiovascular structures remain arduous for one to visualize because of diminutive organ sizes, elevated heart rates, and also intrinsic anatomical variance. Segmentation is necessary for effective surgical planning and reliable diagnosis. In CHD cases, segmentation that is accurate is critical. Pediatric imaging is, however, often obstructed by motion artifacts, low signal-to-noise ratios, and limited contrast, making it difficult to delineate small structures such as the pulmonary arteries [7].



B. Limitations of Conventional Segmentation Techniques

Manual segmentation by clinicians is common yet untenable for large datasets and time-consuming. Since conventional techniques often fail for complex pediatric anatomy, inaccurate boundaries and distorted segmentation happen in low-contrast regions [8].

C. V-Net for 3D Medical Image Segmentation

V-Net like networks have exhibited strong performance during 3D medical image segmentation. V-Net conserves spatial detail via skip connections coupled with residual blocks, intended for volumetric data. Contemporary augmentations have exhibited refined precision within juvenile cardiac data sets [9]. They represent attention gates and expanded convolutions.

D. Architectural Solutions to Class Imbalance

Segmentation models often grapple with class imbalance; wherein larger structures dominate training while smaller pulmonary arteries are underrepresented. Refining model susceptibility to these underrepresented classifications without impairing overall efficacy has demonstrated merit through the use of structural models [10].

III. METHOD

Data compilation officially began once the institutional ethics committee approved this study's procedure. The completed procedures are described in each of the following sections.

A. Dataset and Cohort Description

The dataset used in the proposed work comprises anonymized pediatric cardiovascular cases obtained from four medical centers and also included contrast-improved CT and cardiac MRI scans that represented various CHD phenotypes. Each volume was annotated by at least two expert radiologists. Therefore multi-class ground truth masks for the aorta and heart chambers and pulmonary arteries were generated. Preparation of uniforms occurred because the scan settings differed.

B. Preprocessing and Data Augmentation

Consistent input was ensured via cardiac region truncation and volume resampling to $1 \times 1 \times 1$ mm³ voxel spacing. Additionally, volumes did undergo intensity standardization to have a zero mean and unit variance and resizing of them to $128 \times 128 \times 64$ since the researchers had used center cropping or padding. Data were augmented in real-time with random affine transformations, gamma correction, and elastic deformations using TorchIO library. The model gained some extrapolation abilities by way of these enhancements. The model also gained in resilience.

C. Data Anonymization and Management

Adhering to protocols approved by the Institutional Review Board (IRB), patient data were initially anonymized, also data were stored securely under the General Data Protection Regulation (GDPR). Metadata were tracked using a structured database. The database included patient age as well as modality and institution plus segmentation status in addition to preprocessing progress.

D. Labeling Protocol and Auxiliary Masking

Segmentation masks for multiple classes designated each scan in this manner: 0 signified background, 1 signified aorta, 2 signified heart chambers, 3 signified pulmonary arteries. For supervising a specialized decoder segment inside the model, an ancillary binary mask separating the pulmonary artery was fashioned. This helped the model focus on learning small anatomical structures more effectively.

E. 3D Surface Model Generation

SimpleITK and 3D Slicer transform anticipated segmentation results into STL surface models for qualitative assessment utilizing the marching cubes algorithm. These models were visually validated through anatomical correctness and spatial precision by comparing them across sagittal, coronal, and axial planes to the ground truth masks.

IV. NETWORK ARCHITECTURE

To address the challenges intrinsic to segmenting small complex cardiovascular structures along with it, a modified V-Net architecture in pediatric imaging was developed. This model utilizes a three-dimensional encoder-decoder framework, plus it presents four encoding and decoding strata. The architecture strides also transposes convolutions then activates ReLU. It also normalizes the instances also it skips the connections so as to preserve spatial information for an improved gradient flow.

For improvement regarding the partitioning of some pulmonary arteries, they incorporated such a subordinate decoder division. An exclusive Dice loss for the pulmonary artery class supports autonomous optimization of the auxiliary branch. This independent enhancement betters output on obscure areas since it shows a shared encoder with the core decoder. Because it combines Dice with cross-entropy in multi-class segmentation the central decoder employs a hybrid loss. Adam with cosine annealing optimized the training across 1000 epochs. A GPU such as an NVIDIA RTX 3090 was used for the training. Toward improved generalizability, researchers utilized real-time data enrichment through TorchIO that included affine and intensity transformations, and early stopping occurred on validation loss.

V. RESULTS

A. Segmentation Performance

The dual-head V-Net model was trained on a multi-center pediatric dataset over wide-ranging epochs, as it was evaluated through the use of standard segmentation metrics including Intersection over Union (IoU) and Dice Similarity Coefficient (DSC). The results demonstrated strong overall performance especially for background and heart chambers as anatomical structures well-defined and large. On account of its small size in conjunction with class imbalance, the pulmonary arteries were still the most challenging region whereas the aorta showed moderate segmentation accuracy. Importantly, the additional decoder branch performed better at segmenting these underrepresented and morphologically complex structures. The model can improve in consistency and accuracy. These results stress its promise in heart imaging for children.

B. Training and Validation Performance

Training and validation losses gradually converged over the 1000-epoch training cycle. The hybrid loss function fostered an improved convergence speed. This function curtailed overfitting, coupled with structure-specific supervision through the auxiliary decoder. Figure 1 displays the training curves, and it incorporates a heatmap of class-specific segmentation measures that depict uniform efficacy throughout all anatomical targets.

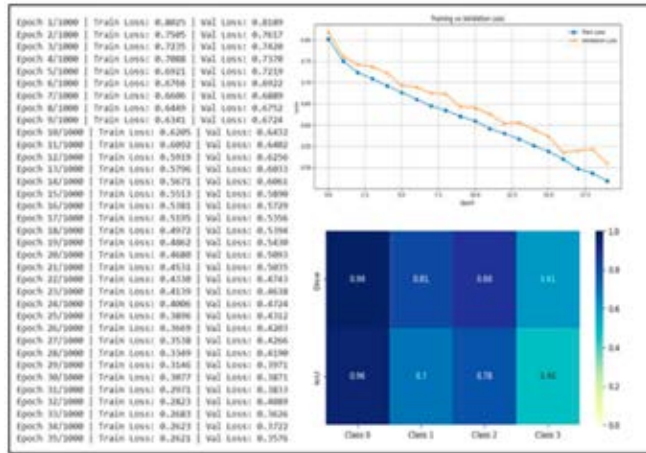


Fig 1. Training vs. validation loss curve and class-wise segmentation heatmap (Dice and IoU) for the multi-class model.

C. Visual and Structure-Specific Assessment

Segmentations were visually evaluated by comparing the predicted and the ground-truth masks in axial, coronal and sagittal planes. This model precisely circumscribed global cardiovascular architecture with smaller conduits like pulmonary arteries as depicted within Figures 2 and 3. Within representative instances, manual annotations displayed large spatial congruence alongside the predicted masks like patient 1 and patient 2. The supplementary decoder helped defining refined borders along with diminishing inaccuracies throughout detailed regions. It proved important for conserving great anatomical detail, which accentuated this contribution.

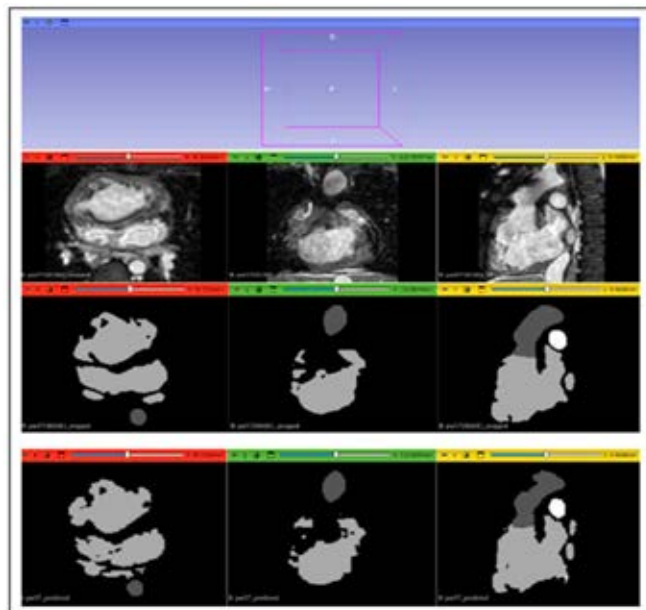


Fig 2. Multi-plane segmentation results for patient 1, showing ground truth and predicted masks.

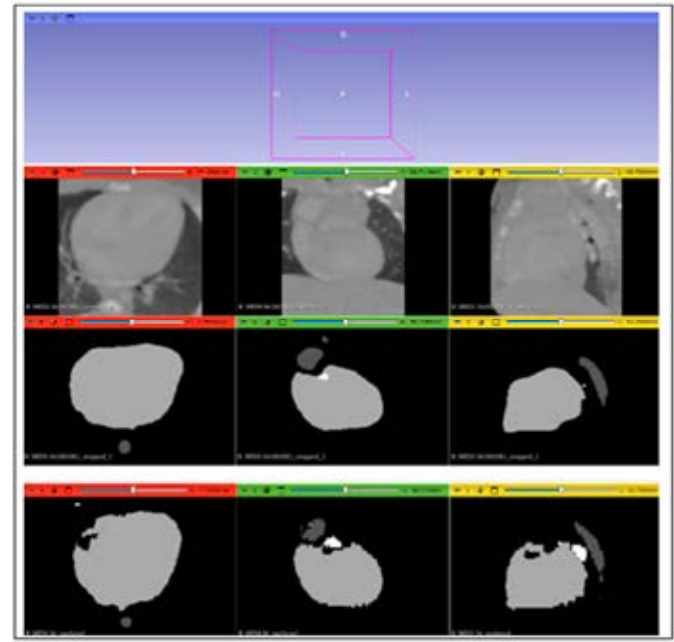


Fig 3. Multi-plane segmentation results for patient 2, showing ground truth and predicted masks.

VI. DISCUSSION AND CONCLUSION

This study introduces a structure-aware V-Net architecture for multi-class segmentation of pediatric cardiovascular structures, in which is tailored with particular emphasis on anatomically small complex regions like the pulmonary arteries. The dual-head design has within it an auxiliary decoder that targets at the pulmonary artery. This design addresses class imbalance as well as anatomical variability since they do present key challenges in pediatric imaging. The framework improves generalization and class-wise accuracy by integrating standardized preprocessing and a hybrid loss function along with TorchIO-based data augmentation.

The results showed consistent IoU and Dice scores for the aorta and heart chambers, which reflects strong segmentation. Pulmonary artery segmentation, which is typically underrepresented in training data, was greatly improved via the auxiliary decoder. Steady convergence over 1000 epochs confirmed effective learning. Because predicted masks showed clear boundaries and structural integrity across all planes, qualitative results further demonstrated anatomical accuracy. Surface reconstructions of 3D confirmed the model could capture vascular morphology precisely.

Through this automated approach, clinical workflows may be supported especially for CHD. Manual labor and inter-observer variability may decrease. Planning before surgery may be improved in fact, even in complex anatomies too. Dataset diversity and annotator variability are current limitations. To address these limits, future work should validate performance via multi-center studies, expand CHD subtypes, and incorporate deeper attention-based architectures.

VII. ACKNOWLEDGMENT

This work was supported by the Scientific and Technological Research Council of Türkiye (TUBITAK) under the 1001 Grant Program (Project No. 122E488).



VIII. AFFILIATIONS

This study was conducted in collaboration with Istinye University, Istanbul Medipol University, and Anadolu Medical Center.

REFERENCES

- [1] Karimi-Bidhendi, S., Arafati, A., Cheng, A. L., Wu, Y., Kheradvar, A., & Jafarkhani, H. (2020). Fully-automated deep-learning segmentation of pediatric cardiovascular magnetic resonance of patients with complex congenital heart diseases. *Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance*, 22(1), 80. <https://doi.org/10.1186/s12968-020-00678-0>
- [2] Shi, F., Wang, J., Shi, J., Wu, Z., Wang, Q., Tang, Z., He, K., Shi, Y., & Shen, D. (2021). Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation, and Diagnosis for COVID-19. *IEEE reviews in biomedical engineering*, 14, 4–15. <https://doi.org/10.1109/RBME.2020.2987975>
- [3] Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature methods*, 18(2), 203–211. <https://doi.org/10.1038/s41592-020-01008-z>
- [4] Ma, X., Song, H., Jia, X., & Wang, Z. (2024). An improved V-Net lung nodule segmentation model based on pixel threshold separation and attention mechanism. *Scientific reports*, 14(1), 4743. <https://doi.org/10.1038/s41598-024-55178-3>
- [5] Cui, L., Xu, M., Liu, C., Liu, T., Yan, X., Zhang, Y., & Yang, X. (2025). Towards Reliable Healthcare Imaging: A Multifaceted Approach in Class Imbalance Handling for Medical Image Segmentation. *Interdisciplinary sciences, computational life sciences*, 17(3), 614–633. <https://doi.org/10.1007/s12539-025-00726-2>
- [6] Pérez-García, F., Sparks, R., & Ourselin, S. (2021). TorchIO: A Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. *Computer methods and programs in biomedicine*, 208, 106236. <https://doi.org/10.1016/j.cmpb.2021.106236>
- [7] Leiner, T., Bogaert, J., Friedrich, M. G., Mohiaddin, R., Muthurangu, V., Myerson, S., Powell, A. J., Raman, S. V., & Pennell, D. J. (2020). SCMR Position Paper (2020) on clinical indications for cardiovascular magnetic resonance. *Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance*, 22(1), 76. <https://doi.org/10.1186/s12968-020-00682-4>
- [8] Chen, X., Wang, X., Zhang, K., Fung, K. M., Thai, T. C., Moore, K., Mannel, R. S., Liu, H., Zheng, B., & Qiu, Y. (2022). Recent advances and clinical applications of deep learning in medical image analysis. *Medical image analysis*, 79, 102444. <https://doi.org/10.1016/j.media.2022.102444>
- [9] Li, H., Li, D., Dong, Q., Han, X., & Dong, S. (2025). Multi-view semi-supervised attention network for 3D cardiac image segmentation. *Frontiers in cardiovascular medicine*, 12, 1461774. <https://doi.org/10.3389/fcvm.2025.1461774>
- [10] Yeung, M., Sala, E., Schönlieb, C. B., & Rundo, L. (2022). Unified Focal loss: Generalising Dice and cross entropy-based losses to handle class imbalanced medical image segmentation. *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society*, 95, 102026. <https://doi.org/10.1016/j.compmedimag.2021.102026>