



A Unified Pipeline for Consistent Multi-Label Mask Generation in Pediatric Cardiac Segmentation

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

[u.tr](mailto:mohamad.alzaeim@std.medipol.edu.tr)

Beyza Nur Yazici

Computer Engineering

Istinye University

Istanbul, Turkey

beyzanur.yazici@stu.istinye.edu.tr

Sakip Onder

Biomedical Engineering

Yildiz Technical University

Istanbul Turkey

sonder@yildiz.edu.tr

Doruk Gunveren

Cemberlitas Anatolian High School

Istanbul, Turkey

gunverendoruk@gmail.com

Senol Piskin*

Mechanical

Engineering Department

Istinye University

Istanbul, Turkey

*Corresponding author:

senol.piskin@istinye.edu.tr

Fatma Zeynep Gungoren

Department of Radiology

The Anadolu Medical Center

Istanbul, Turkey

fzgungoren@yahoo.com

Ruqaiyah Mirza

Software Engineering

Istinye University

Istanbul, Turkey

ruqaiyah.mirza@stu.istinye.edu.tr

Irem Yilmaz

Biomedical Engineering

Yildiz Technical University

Istanbul, Turkey

irem.yilmaz@std.yildiz.edu.tr

Ibrahim Faress

Software Engineering

Istinye University

Istanbul, Turkey

ibrahim.faress@stu.istinye.edu.tr

Kevser Banu Kose

Biomedical

Engineering Department

Istanbul Medipol University

Istanbul Turkey

kbkose@medipol.edu.tr

JS Femilda Josephin

Computer Engineering Department

Istinye University

Istanbul, Turkey

femilda.bai@istinye.edu.tr

Musab Ahmed Khan Umair

Artificial Intelligence

Istinye University

Istanbul, Turkey

musabkhan2003@gmail.com

Fatmanur Aydogdi

Zuhtu Kurtulmus Anatolian High School

Istanbul, Turkey

nurraydogdi21@gmail.com

Abstract— Artificial Intelligence based cardiac image segmentation requires consistent labels, harmonized volumes, and computationally efficient dataset. The present study proposes a reproducible pipeline that (i) merges segmented cardiac structures into a single multi-label mask (0: background, 1: aorta, 2: chambers, 3: pulmonary structures), (ii) resamples masks to perfectly match scan dimensions using 3D Slicer’s “Resample Image (BRAINS)” module (nearest-neighbor interpolation), and (iii) performs automatic region-of-interest (ROI) cropping by adding 10-pixel margin around mask-derived bounds. On our internal dataset, the workflow standardizes inputs, reduces data volume substantially, and preserves full cardiac anatomy. The resulting data enabled stable multi-class and task-specific training and contribution to a better Dice score in downstream models. The ROI script operate on a simple folder structure and can be executed batch-wise, supporting anonymity, interoperability, and rapid experimentation.

Keywords— multi-label segmentation; deep learning; convolutional neural networks (CNNs); image resampling; ROI cropping; data preprocessing; medical image segmentation; pediatric cardiac image; segmentation AI

I. INTRODUCTION

For reliable supervised segmentation in pediatric and complex cardiac Magnetic Resonance Imaging (MRI) and Computed Tomography (CT), the process heavily relies on high-quality and spatially aligned labels. However, model accuracy is often degraded while training complexity often increases because of variations in label encoding, inconsistent mask dimensions, and the presence of irrelevant background voxels. Hence, significant mismatches between image volumes and segmentation masks (e.g., 129-slice scans paired with 646-slice masks), heterogeneous labeling conventions, and extraneous anatomical structures beyond the cardiac region. Now, to address these challenges, a reproducible three-stage preprocessing helps by: (1) generating standardized multi-label masks; (2) harmonizing spatial dimensions via isotropic resampling; and (3) automatically cropping the region of interest (ROI) using margin-expanded mask bounds. This paper details the design, implementation, and benefits of this preprocessing workflow, especially highlighting its impact on data uniformity and segmentation performance across cardiac dataset.

II. RELATED WORK

Classical pipelines often rely on manual re-labeling and ad hoc cropping, which introduce variability and risk reproducibility [1]. Even though automated multi-label strategies exist, many do not enforce strict spatial alignment or incorporate margin-controlled cropping critical for consistent downstream analysis for 3D cardiac data [2]. In this work, established tools such as 3D Slicer and BRAINS Resample [3] were utilized to present a unified, scriptable pipeline that integrates label standardization, precise volumetric alignment, and reproducible cropping into a single automated workflow.

III. DATASETS

Original masks and corresponding scans were generated through in-house segmentations.

IV. METHODS

A. Multi-Label Mask Generation

Initially, existing segmentations were imported into 3D Slicer in a defined order to ensure consistent label assignment and exported as binary labelmaps [3]. A unified mask with fixed integer codes was adopted: 0 (background), 1 (aorta), 2 (chambers), 3 (pulmonaries) as shown in Fig. 1.

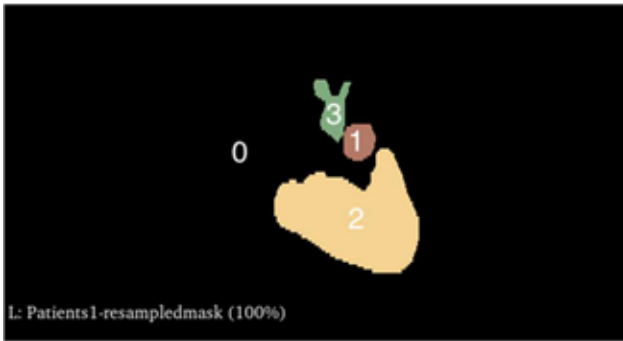


Fig. 1. Label Assignment; 0 = background, 1 = aorta, 2 = cardiac chambers (atria + ventricles), and 3 = pulmonary arteries.

B. Size Harmonization via Resampling

Mismatches in slice count, spacing, and origin between scans and masks are illustrated in Fig. 2. The scan (top) is $512 \times 512 \times 129$ voxels at $0.37 \times 0.37 \times 1$ mm spacing with origin (58, -100, -280 mm) and a flipped IJK→RAS direction, whereas the label volume (bottom) is $465 \times 462 \times 646$ voxels at 0.127 mm isotropic spacing, origin (-78, -226, -261 mm), and a different orientation. These mismatches in geometry must be harmonized to avoid misalignment and were corrected using the “Resample Image (BRAINS)” module within 3D Slicer [3]. Nearest-neighbor interpolation and uchar pixel type were applied to retain integer labels [4]. Masks were aligned to the reference scan geometry. An example as shown in Fig. 3, both scan (top) and label (bottom) now share identical image dimensions, isotropic spacing, origin, and orientation, confirming successful resampling and alignment. This step is essential for correct model training, as misaligned tensors can silently corrupt losses and gradients.

Active Volume: Patients1(SCAN)			
Volume Information			
Image Dimensions:	512	512	129
Image Spacing:	0.3671875mm	0.3671875mm	1.0000000mm
Image Origin:	58.0000mm	-100.0000mm	-279.9000mm
IJK to RAS Direction Matrix:	-1.0000	0.0000	0.0000
	0.0000	-1.0000	0.0000
	0.0000	0.0000	1.0000

Active Volume: Patients1-label			
Volume Information			
Image Dimensions:	465	462	646
Image Spacing:	0.1272mm	0.1272mm	0.1272mm
Image Origin:	-78.2219mm	-226.6791mm	-261.0427mm
IJK to RAS Direction Matrix:	1.0000	0.0000	0.0000
	0.0000	1.0000	0.0000
	0.0000	0.0000	1.0000

Fig. 2. Mismatch in volume geometry metadata between the scan and its label before preprocessing

Active Volume: Patients1-resampledmask			
Volume Information			
Image Dimensions:	512	512	129
Image Spacing:	0.3671875mm	0.3671875mm	1.0000000mm
Image Origin:	58.0000mm	-100.0000mm	-279.9000mm
IJK to RAS Direction Matrix:	-1.0000	0.0000	0.0000
	0.0000	-1.0000	0.0000
	0.0000	0.0000	1.0000

Active Volume: Patients1(SCAN)			
Volume Information			
Image Dimensions:	512	512	129
Image Spacing:	0.3671875mm	0.3671875mm	1.0000000mm
Image Origin:	58.0000mm	-100.0000mm	-279.9000mm
IJK to RAS Direction Matrix:	-1.0000	0.0000	0.0000
	0.0000	-1.0000	0.0000
	0.0000	0.0000	1.0000

Fig. 3. Geometry metadata after preprocessing.

C. ROI Cropping with Margin Preservation

To reduce computational load and focus learning on cardiac anatomy in segmentation [5], min-max coordinates were extracted from masks to define a tight 3D bounding box, then added margin by 10 pixels on all sides to avoid cutting critical structures [6,7]. Both scan-mask pairs were cropped identically. A Python Jupyter Notebook script executes batch cropping based on a simple folder structure (input/, output/, scripts/), enabling reproducible processing for each patient. Fig. 4 illustrates an example scan output, Top: original (non-cropped) scan with extensive background. Bottom: cropped volume centered on the cardiac anatomy with a fixed safety margin, preserving spacing/orientation while removing extraneous voxels to reduce memory use and accelerate training. Similarly, Fig. 5 demonstrates the corresponding ground truth mask, where the cropped volume is focused on the main structures (aorta, chambers, pulmonary arteries), allowing faster learning and keeping input dimensions input consistent among data.

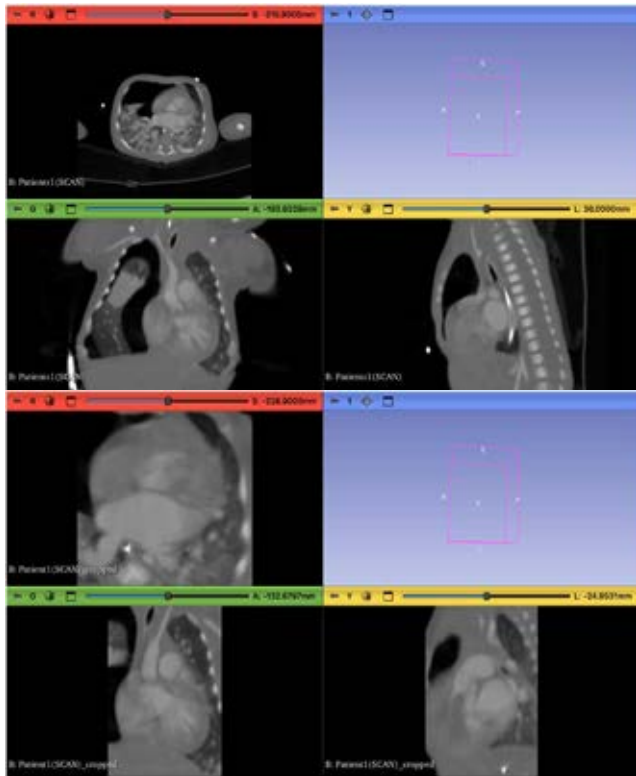


Fig. 4. Field-of-view comparison for compute-aware scan cropping.

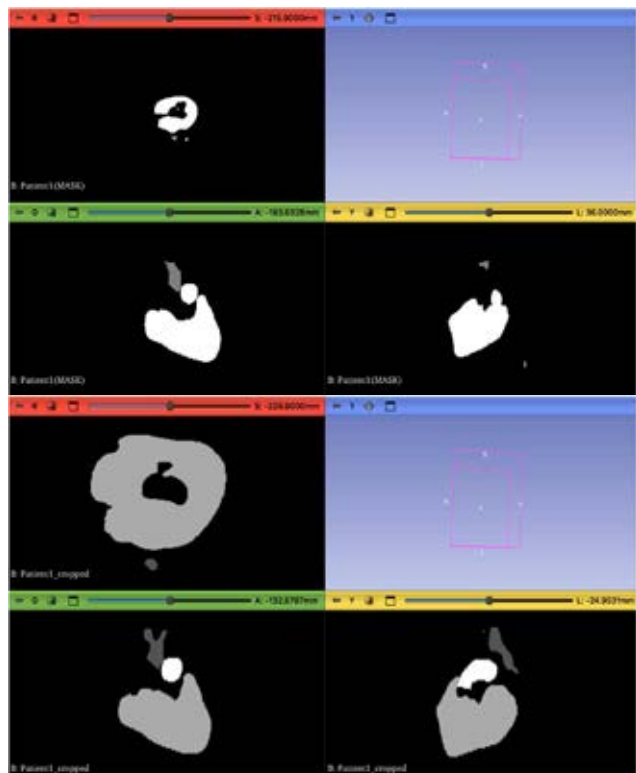


Fig. 5. Effect of cropping on label volumes

V. IMPLEMENTATION DETAILS

Using pseudonymous IDs (e.g., Patient1, Patient2), patient files were organized. Each folder contained a single raw scan and the multi-label mask paired it. The proposed pipeline loops through each folder, preserves the original scan and outputs harmonized and cropped volumes that are prepared for analysis in order to guarantee uniform preparation. The widely utilized scientific packages NumPy to facilitate efficient array manipulation was employed in the implementation, which had been carried out in Python 3.8+.

VI. EXPERIMENTS & RESULT

A. Qualitative Verification

Middle-slice overlays confirmed accurate alignment of post-resampling and cropping. No spatial drift was noted between labels and intensity volumes.

B. Quantitative Effects

Data volume was significantly reduced while retaining all cardiac anatomy, leading to faster training and inference. Although exact voxel count reductions vary by case, the approach consistently trimmed irrelevant body regions.

VII. DISCUSSION

The proposed pipeline closes three common gaps in medical image segmentation preparation: unified label encoding, geometric consistency, and compute-aware cropping. By standardizing these steps, minimized human error and improved repeatability. While effective, the method assumes initial segmentations are anatomically valid; future work could integrate QA metrics (e.g., histogram-based sanity checks) directly into the pipeline. Additionally, explicit logging of voxel-count reductions and time savings would strengthen the cost-benefit analysis.

VIII. CONCLUSION

The overall methodology integrates multi-label mask generation, precise nearest neighbor resampling in 3D Slicer ("Resample Image (BRAINS)", unsigned char (uchar) labels), and ROI cropping with a fixed 10 pixel safety margin into a single, aligned, and reproducible preprocessing pipeline for cardiac MRI segmentation. Across internal scan-mask pairs, the workflow ensured exact label-image alignment, curtailed non-cardiac volume to reduce data size and time, and produced standardized inputs that supported stable, strong performance in both multi-class and task-specific models. These results suggest straightforward transferability to other multi-structure organs; and support adoption as a reproducible preprocessing standard. The method assumes anatomically valid initial segmentations; next steps include integrating fully automated quality assurance checks, systematically logging and profiling voxel count reductions and time/memory usage, and evaluating adaptive margin sizing.



IX. ACKNOWLEDGMENT

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

X. AFFILIATIONS

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

REFERENCES

- [1] F. Renard, S. Guedria, N. D. Palma, and N. Vuillerme, “Variability and reproducibility in deep learning for medical image segmentation,” *Scientific Reports*, vol. 10, no. 1, p. 13724, Aug. 2020, doi: <https://doi.org/10.1038/s41598-020-69920-0>.
- [2] F. Maleki et al., “RIDGE: Reproducibility, Integrity, Dependability, Generalizability, and Efficiency Assessment of Medical Image Segmentation Models,” *Deleted Journal*, Nov. 2024, doi: <https://doi.org/10.1007/s10278-024-01282-9>.
- [3] A. Fedorov et al., “3D slicer as an image computing platform for the Quantitative Imaging Network,” *Magn. Reson. Imaging*, vol. 30, no. 9, pp. 1323–1341, Nov. 2012. doi: 10.1016/j.mri.2012.05.001.
- [4] W. Kim, Y. Kang, S. Lee, H.-J. Lee, and Y. Nam, “Label-Preserving Data Augmentation for Robust Segmentation of Thin Structure in MRI,” *Investigative Magnetic Resonance Imaging*, vol. 28, no. 3, p. 107, 2024, doi: <https://doi.org/10.13104/imri.2024.0014>.
- [5] R.-R. Galea, L. Diosan, A. Andreica, L. Popa, S. Manole, and Z. Bálint, “Region-of-Interest-Based Cardiac Image Segmentation with Deep Learning,” *Applied Sciences*, vol. 11, no. 4, p. 1965, Feb. 2021, doi: <https://doi.org/10.3390/app11041965>.
- [6] I. S. Isler, D. Mohaisen, C. Lisle, D. Turgut, and U. Bagci, “Uncertainty-Guided Coarse-to-Fine Tumor Segmentation with Anatomy-Aware Post-Processing,” *arXiv.org*, 2025. <https://arxiv.org/abs/2504.12215>
- [7] T. Zhou, L. Li, G. Bredell, J. Li, and Ender Konukoglu, “Quality-Aware Memory Network for Interactive Volumetric Image Segmentation,” *arXiv (Cornell University)*, Jan. 2021, doi: <https://doi.org/10.48550/arxiv.2106.10686>.