Optimized U-Net Architecture for Precise Edema ROI Definition in MRI Image Segmentation

Jitendra Kumar Mishra[*1,2] JMISHRA@DEPAUL.EDU
Graduate Student, Data Scientist
Jacob Furst[*1,2] JFURST@DEPAUL.EDU
Professor, Visual Computing
[1] *College Of Computing and Digital Media*
[2] *DePaul University, Chicago, IL 60604, USA*
Mark Conneely[†3] MARK.CONNEELY@VA.GOV
MD, CIIP, Head of Radiology and Nuclear Medicine
[3] *Captain James A.Lovell Federal Health Care Center, Chicago, IL 60202, USA*

Editors: Under Review for MIDL 2025

Abstract

Medical image segmentation is crucial in computer-aided diagnosis, enabling precise identification of anatomical structures and pathological regions. This study presents an optimized U-Net architecture to enhance segmentation accuracy and efficiency in identifying edema in MRI images, emphasizing its **novel application** in defining regions of interest (ROI) for edema.

The improved U-Net model features refined upsampling paths with convolutional blocks and padding strategies, addressing spatial misalignment issues. Data augmentation techniques, such as random flips and rotations, were consistently applied to both images and masks, increasing the variability and robustness of the dataset. Different loss functions, including Mean Squared Error (MSE) and Dice Loss, were evaluated for their impact on performance, highlighting their role in mitigating overfitting.

The performance of the model, assessed using Intersection over Union (IoU) scores and other segmentation metrics 2.5, demonstrates superior segmentation accuracy and robustness under varying conditions. This refined U-Net architecture promises significant improvements in clinical workflows and diagnostic precision for the detection of edema on MRI images.

Keywords: Medical Image Segmentation, U-Net, Data Augmentation, Dice Loss, IoU Score, Edema, RAVD, DSC, ASSD, MSSD, Fracture, MRI Scan.

1. Introduction

Medical image segmentation [5; 14] plays a vital role in computer-aided diagnosis (CAD) systems, providing detailed and accurate delineation of anatomical structures and pathological regions [1; 12]. Accurate segmentation is essential for the diagnosis and treatment planning of various medical conditions. The U-Net architecture is widely recognized for its effectiveness in biomedical image segmentation due to its symmetric contracting and expanding paths that facilitate precise localization and context capture. However, the standard U-Net architecture [3; 11; 13] has limitations, particularly in segmenting small or poorly defined structures. This study aims to optimize the U-Net architecture to improve

[*] Contributed equally
[†] Contributed equally

its performance in medical image segmentation, focusing on the identification of edema in MRI images. Segmentation of edema is crucial as it serves as a significant indicator of various medical conditions, requiring a precise definition of ROI for effective diagnosis and treatment.

To enhance the U-Net architecture Figure 1, we introduced several modifications. The upsampling path was refined with carefully tuned convolutional blocks and padding strategies, ensuring spatial alignment and accurate feature integration. Using convolutional transpose and bilinear interpolation in the upsampling layers, we addressed common misalignment issues between feature maps, thus improving the model's ability to preserve spatial details and maintain robust contextual understanding.

Additionally, we implemented an augmented dataset approach [16; 15], employing data-augmentation techniques such as random horizontal and vertical flips and rotations. These augmentations were applied consistently to both input images and their corresponding masks, increasing dataset variability and robustness. We also evaluated the impact of different loss functions, including Mean Squared Error (MSE) and Dice Loss [2; 17], on model performance. The results underscored the importance of loss function selection in mitigating overfitting and improving generalization on unseen data.

The optimized U-Net model was rigorously assessed using Intersection over Union (IoU) scores [4] across training and validation datasets. We also analyzed the effect of batch size on IoU during testing to demonstrate the model's stability and robustness under varying conditions. Our findings reveal that the optimized U-Net architecture consistently performed better during each phase of our experiment. With no direct comparison to any previous work, this represents a novel application, offering improved segmentation accuracy and consistency. These advances are crucial to improving clinical workflows and diagnostic precision, particularly in the context of edema detection in MRI images.

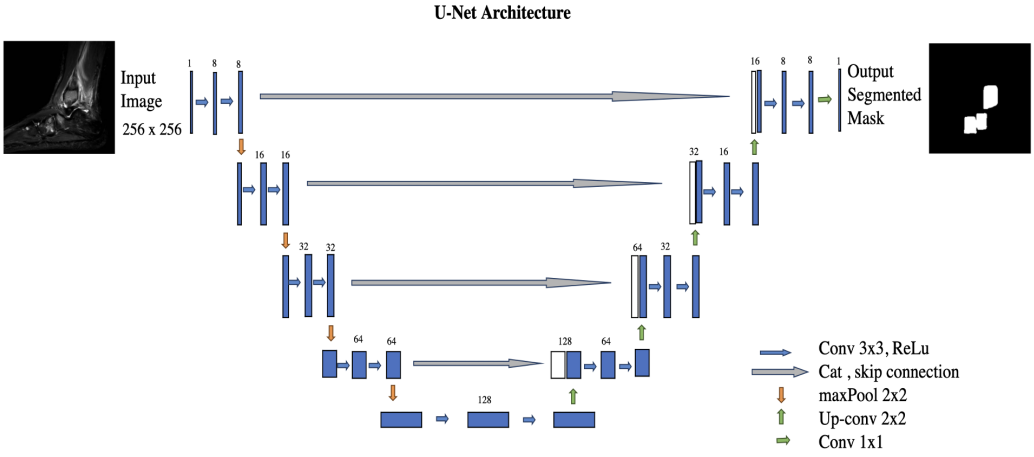

Figure 1: A U-Net architecture, showcasing the convolutional neural network (CNN) layers for feature extraction and upsampling with skip connections for precise segmentation

2. Methods

We employed a custom dataset for training and evaluation purposes, consisting of pairs of gray-scale images and their corresponding binary masks.

2.1. Dataset

The data set used in this study consists of foot MRI scans of the ankle and foot, collected from 96 patients diagnosed with acute stress fractures. DICOM files from each MRI scan were meticulously converted into individual JPEG images without any compression, ensuring a lossless transformation process, which serve as input data for the segmentation models. The data set includes both images and the corresponding binary mask images. A total of 861 image files from 96 unique patients were divided into training (559), testing (186), and validation (116) sets, to ensure that there was no patient overlap between the splits.

2.2. Binary Mask Generation

The binary masks are created from coordinate information provided along with the image data. The masks were generated by trained medical student research assistants and the precision was checked by a board certified radiologist with fellowship training in musculoskeletal imaging. Each mask represents specific regions of interest on the MRI slices. The coordinate data specifies the locations of relevant anatomical structures, which are used to generate the binary masks. These masks are created by marking the specified coordinates on a blank image, resulting in binary masks where the regions of interest are highlighted. Figure 2

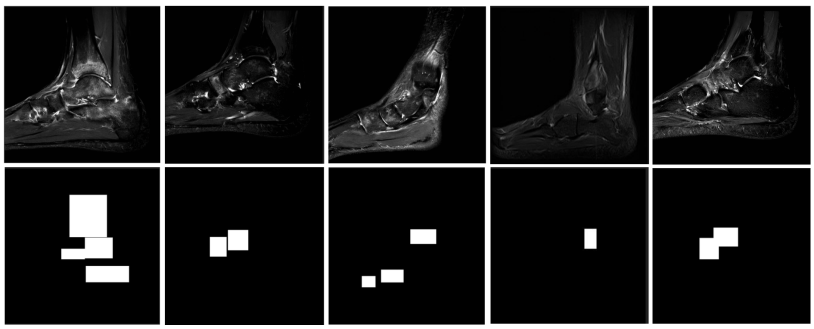

Figure 2: Images and their corresponding masks, highlighting single/multiple regions of interest (ROIs) of edemas in MRI scans with bounding boxes.

2.3. Data processing

The images were prepossessed [6; 9], which involved resizing to 256x256 pixels. To standardize the input format, each image was converted to gray scale. The data set was then divided into training, validation and test sets to ensure robust model evaluation.

To enhance the robustness of the model and prevent overfitting[10], **data augmentation** techniques are applied. These techniques include horizontal and vertical flips, as well as random rotations. The enhancements are applied consistently to both images and masks to maintain the integrity of the data while increasing the limited data to 2x. To ensure we maintain the aspect ratio of the images, we perform functional padding with bilenear interpolation during upsampling in the U-Net architecture during training.

2.4. Model Architecture and Parameters

The U-Net architecture is a powerful model for image segmentation, particularly effective in biomedical imaging. It features an encoder-decoder structure with skip connections, capturing both fine details and contextual information. The encoder progressively reduces spatial dimensions while increasing feature maps through Down blocks, which consist of **max pooling** layers followed by DoubleConv blocks (each with two convolutional layers, **batch normalization**, and **ReLU** activation). In the decoder, upsampling is performed using **bilinear interpolation** followed by functional **padding** to maintain the aspect ratio of the resized images, and is followed by DoubleConv blocks. Skip connections concatenate feature maps from the encoder to the decoder, improving spatial accuracy. The final convolutional layer reduces the number of channels to match the number of classes for segmentation, producing the output mask. Designed in **PyTorch**, U-Net is renowned for its effectiveness in detailed and accurate segmentation tasks.
Two loss functions are evaluated: **Mean Squared Error (MSE)** and **Dice loss** Equation (1). The choice of loss function is based on the nature of the task. Dice Loss is used for segmentation tasks where the overlap between predicted and true masks is critical, while MSE is utilized for image-to-image translation. The **Adam** optimizer is used with an optional experimental **L2** regularization term to combat overfitting. The learning rate is set to **0.001**, and the weight decay parameter for L2 regularization is set to **1e-5** when applicable.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \qquad , \qquad DiceLoss = 1 - \frac{2 \cdot |y_i \cap \hat{y}_i|}{|y_i| + |\hat{y}_i|} \tag{1}$$

The model is trained for a predefined number of epochs with early stopping implemented to halt training when validation loss plateaus. Training includes the computation of **Intersection over Union (IoU)** scores to monitor the model's performance on both training and validation datasets. Additionally, threshold techniques are applied to optimize the IoU scores by varying the threshold values and selecting the optimal threshold based on performance metrics.
The implementation is carried out using the PyTorch and torch vision libraries. Custom data augmentation and data handling classes are developed to pre-process the MRI slices and apply the necessary transformations. The data set is loaded using PyTorch's DataLoader, which manages batching and shuffling of the data during training.

2.5. Performance Metrics

To evaluate the segmentation performance between the ground truth and the segmentation map of the proposed network, we utilize the performance metrics based on the volumetric

size similarity and surface distance measures ([7], [8]). In definition, ground truth is denoted by A, and the segmented result is denoted by B. The Dice Similarity Coefficient (DSC) and the (IOU) score expressed as Equation (2)

$$\mathrm{DSC}(A,B) = \frac{2|A \cap B|}{|A| + |B|} \times 100 \qquad , \qquad \mathrm{IOU}(A,B) = \frac{|A \cap B|}{|A \cup B|} \tag{2}$$

Relative Absolute Volume Difference (RAVD) denoted as Equation (3)

$$\mathrm{RAVD} = \mathrm{Abs}\left(\frac{|B| - |A|}{|A|}\right) \times 100 \tag{3}$$

The surface distance measures are Average Symmetric Surface Distance (ASSD) and Maximum Symmetric Surface Distance (MSSD). ASSD denoted as Equation (4), Let S(A) indicates the number of surface voxels of A. The shortest distance of a random voxel V to S(A) is expressed in (Appendix B)

$$\mathrm{ASSD}(A,B) = \frac{1}{|S(A)| + |S(B)|}\left(\sum_{S_A \in S(A)} d(S_A, S(B)) + \sum_{S_B \in S(B)} d(S_B, S(A))\right) \tag{4}$$

Maximum symmetric surface distance (MSSD) denoted as Equation (5)

$$\mathrm{MSSD}(A,B) = \max\left\{\max_{S_A \in S(A)} d(S_A, S(B)), \max_{S_B \in S(B)} d(S_B, S(A))\right\} \tag{5}$$

The DSC, IoU, and RAVD measured in percentage, and surface distance measures (ASSD and MSSD) measured in millimeters (mm). For DSC and IoU 100% is the best segmentation, and 0% is the worst segmentation, RAVD 0% is the best segmentation, and 100% is the worst segmentation. These metrics expressed in percentage. The ASSD and MSSD measured in millimeters (mm). For ASSD and MSSD 0 mm is the best segmentation, and the upper value has not bound; the maximum value shows the worst segmentation.

We also introduced a **custom scoring system** to evaluate model performance by comparing the binary MRI image (I) with the predicted segmentation mask (P) and the actual bounding box mask (A). The Prediction Score and Actual Score are calculated as the overlap of I with P and A, respectively. This approach accounts for the irregular shapes of predicted masks compared to the square ground truth masks, providing a robust evaluation of the model's segmentation accuracy. (Appendix C)

3. Results

The experiments were conducted in multiple phases, each designed to enhance different aspects of the segmentation model. Each phase focused on refining specific components of the workflow and improvements were observed as a result of these iterative enhancements. In this study, we explored various configurations of a U-Net model for foot magnetic resonance imaging segmentation to improve generalization performance and achieve higher test metrics. The models were trained, validated, and tested on unseen data, with each phase incorporating additional measures to address overfitting. The BaseModel (M) showed

Table 1: Results, (M) Base model , (M1) Base model trained with augmented data, (M2) model trained with L2 regularization, early stopping, (M3) Model using Dice loss for improvement. Every Model is an incremental step to the previous model

Model	**Tr.Loss**	**Tr.IoU**	**Val.Loss**	**Val.IoU**	**Tst.IoU**	**Tr.Time**
BaseModel(M)	0.001	0.85	0.01	0.49	0.43	21m 30s
(M)Augment (M1)	0.002	0.83	0.01	0.57	0.49	43m 11s
(M1)L2 reg (M2)	0.01	0.68	0.01	0.51	**0.54**	7m 24s
(M2)DiceLoss (M3)	0.20	0.66	0.37	0.52	**0.57**	4m 40s

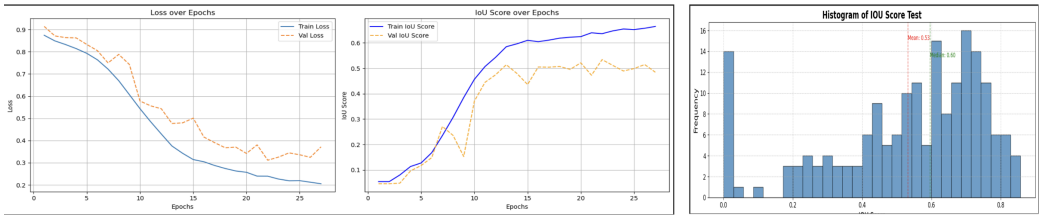

Figure 3: Results of best Test score IOU model (M3) showing Training, validation and frequency distribution of IoU scores.

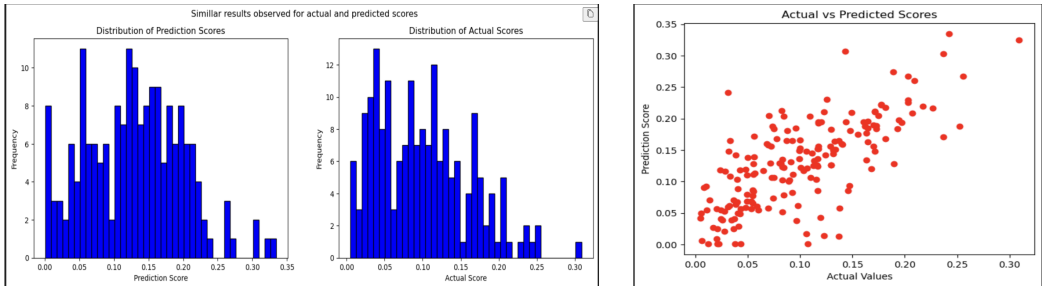

Figure 4: Results of best model (M3) displaying Custom calculated Scores with closer resemblance for prediction and actual scores. (Appendix C)

overfitting despite a high training IoU. Adding data augmentation (M1) improved generalization, while further incorporating L2 regularization (M2) increased test IoU. The best performance was achieved with the model augmented with L2 regularization and Dice loss (M3) demonstrating the effectiveness of these techniques in improving model generalization. Figure 3, Table 1 shows the training and Testing data, Table 2 shows comparison results, Figure 5 shows the test images displaying the actual image and mask with predicted binary mask and the threshold mask.

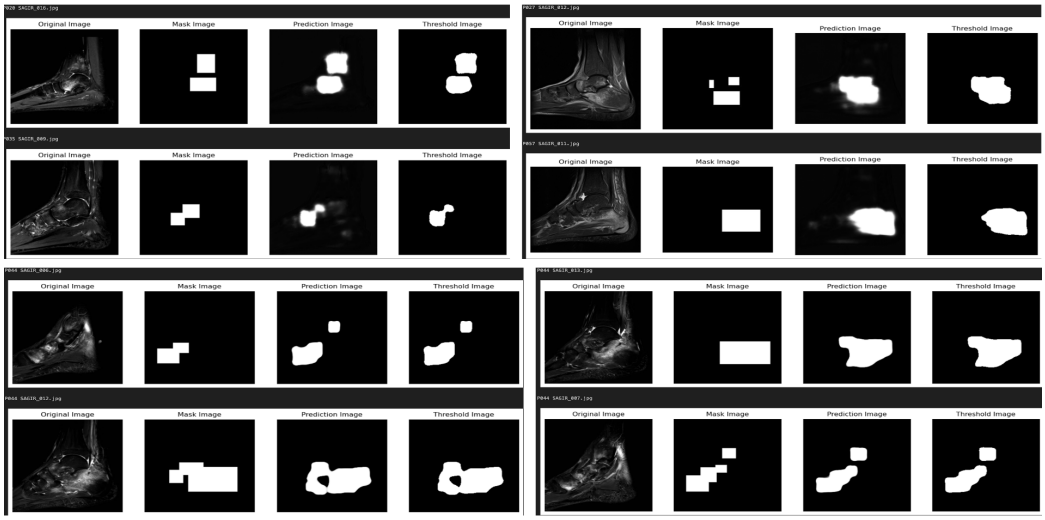

Figure 5: M3 Test results showing actual image, true mask, predicted and thresholded binary image, Threshold for segmentation was set at 0.4 after threshold analysis

3.1. Model Comparison

In this study, we compared our optimized U-Net model(M3) with other state-of-the-art (SOTA) models, including YOLO, nnDetect to evaluate its relative performance in foot MRI segmentation Table 2. These models were selected due to their proven effectiveness in object detection and segmentation tasks, providing a robust baseline for comparison.
YOLO (You Only Look Once) is a real-time object detection model known for its speed and accuracy in detecting objects in images. While YOLO is highly efficient in object detection tasks, it was not initially designed for fine-grained medical image segmentation. When applied to our MRI segmentation task, YOLO's performance was limited, as it struggled with the precision required to accurately delineate edema regions in the foot MRI scans, resulting in lower Intersection over Union (IoU) and other segmentation-specific metrics.
nnDetect, despite being designed for object detection, performed the worst in our comparison. The model was unable to accurately segment the MRI images, failing to detect or delineate the edema regions. This failure was attributed to nnDetect's inability to handle the intricate structures and subtle features typically found in medical images, leading to poor segmentation performance. Metrics such as IoU and other metrics were significantly

lower compared to all other models tested, highlighting nnDetect's limitations for medical image segmentation tasks.
We also experimented with Mask R-CNN with a ResNet-50 backbone, which is traditionally used for segmenting multiple objects in an image. However, this model was not suitable for our problem statement, as it struggled to segment the complex tissue and water pocket regions in MRI images. The presence of multiple tissue types and small, varied water pockets in the foot MRI scans posed significant challenges for Mask R-CNN.(experiment results not included in table comparison).

Table 2: Comparative Result analysis for Edema Segmentation with ours M3 Model. with IOU as avg value and other metrics with median value.

Model	**IOU(%)**	**DSC(%)**	**RAVD (%)**	**ASSD(mm)**	**MSSD(mm)**
nnDetect	22.3	24.3	63.6	16.6	64.5
YOLO	41.7	67.0	52.2	**4.8**	38.6
Our U-net M3	**57.0**	**81.1**	**32.6**	6.8	**35.0**

4. Conclusion

The experimental results demonstrate that the incorporation of additional measures such as data augmentation, L2 regularization, and Dice loss significantly improves the performance of the model on unseen test data, particularly in this **novel application of MRI segmentation** for defining regions of interest **(ROI) for edema**. Among the models tested, the BaseModel with Augmentation, L2 Regularization and Dice Loss (M3) emerged as the best model, achieving the highest test IoU of 0.57. This indicates that the novel combination of these techniques in the training configuration effectively addressed issues of overfitting and class imbalance, leading to superior performance in accurately segmenting edema regions on MRI scans. **Notably, our experiment has no comparison with any other work as this represents a novel application in this specific context**.
In an effort to benchmark our model against state-of-the-art (SOTA) segmentation models, we compared it with YOLO, nnDetect, and Mask R-CNN with a ResNet-50 backbone. Our optimized U-Net model consistently outperformed these SOTA models in this specific application. The findings highlight the potential of this approach to improve clinical workflows for edema assessment, setting a benchmark for future research in this domain.
A limitation of this study is the relatively small dataset, which may have constrained the model's performance. We believe that training on a larger dataset could have a significant positive impact on the results, further enhancing the model's generalization capabilities and robustness in clinical applications for edema assessment. In the ongoing efforts, we aim at 1) Train our model with a bigger dataset to compare the results and potentially achieve better performance. 2) Extend this approach to include a knee dataset, thereby broadening the applicability of our methodology to different types of MRI segmentation tasks.

Acknowledgments

We extend our heartfelt gratitude to the team of radiologists and medical imaging specialists whose invaluable insights and guidance were crucial to the success of this study. Their expertise significantly contributed to the refinement and accuracy of our research.

We also acknowledge the support of our institutional partners at **Chicago Medical School Rosalind Franklin University of Medicine and Science** for granting us access to MRI datasets, which were fundamental to our study. Special thanks to RA colleagues who worked in collecting this data **1) Alexander Winters, 2) Annemarie Nwajei, 3)Chirag Divecha 4) Rikhil Makwana 5) Ju Oh 6) Logan Holland 7) Chris Yun 8) Kamil Bosak.**

Special thanks to the technical team at DePaul University for their assistance with computational resources, which were essential for conducting our experiments.

Additionally, we are grateful to our colleagues **1) Prof. Thiru Ramaraj, 2) Michael Robinson, 3) Joanne Lin and 4) Om Dange** for their constructive feedback and encouragement throughout the research process. We extend a special thanks to **5) Rashi Jain** for her invaluable contributions to the model comparison section, which greatly aided in the completion of this research.

This research was supported by funding from the **DePaul University College of Computing and Digital Media (CDM)**. Their financial support was instrumental in advancing our work and achieving the results presented in this paper.

References

Majid M. Qayyum A. et al Anwar, S.M. Medical image analysis using convolutional neural networks: A review. *J Med Syst 42, 226, https://doi.org/10.1007/s10916-018-1088-1*, 2018.

Martha Büttner, Lisa Schneider, Aleksander Krasowski, Vinay Pitchika, Joachim Krois, Hendrik Meyer-Lueckel, and Falk Schwendicke. Conquering class imbalances in deep learning-based segmentation of dental radiographs with different loss functions. *Journal of Dentistry*, 148:105063, 2024. ISSN 0300-5712. doi: https://doi.org/10.1016/j.jdent.2024.105063. URL https://www.sciencedirect.com/science/article/pii/S030057122400232X.

Wanli Chen, Yue Zhang, Junjun He, Yu Qiao, Yifan Chen, Hongjian Shi, and Xiaoying Tang. Prostate segmentation using 2d bridged u-net. https://arxiv.org/abs/1807.04459, 2018.

Yeong-Jun Cho. Weighted intersection over union (wiou) for evaluating image segmentation. *Pattern Recognition Letters*, 185:101–107, 2024. ISSN 0167-8655. doi: https://doi.org/10.1016/j.patrec.2024.07.011. URL https://www.sciencedirect.com/science/article/pii/S0167865524002149.

Marleen de Bruijne. Machine learning approaches in medical image analysis: From detection to diagnosis. *Medical Image Analysis*, 33:94–97, 2016. ISSN 1361-8415. doi: https://doi.org/10.1016/j.media.2016.06.032. URL https://www.sciencedirect.com/science/article/pii/S1361841516301098. 20th anniversary of the Medical Image Analysis journal (MedIA).

K. M. Hanson. Introduction to Bayesian image analysis. In M.H̃. Loew, editor, *Medical Imaging: Image Processing*, volume 1898 of *Proc. SPIE*, pages 716–731, 1993.

Tobias Heimann, Bram van Ginneken, Martin A. Styner, Yulia Arzhaeva, Volker Aurich, Christian Bauer, Andreas Beck, Christoph Becker, Reinhard Beichel, GyÖrgy Bekes, Fernando Bello, Gerd Binnig, Horst Bischof, Alexander Bornik, Peter M. M. Cashman, Ying Chi, AndrÉs Cordova, Benoit M. Dawant, MÁrta Fidrich, Jacob D. Furst, Daisuke Furukawa, Lars Grenacher, Joachim Hornegger, Dagmar KainmÜller, Richard I. Kitney, Hidefumi Kobatake, Hans Lamecker, Thomas Lange, Jeongjin Lee, Brian Lennon, Rui Li, Senhu Li, Hans-Peter Meinzer, GÁbor Nemeth, Daniela S. Raicu, Anne-Mareike Rau, Eva M. van Rikxoort, MikaËl Rousson, LÁszlÓ Rusko, Kinda A. Saddi, GÜnter Schmidt, Dieter Seghers, Akinobu Shimizu, Pieter Slagmolen, Erich Sorantin, Grzegorz Soza, Ruchaneewan Susomboon, Jonathan M. Waite, Andreas Wimmer, and Ivo Wolf. Comparison and evaluation of methods for liver segmentation from ct datasets. *IEEE Transactions on Medical Imaging*, 28(8):1251–1265, 2009. doi: 10.1109/TMI.2009.2013851.

Huiyan Jiang, Shaojie Li, and Siqi Li. Registration-based organ positioning and joint segmentation method for liver and tumor segmentation. *BioMed Research International*, 2018(1):8536854, 2018. doi: https://doi.org/10.1155/2018/8536854. URL https://onlinelibrary.wiley.com/doi/abs/10.1155/2018/8536854.

Limiao Li, Keke He, Xiaoyu Zhu, Fangfang Gou, and Jia Wu. A pathology image segmentation framework based on deblurring and region proxy in medical decision-making system. *Biomedical Signal Processing and Control*, 95:106439, 2024. ISSN 1746-8094. doi: https://doi.org/10.1016/j.bspc.2024.106439. URL https://www.sciencedirect.com/science/article/pii/S174680942400497X.

Zeju Li, Konstantinos Kamnitsas, and Ben Glocker. Analyzing overfitting under class imbalance in neural networks for image segmentation. *IEEE Transactions on Medical Imaging*, 40(3):1065–1077, 2021. doi: 10.1109/TMI.2020.3046692.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015. URL http://arxiv.org/abs/1505.04597.

Dinggang Shen, Guorong Wu, and Heung-Il Suk. Deep learning in medical image analysis. *Annual review of biomedical engineering*, 19(1):221–248, 2017.

Jason Walsh, Alice Othmani, Mayank Jain, and Soumyabrata Dev. Using u-net network for efficient brain tumor segmentation in mri images. *Healthcare Analytics*, 2:100098, 2022. ISSN 2772-4425. doi: https://doi.org/10.1016/j.health.2022.100098. URL https://www.sciencedirect.com/science/article/pii/S2772442522000429.

Guotai Wang, Wenqi Li, Maria A. Zuluaga, Rosalind Pratt, Premal A. Patel, Michael Aertsen, Tom Doel, Anna L. David, Jan Deprest, Sébastien Ourselin, and Tom Vercauteren. Interactive medical image segmentation using deep learning with image-specific fine tuning. *IEEE Transactions on Medical Imaging*, 37(7):1562–1573, 2018. doi: 10.1109/TMI.2018.2791721.

Mingle Xu, Sook Yoon, Alvaro Fuentes, and Dong Sun Park. A comprehensive survey of image augmentation techniques for deep learning. *Pattern Recognition*, 137:109347, 2023. ISSN 0031-3203. doi: https://doi.org/10.1016/j.patcog.2023.109347. URL https://www.sciencedirect.com/science/article/pii/S0031320323000481.

Zhenghua Xu, Shengxin Wang, Gang Xu, Yunxin Liu, Miao Yu, Hongwei Zhang, Thomas Lukasiewicz, and Junhua Gu. Automatic data augmentation for medical image segmentation using adaptive sequence-length based deep reinforcement learning. *Computers in Biology and Medicine*, 169:107877, 2024. ISSN 0010-4825. doi: https://doi.org/10.1016/j.compbiomed.2023.107877. URL https://www.sciencedirect.com/science/article/pii/S0010482523013422.

Rongjian Zhao, Buyue Qian, Xianli Zhang, Yang Li, Rong Wei, Yang Liu, and Yinggang Pan. Rethinking dice loss for medical image segmentation. In *2020 IEEE International Conference on Data Mining (ICDM)*, pages 851–860, 2020. doi: 10.1109/ICDM50108.2020.00094.

Appendix A. Computation Environment and Hardware Specifications

For the computation of our results, we utilized MacBook Air(M3 Chip) equipped with an 8-core CPU and up to a 10-core GPU. This configuration offers substantial processing power and enhanced graphical performance, essential for handling the complex computations involved in our experiments. Additionally, the MacBook Air supports up to 24GB of unified memory, which facilitates efficient data processing and management across different tasks. This combination of high-performance CPU, GPU, and ample unified memory provided a robust environment for executing our computational workloads and achieving reliable results.

Appendix B. Distance of a Random voxel V to S(A) equation

$$d(v, S(A)) = \min_{S_A \in S(A)} \|v - S_A\|, \text{ where } \|\| \text{ denotes the Euclidean distance.} \tag{6}$$

Appendix C. Custom Scoring System

$$\text{Prediction Score} = \frac{\sum I \cdot P}{\sum I} \qquad \text{Actual Score} = \frac{\sum I \cdot A}{\sum I} \tag{7}$$

- I: Binary representation of the actual MRI image.
- P: Predicted segmentation mask.
- A: Actual segmentation mask.
- $\cdot$: Element-wise multiplication.
- $\sum$: Summation over all pixel values.