Enhancing UNet-driven T1-weighted Brain Magnetic Resonance Image Segmentation using Post-processing Techniques

Mahsa Dibaji, Electrical and Software Engineering Department, University of Calgary

***Abstract*— Brain magnetic resonance image (MRI) segmentation is a crucial preprocessing step for various brain analysis tasks. This study aims to improve the quality of segmentation masks predicted by a UNet model by integrating traditional methods and post-processing techniques. The research utilizes two datasets of T1-weighted healthy brain images and their corresponding masks. The UNet model is trained on the CC359 dataset, and its performance is evaluated on a separate set of images from the LPBA40 dataset. A post-processing pipeline, comprising connected component analysis, morphological operations, and region growing, is employed to refine the initial segmentation masks. The results demonstrate enhanced segmentation mask quality for the majority of test images, as evidenced by both visual inspection and evaluation metrics. Limitations encompass potential dataset dependency and method's generalizability to other brain structures and pathologies. This study contributes to more accurate and reliable brain MRI segmentation by leveraging the benefits of both deep learning-based methods and traditional techniques.**

***Index Terms*— Brain Magnetic Resonance Imaging, Image Processing, Convolutionl Neural Network, Skull Stripping**

I. Introduction

MAGNETIC Resonance Imaging (MRI) is a non-invasive medical imaging technique that revolutionized medical diagnosis and treatment by providing detailed images of internal organs, soft tissues, and bones. MRI is now widely used in various medical fields, such as neurology, cardiology, oncology, and orthopedics. MRI segmentation, an essential step in medical image analysis, separates different regions or structures within an MRI image for accurate quantification and characterization of anatomical structures. Brain MRI segmentation plays a critical role in diagnosing and monitoring neurological disorders and influences surgical planning [1].

Traditional methods for brain MRI segmentation, such as manual, semi-automatic, and some fully automatic methods, can be time-consuming, labor-intensive, and prone to errors, especially when dealing with complex brain structures or pathologies. They also require domain expertise and are subject to inter-observer variability. In contrast, deep learning-based approaches like convolutional neural networks (CNNs) offer significant improvements by automatically learning complex patterns from large datasets. CNNs are more versatile, robust, and provide more consistent and reproducible results. However, they can also produce errors or fail to capture fine-grained structures in challenging cases [2].

This study aims to improve the quality of segmentation masks predicted by a UNet model by integrating traditional methods and post-processing techniques. This approach leverages the advantages of both methods while addressing some of their limitations, ultimately contributing to more accurate and reliable brain MRI segmentation.

II. Related Works

Various methods have been developed for brain segmentation, each with different advantages and limitations. Traditional methods, reviewed by Balafar et al. [3] and Despotović et al. [4], include intensity-based, region-based, and boundary-based approaches. While some are computationally efficient or adapt to intensity variations, they may also be sensitive to image noise or require manual initialization.

Deep learning, particularly convolutional neural networks (CNNs), has significantly improved brain MRI segmentation by capturing complex spatial patterns and hierarchies, as discussed by Akkus et al. [5]. However, these methods can be computationally expensive, requiring large amounts of annotated data, and may produce segmentation errors in ambiguous or noisy image regions.

This project aims to address some limitations of existing deep learning-based approaches by combining them with post-processing techniques. By integrating deep learning models with post-processing techniques, we aim to enhance segmentation results, especially in challenging cases. In conclusion, our project builds upon recent advances to develop an improved approach to brain MRI segmentation, combining deep learning methods with post-processing techniques.

The remainder of this paper is organized as follows: first, we present the materials and methods used in our study; next, we describe the experiments conducted and the results obtained; finally, we discuss our findings and draw conclusions from the research.

III. Materials and Methods

A. Datasets

In this study, we employ two datasets for our experiments: the Calgary-Campinas-359 (CC-359) dataset [6] and LONI Probabilistic Brain Atlas 40 (LPBA40) dataset [7].

The CC-359 dataset is a volumetric brain MRI dataset that includes 359 T1-weighted three-dimensional (3D) volumes acquired using different magnetic resonance imaging (MRI) scanners from GE, Philips, and Siemens, with varying magnetic field strengths (1.5 T and 3 T). The dataset encompasses samples from both biological sexes, with 183 (50.97%) female subjects and 176 (49.03%) male subjects, ranging in age from 29 to 80 years. Brain segmentation masks are provided for this dataset, were generated using the consensus algorithm Simultaneous Truth and Performance Level Estimation (STAPLE) and are considered silver standard.

The LPBA40 dataset is a valuable resource in medical imaging, consisting of T1-weighted MRI scans and manually delineated brain structures of 40 healthy adult subjects. Its distinctiveness lies in providing brain masks in native space, allowing researchers to evaluate segmentation and registration algorithms in real-world scenarios. Maintained by the Laboratory of Neuro Imaging (LONI) at the University of Southern California (USC), the LPBA40 dataset is widely used in brain image analysis, offering high-quality ground truth for comparisons and enabling the development and validation of novel medical image processing methods. The authors produced brain segmentation masks by first aligning the MRI volumes to the MNI-305 average brain and then removing extra-meningeal tissue using the Brain Extraction Tool (BET). Any errors in the automated segmentation and removal of the cerebellum and brainstem were manually corrected to obtain the final cerebrum masks.

The CC359 dataset served as the development set, with 70% of the data allocated for training and 30% for validation. The LPBA40 dataset was employed as the test set to assess the effectiveness of the proposed method.

B. Preprocessing and Augmentation

The preprocessing of data involved intensity scaling to standardize the range of pixel intensities across the dataset. This step is crucial for ensuring consistency and facilitating the learning process for the model. In addition to preprocessing, several data augmentation techniques were employed to enhance the model's robustness and generalizability. These augmentations included random cropping to a specific size, random flipping along an axis, random rotation, and random application of Gaussian Noise and Gibbs Noise to selected images.

Gaussian noise, a random noise that follows a Gaussian distribution, is often used in data augmentation to increase a model's robustness against variations in pixel intensities [8]. This type of noise can help the model adapt to real-world scenarios where images might be corrupted by random fluctuations in intensity values. On the other hand, Gibbs noise manifests as ringing patterns near sharp intensity transitions in MRI scans [9]. By introducing Gibbs noise during data augmentation, the model's performance on real-world images containing such artifacts can be improved. These artifacts are common in MRI scans and can pose challenges for automatic segmentation algorithms.

Furthermore, the contrast of some randomly chosen images was adjusted to ensure the model's effectiveness on images with varying contrasts. This step is particularly important, as the model needs to perform well on images acquired from different MRI scanners or with different acquisition parameters, which might result in varying levels of contrast.

C. Initial Segmentation Using UNet

UNet is a popular convolutional neural network architecture, specifically designed for biomedical image segmentation tasks [10]. The architecture consists of an encoding path, which progressively reduces the spatial dimensions while increasing the number of feature channels, and a decoding path, which performs the inverse operation, gradually restoring the spatial dimensions while reducing the number of feature channels. Additionally, skip connections between the encoding and decoding paths help retain fine-grained information, which is crucial for accurate segmentation. The advantages of the UNet model include its ability to handle images with varying input sizes, efficient training on small datasets, and precise localization of segmented structures.

D. Post-Processing

1) Morphological Operations: Morphological operations are essential tools in image processing that can be used to refine and enhance the results of image segmentation tasks. These operations are particularly effective in eliminating noise, filling gaps, and correcting errors in the segmented regions. In this project, we employed several morphological operations, including dilation, erosion, closing, and opening, to improve the initial segmentation masks generated by the UNet model.

Dilation: Dilation expands the boundaries of a region by adding pixels to the perimeter, filling small holes and connecting neighboring regions.

Erosion: Erosion is the opposite of dilation, shrinking the boundaries by removing pixels from the perimeter, eliminating small noise and separating connected regions [11].

Closing: Closing combines dilation followed by erosion, filling small holes or gaps within a region, connecting neighboring regions, and smoothing object boundaries without causing significant size changes.

Opening: Opening involves erosion followed by dilation, effectively removing small noise, separating connected regions, and smoothing object boundaries with minimal impact on the segmented area size [12].

Applying these morphological operations to the initial segmentation masks helped refine the segmented regions, reduce noise, and improve the overall quality of the segmentation results.

2) Connected Component Analysis: Connected Component Analysis (CCA) is a valuable image processing technique for identifying and labeling distinct regions or objects within a binary image. The algorithm groups connected pixels based on predefined connectivity criteria and assigns unique labels to separate regions. CCA is particularly beneficial in refining segmentations, as it enables the removal of small, disconnected regions that may arise from noise or errors during the segmentation process, leaving only the most significant connected components representing the actual structures of interest [13].

3) Confidence Connected Region Growing: Confidence Connected Region Growing (CCR) is an adaptive image segmentation technique that focuses on expanding a seed region based on local statistical properties of the image. The algorithm starts with a set of seed points within the region of interest and iteratively incorporates neighboring pixels that meet specific intensity criteria. By considering the local mean and standard deviation of pixel intensities, CCR adapts to local variations in intensity, making it more robust to noise and intensity inhomogeneities commonly found in medical images, such as MRI scans. [14]

The key advantage of the CCR algorithm lies in its adaptive growth and region specificity. By starting with seed points within the region of interest, the growth is focused on the desired structures, reducing the risk of merging with neighboring regions. This approach allows for more accurate segmentation results, particularly when working with images that exhibit intensity inhomogeneities and noise. [15]

E. Evaluation Metrics

In this project, various evaluation metrics were employed to assess the segmentation quality. These metrics include the Dice Coefficient, Jaccard Index, Hausdorff distance, precision, and recall.

The Dice Similarity Coefficient (also known as the Sørensen-Dice index) is a commonly used metric in medical image segmentation that measures the similarity between two sets. It is defined as the ratio of twice the intersection of two sets divided by the sum of their sizes according to the below formula [16].

$$DSC(A,B) = \frac{2|A \cap B|}{|A| + |B|} \tag{1}$$

The Jaccard Index, also known as the Intersection over Union (IoU), is another metric used to evaluate the similarity between two sets. It is defined as the size of the intersection of two sets divided by the size of their union [16]. The formula for the Jaccard Index is:

$$JI(A,B) = \frac{|A \cap B|}{|A \cup B|} \tag{2}$$

Hausdorff distance is a measure of dissimilarity between two sets, capturing the largest distance between points in one set to the closest points in the other set [16]. It is defined as:

$$HD(A,B) = \max\{\sup_{a \in A} \inf_{b \in B} d(a,b), \sup_{b \in B} \inf_{a \in A} d(a,b)\} \tag{3}$$

Precision and recall are two widely used metrics for evaluating classification performance. Precision measures the fraction of true positive predictions among all positive predictions, while recall quantifies the fraction of true positive predictions among all actual positive instances [16]. The formulas for precision and recall are:

$$Precision = \frac{TP}{TP + FP}, Recall = \frac{TP}{TP + FN} \tag{4}$$

where TP, FP, and FN represent the number of true positives, false positives, and false negatives, respectively.

F. Experimental Environment

To carry out this project, various Python libraries were employed. The PyTorch framework was used for implementing the UNet model, while MONAI was utilized for performing preprocessing and augmentation steps and handling data ingestion. For the post-processing pipeline, SimpleITK and NumPy were leveraged to execute the various stages of the pipeline.

The University of Calgary Advanced Research Computing (ARC) cluster, specifically gpu-a100 and gpu-v100 nodes, was utilized for this project. These powerful GPU resources allowed for efficient processing and analysis of the dataset.

IV. Experiments and Results

In this project, the UNet model was trained on the entire CC-359 dataset for 100 epochs using the ADAM optimizer with a learning rate of 0.001. During the training process, Dice Loss was employed as an evaluation metric, calculating the overlap between prediction and ground truth [17]. Upon completion of the training, the top-performing model was selected based on the evaluation metric on the validation set. Table I shows the loss values on the training and validation set in epoch 99, when the best model was saved. This model was then utilized to generate initial segmentation masks for the images in the LPBA40 dataset. Inference on a separate dataset, such as LPBA40, allows for the assessment of the model's generalization capabilities and performance on unseen data. These masks served as a starting point for further refinement and post-processing in the subsequent steps of the project.

TABLE I
Training and Validation Loss at epoch 99 when the best model was saved

Training Loss	Validation Loss
0.0385	0.0229

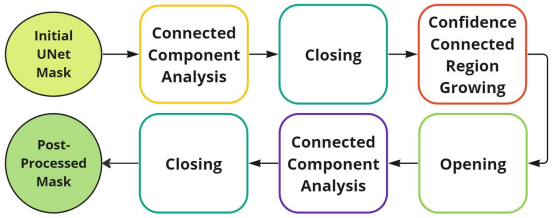

Fig. 1. The Post-processing pipeline applied after the initial segmentation using UNet.

In the post-processing phase, a pipeline was employed to apply a similar set of post-processing techniques on all the images in the test set. Fig. 1 shows the order in which steps were applied in this pipeline. Initially, connected component analysis was performed on the initial masks predicted by UNet to remove any irrelevant segmented regions, retaining only the largest region. This step was followed by a closing operation to fill any holes. Confidence connected region growing was then applied in order to expand the segmentation mask to cover more regions in the brain. In order to select a suitable seed point, two criteria were considered: intensity and position within the image. The chosen seed point had an intensity

close to the average pixel intensity of the initial segmented region and was also located near the center of that region. An opening operation was then applied to remove any small unwanted segmented regions, followed by another connected component analysis to retain the largest segmented region. Finally, another closing operation was performed to fill any possible holes created in previous steps.

Fig. 2 shows a sample image from LPBA40 dataset along with its provided ground truth mask. Figures 3-8 illustrate the initial segmentation mask predicted by UNet for this sample as well as the result of applying each post-processing step in the pipeline on this mask.

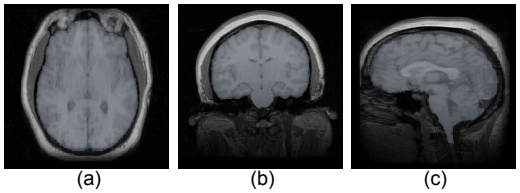

Fig. 2. Ground truth segmentation mask. Center Slices from (a) Axial, (b) Coronal, and (c) Sagittal views.

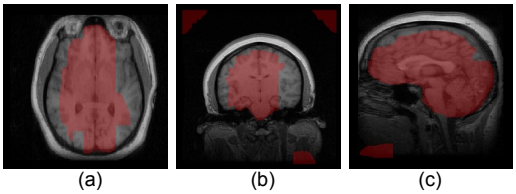

Fig. 3. Initial brain segmentation mask predicted by UNet model. Center Slices from (a) Axial, (b) Coronal, and (c) Sagittal views.

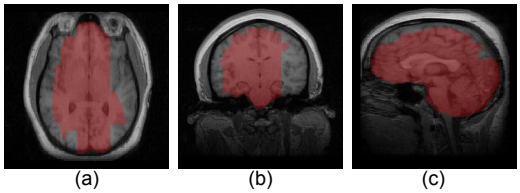

Fig. 4. Brain segmentation mask after applying step 1 and 2: connected component analysis and closing operation with kernel size (2,2,2). Center Slices from (a) Axial, (b) Coronal, and (c) Sagittal views.

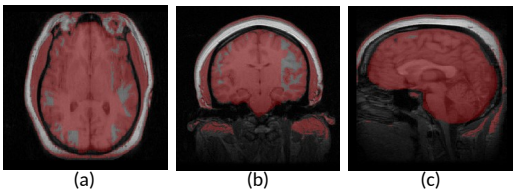

Fig. 5. Brain segmentation mask after applying step 3: confidence connected region growing. Center Slices from (a) Axial, (b) Coronal, and (c) Sagittal views.

Fig. 9 and Fig. 10 illustrates the average value of each metric after getting initial segemntations from UNet and each step in the post-processing pipeline. Furthermore, Fig. 11 demonstrates the Jaccard similarity coefficient for each image before and after applying the entire post-processing pipeline.

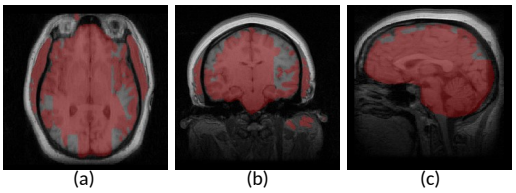

Fig. 6. Brain segmentation mask after applying step 4: opening operation with kernel size (2,2,2). Center Slices from (a) Axial, (b) Coronal, and (c) Sagittal views.

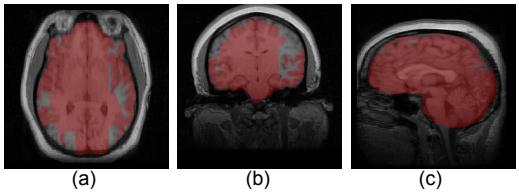

Fig. 7. Brain segmentation mask after applying step 5: connected component analysis. Center Slices from (a) Axial, (b) Coronal, and (c) Sagittal views.

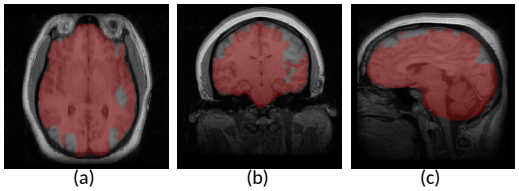

Fig. 8. Brain segmentation mask after applying step 6: closing operation with kernel size (3,3,3). Center Slices from (a) Axial, (b) Coronal, and (c) Sagittal views.

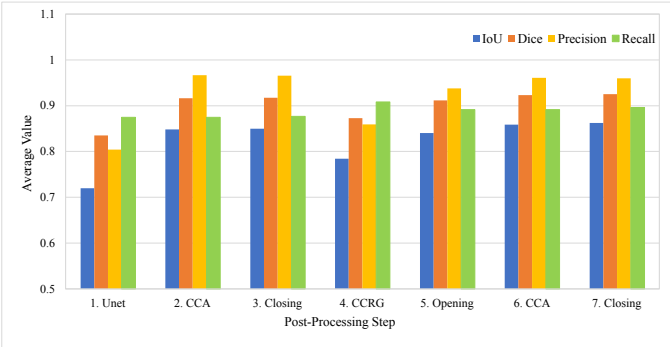

Fig. 9. Comparison of Dice Coefficient Similarity, Jaccard Index (IoU), Precision, and Recall after applying each step in the post-processing pipeline. Steps include: connected component analysis (CCA), and confidence connected region growing (CCRG).

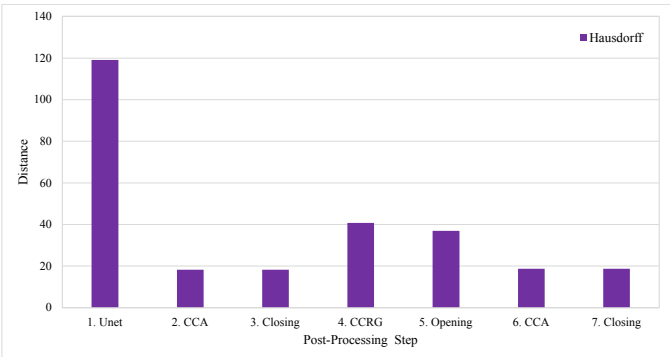

Fig. 10. Comparison of Hausdorff distance after applying each step in the post-processing pipeline. Steps include: connected component analysis (CCA), and confidence connected region growing (CCRG).

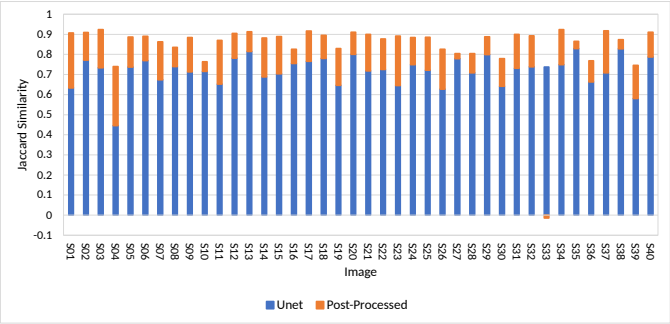

Fig. 11. Comparison of Jaccard similarity coefficients for each image before and after applying the post-processing pipeline.

V. Discussion

In this project, we aimed to improve the quality of segmentation masks predicted by a UNet model using a post-processing pipeline. The results indicate that the proposed method was successful in enhancing the segmentation masks for almost all images in the test set compared to the initial predicted masks by the UNet model. We observed that the average value of evaluation metrics increased after each post-processing step, except for the confidence connected region-growing step. This step initially caused over-segmentation in some regions, which temporarily reduced the quality of the mask. However, subsequent post-processing steps were able to address this issue effectively. Overall, the proposed method significantly refined poor-quality masks while preserving and even slightly enhancing the good quality masks.

This project has some limitations that should be addressed. Firstly, the post-processing methods employed may not perform as well on other datasets, although the goal was to create a pipeline as general as possible. Secondly, all images used in this project were of healthy brains, which raises concerns about the method's generalizability to other brain structures and pathologies. Thirdly, combining traditional segmentation methods with CNNs may increase the required computational resources compared to using each method alone. Lastly, the effectiveness of the post-processing methods is dependent on the performance of the CNN model. In rare cases where the model predicts a very poor segmentation mask, the post-processing steps might not be able to compensate for that performance completely and therefore not significantly enhance the mask.

Some future work includes further optimization of hyperparameters and the most effective order to apply post-processing steps. It is also worth experimenting with other CNN architectures, post-processing methods, and different orders of applying post-processing steps to yield better outcomes. We can analyze the performance of this method on other datasets and compare its effectiveness with other methods in the literature. By addressing these limitations and exploring future directions, we aim to continue enhancing the segmentation results and contribute to more accurate and reliable brain MRI segmentation.

VI. Conclusion

In conclusion, this study presents a method that combines the strengths of a UNet model and a post-processing pipeline to improve brain MRI segmentation. The results demonstrate the effectiveness of this approach in refining segmentation masks, particularly in challenging cases where the initial predictions may be of lower quality. This research contributes to the ongoing effort to enhance the accuracy and reliability of medical image analysis.

References

[1] P. Kalavathi and V. S. Prasath, "Methods on skull stripping of mri head scan images—a review," *Journal of digital imaging*, vol. 29, pp. 365–379, 2016.

[2] L. Pei, M. Ak, N. H. M. Tahon, S. Zenkin, S. Alkarawi, A. Kamal, M. Yilmaz, L. Chen, M. Er, N. Ak, *et al.*, "A general skull stripping of multiparametric brain mris using 3d convolutional neural network," *Scientific Reports*, vol. 12, no. 1, p. 10826, 2022.

[3] M. A. Balafar, A. R. Ramli, M. I. Saripan, and S. Mashohor, "Review of brain mri image segmentation methods," *Artificial Intelligence Review*, vol. 33, pp. 261–274, 2010.

[4] I. Despotović, B. Goossens, and W. Philips, "Mri segmentation of the human brain: challenges, methods, and applications," *Computational and mathematical methods in medicine*, vol. 2015, 2015.

[5] Z. Akkus, A. Galimzianova, A. Hoogi, D. L. Rubin, and B. J. Erickson, "Deep learning for brain mri segmentation: state of the art and future directions," *Journal of digital imaging*, vol. 30, pp. 449–459, 2017.

[6] R. Souza, O. Lucena, J. Garrafa, D. Gobbi, M. Saluzzi, S. Appenzeller, L. Rittner, R. Frayne, and R. Lotufo, "An open, multi-vendor, multi-field-strength brain mr dataset and analysis of publicly available skull stripping methods agreement," *NeuroImage*, vol. 170, pp. 482–494, 2018.

[7] D. W. Shattuck, M. Mirza, V. Adisetiyo, C. Hojatkashani, G. Salamon, K. L. Narr, R. A. Poldrack, R. M. Bilder, and A. W. Toga, "Construction of a 3d probabilistic atlas of human cortical structures," *Neuroimage*, vol. 39, no. 3, pp. 1064–1080, 2008.

[8] R. Verma and J. Ali, "A comparative study of various types of image noise and efficient noise removal techniques," *International Journal of advanced research in computer science and software engineering*, vol. 3, no. 10, 2013.

[9] L. F. Czervionke, J. M. Czervionke, D. L. Daniels, and V. M. Haughton, "Characteristic features of mr truncation artifacts," *American journal of neuroradiology*, vol. 9, no. 5, pp. 815–824, 1988.

[10] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pp. 234–241, Springer, 2015.

[11] P. Soille, *Erosion and Dilation*, pp. 63–103. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004.

[12] P. Soille, *Opening and Closing*, pp. 105–137. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004.

[13] L. He, X. Ren, Q. Gao, X. Zhao, B. Yao, and Y. Chao, "The connected-component labeling problem: A review of state-of-the-art algorithms," *Pattern Recognition*, vol. 70, pp. 25–43, 2017.

[14] D. L. Pham, C. Xu, and J. L. Prince, "Current methods in medical image segmentation," *Annual review of biomedical engineering*, vol. 2, no. 1, pp. 315–337, 2000.

[15] R. Adams and L. Bischof, "Seeded region growing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 6, pp. 641–647, 1994.

[16] A. A. Taha and A. Hanbury, "Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool," *BMC medical imaging*, vol. 15, no. 1, pp. 1–28, 2015.

[17] S. Jadon, "A survey of loss functions for semantic segmentation," in *2020 IEEE conference on computational intelligence in bioinformatics and computational biology (CIBCB)*, pp. 1–7, IEEE, 2020.