Volumetric Quantification of Stroke Lesions Using DW-MRI and 3D U-Net Approach

Wanis BARREH[1*], Ines BEN ALAYA[1], Rayhane BEN AMOR[1], Iyadh RIAHI[1], Ridha BEN SALAH[1]

[1] Laboratory of Biophysics and Medical Technology, Higher Institute of Medical Technology of Tunis
Tunis, Tunisia
*wanis.b@gmail.com

Abstract. Diffusion-Weighted Magnetic Resonance Imaging (DW-MRI) is an effective method for early ischemic stroke detection, often identifying strokes within the first few minutes of onset. Additionally, it enables the quantification of lesion volume, which is a critical factor in determining whether to proceed with endovascular thrombectomy, particularly beyond the 6-hour onset window. However, variability in clinical assessments of lesion volume can lead to missed opportunities for reperfusion therapy.

This research aims to develop Convolutional Neural Network (CNN) algorithms for the automated segmentation of acute ischemic lesions in DW-MRI. Specifically, it focuses on highlighting and evaluating the performance of the 3D U-Net segmentation method, emphasizing its effectiveness in accurately identifying and characterizing ischemic lesions.

The clinical utility of the method was evaluated in patients with confirmed brain infarction. To assess the effectiveness of the proposed approach, metrics including sensitivity and specificity were consistently applied across all segmentation methods. The reference data, serving as the ground truth, were predominantly determined through assessments by radiologists.

The results demonstrate that our proposed segmentation method provides a precise estimation of lesion volume compared to the current manual method used in clinical practice. Specifically, the 3D U-Net segmentation achieved the following results: a precision of 85.57%, a sensitivity of 80.69%, a specificity of 99.92%, an accuracy of 99.809%, and an F1 score of 85.19%. When compared to two other segmentation methods, our approach underscores its utility in achieving accurate lesion volume estimation.

Keywords: Stroke, diffusion MRI, CNN, U-Net, ADC, Artificial intelligence, Deep learning.

1 Introduction

Ischemic stroke is as a significant factor in human disability and mortality rates. Stroke lesions are typically divided into two main areas: the infarct core, comprising irreversibly damaged tissue, and the penumbra, consisting of at-risk tissue that may recover if blood flow is restored. [1]

Successful delineation and localization of acute cores or penumbras in clinical settings are critical, providing valuable insights into the potential salvageable tissue through alternative therapies.

Diffusion Weighted Magnetic Resonance Imaging (DW-MRI) is widely used to detect and quantify acute ischemic stroke lesions in their early stages. Accurate segmentation of ischemic stroke regions is necessary for treatment planning and evaluating disease outcomes. [2]

However, manual segmentation of stroke boundaries in DW-MRI images is tedious and time-consuming, requires neuroanatomical expertise, and is subject to variability [3]. Additionally, segmentation outcomes may be influenced by various factors, including the complex structure of brain tissues and MRI machine quality, similarities between artifacts and stroke lesions, and variations in analysis among analysts. Therefore, there is a pressing need to develop automatic methods capable of recognizing and separating stroke lesions without human intervention.

Automated stroke lesion segmentation can improve clinical practice to better assess and evaluate the risks of each treatment. [4]

In recent years, numerous techniques for automatically segmenting ischemic stroke lesions have been proposed, generally falling into two main categories: those based on handcrafted features and those leveraging automatic feature extraction methods, such as deep learning. Among these, the 3D U-Net has emerged as a state-of-the-art approach for brain lesion segmentation and shows significant promise in accurately identifying and quantifying ischemic stroke lesions. However, challenges and gaps remain in optimizing these models for clinical application, particularly in terms of generalizability and robustness across diverse patient populations. [5,6,7,8]

This study focuses on investigating the variability in infarct volume estimation derived from DW-MRI images, as well as comparing the accuracy and reliability of different segmentation techniques in delineating ischemic stroke lesions from DW-MRI data.

© The Author(s) 2025

S. Krichen et al. (eds.), *Proceedings of International Conference on Decision Aid and Artificial Intelligence (ICODAI 2024)*, Atlantis Highlights in Intelligent Systems 12,
https://doi.org/10.2991/978-94-6463-654-3_8

To achieve this objective, we begin by describing the segmentation method employed in our study, specifically the automatic segmentation using the 3D U-Net model. In the third section, we present the experimental results, and in the fourth section, we discuss these findings in detail.

2 Materials and Methods

Accurate segmentation of stroke lesions on MRI images, along with quantitative assessments, is very important for neurologists in various diagnostic and therapeutic purposes, and they are pertinent to studying brain health and disease. Many segmentation methods have been proposed for predicting stroke patient outcomes. [9]
Fig. 1 illustrates the procedure proposed in our study, which is based on a structured methodology for the detection and volumetric characterization of ischemic stroke. It consists of four main steps: loading the data, extracting acquisition and patient-related parameters, applying lesion quantification methods, and calculating lesion volume.

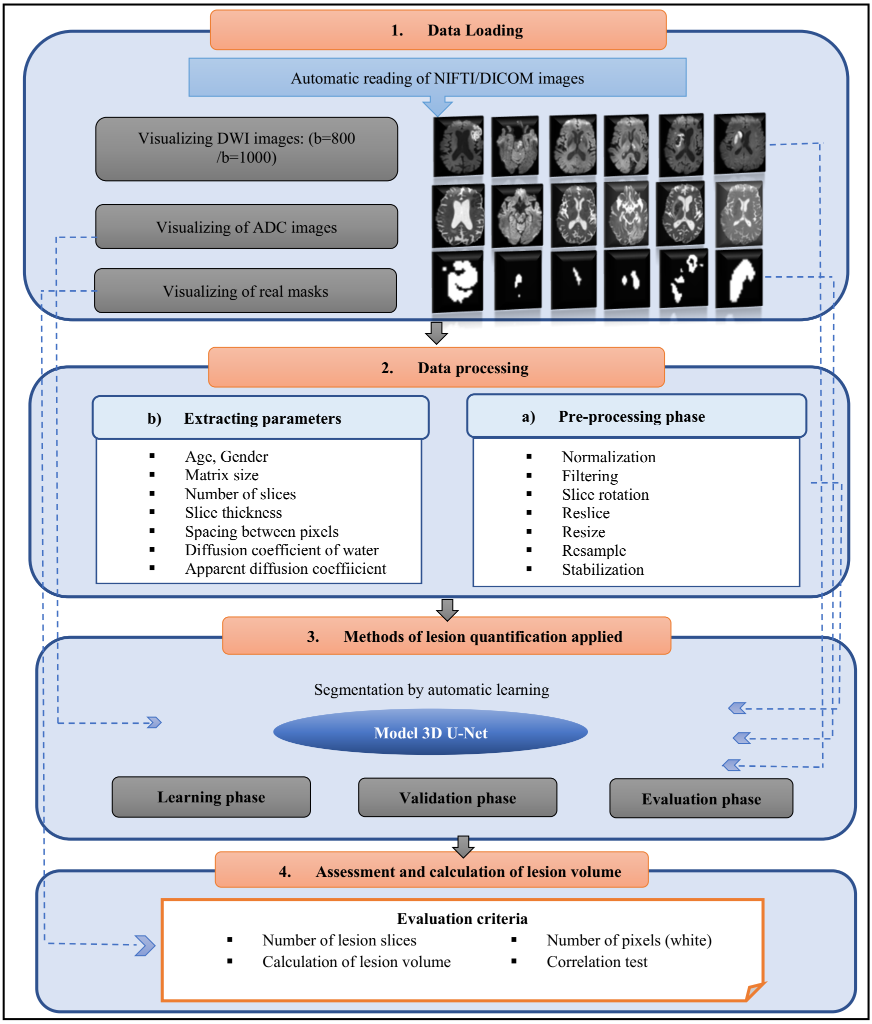


Fig. 1. Representative diagram of the proposed procedure

2.1. Data Loading

The first step involves loading and visualizing the imaging data. Diffusion-weighted images with different b-values (800 and 1000 s/mm²) are loaded to identify restricted diffusion areas, which are indicative of ischemia. ADC maps are then visualized to confirm ischemia by observing regions of low diffusion. The actual masks used for segmentation are also displayed to allow for qualitative and quantitative validation and comparison.

2.2. Data Processing

This step involves extracting parameters for analysis, such as demographic information from JSON files and image attributes from NIfTI files, to ensure compatibility and accuracy in volumetric measurements. Preprocessing steps include normalization, slice rotation, noise filtering, reslicing to a uniform voxel size, resizing for model compatibility, and stabilizing models for improved image alignment and consistency.

2.3. Segmentation Method

Many automated techniques have been developed to enhance the delineation of stroke lesions from DWI images. Machine learning (ML) models have been widely applied and have achieved good segmentation performance. ML techniques use multiple imaging features, including those not visible to humans with a consistent accuracy. Deep learning (DL) models have also been applied to perform segmentation of acute stroke lesions from DWI images. However, none of the CNN architectures developed to date have achieved high accuracy in segmenting ischemic stroke lesions, mainly due to the heterogeneity in location, shape, size, image intensity and texture, especially in this imaging modality. The 3D U-Net architecture is a three-dimensional extension of the classic U-Net architecture, designed for volumetric segmentation of 3D medical images, such as MRI and CT images. It retains the fundamental principles of the U-Net architecture while extending its operations to three dimensions to account for the volumetric nature of the data.
Fig. 2 presents the architecture of 3D U-Net as applied in this research.

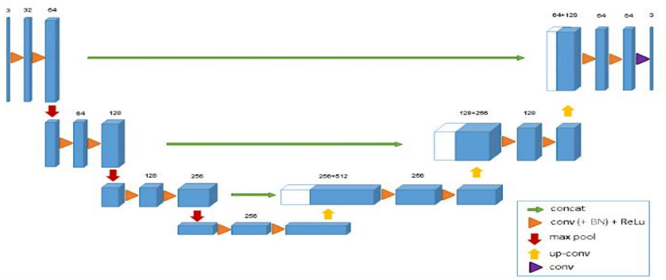


Fig. 2. 3D U-Net architecture

2.4. Optimizer

The neural network's optimizer is an algorithm or method used to change the attributes of your neural network such as weights and learning rate in order to reduce the losses. Various optimization approaches exist, including gradient descent, stochastic gradient descent, Adam, RMSprop, and Adagrad. Adam learning rate = 0.0001 was chosen as the optimizer because it provides great computational efficiency, low memory requirement was an important issue for us, and it is invariant to diagonal scaling of gradients.

2.5. Evaluation Metrics

Dice coefficient. The Dice coefficient DS [10] is one of a number of measures of the extent of spatial overlap between two binary images. It is commonly used in reporting performance of segmentation and gives more weighting to instances where the two images agree. Its values range between 0 (no overlap) and 1 (perfect agreement). In this paper the Dice values are expressed as percentages and obtained using Equation:

$$DS = \frac{2(A \cap G)}{(A \cap G + A \cup G)} \quad \textbf{(1)}$$

2.6. Lesion volume calculation

The final step involves assessing and calculating the lesion volume, which includes measuring the volume of the lesioned areas, counting the number of slices containing lesions, and counting the number of pixels corresponding to lesions in the segmented images. A correlation test is performed to ensure the consistency of the results. These measures enable precise quantification and characterization of ischemic strokes, aiding in accurate diagnosis and patient follow-up.

2.7. Database and Image Collection

In this context, we utilized the dataset from the Sub-Acute Ischemic Stroke Lesion Segmentation (SISS) challenge, which is a subset of the larger Ischemic Stroke Lesion Segmentation(ISLES-2022) dataset [13]. This publicly available dataset comprises 250 heterogeneous patients, and each patient includes three MRI image modalities (DWI, ADC, and Flair), all patients aged 40 to 80 years, with images of varying resolutions (112x112, 115x115, 128x128, 130x130, 192x192 and 256x256 pixels) and different numbers of slices (25, 27, 30, 63, 72, 73, 74 and 76 slices). The slice thickness (2 mm and 1.756 mm), the spacing between slices (1.453, 2 and 2 mm), and the position of the lesions (Middle Cerebral Artery stroke in either the left or right hemisphere, and Vertebrobasilar stroke) also varied. After adequate preprocessing to normalize and homogenize these variations, we selected 140 patients for our study, The others, about 110 patients, were eliminated because they did not present JSON files (matrix size, number of slices and b-factor), and even they did not present any lesion. Out of these 140 patients, 108 were randomly selected for training, 12 for validation and 20 for evaluation of the different segmentation models. This strategy of dividing into training and validation sets is important in the development and evaluation of machine learning models according to the Validation-split = 0.1.
Fig. 3 shows a diagram of the modelling process.

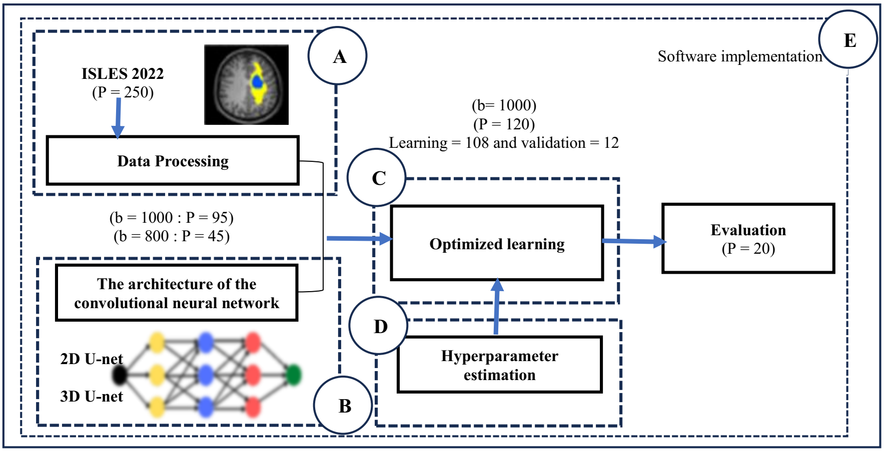


Fig. 3. Validation and evaluation process of applied models

3 Setup and Implementation

The experiments were developed using by the Adam method [11] with initial learning rate of 0.0001 as our optimizer. The learning rate is scaled down by a factor of 0.1 if no progress is made for 15 epochs on validation loss. Early-stopping technique is adopted after 50 epochs with no progress on the validation loss. The experiments are performed on a computer with an Intel Core i7-6800K CPU, 64GB RAM and Nvidia GeForce 1080Ti GPU with 11GB memory. The computer operates on Windows 10 with CUDA 9.0. The network is implemented on Tensorflow (http://www.tensorflow.org/). The MR image files are stored as Neuroimaging Informatics Technology Initiative (NIfTI) format, and processed using Simple Insight ToolKit (SimpleITK) [12]. We use PyQt5, Vedo, and Matplotlib for the visualization of results.

As the MR images were acquired on different MR scanners, the matrix size varies. Therefore, we resample all the MR images to the same dimension (128,128,128) by interpolating the images using "scikit-image" (skimage) library.

4 RESULTS

The 3D U-Net model was trained for 100 epochs, with a total training time of 121751 seconds, equivalent to 33 hours and 49 minutes.

The confusion matrix (Fig. 4) obtained after training the models show that 3D U-Net model has 25 117 989 good predictions against 47 835 bad predictions with the validation data which gives us an accuracy of 99.80%.

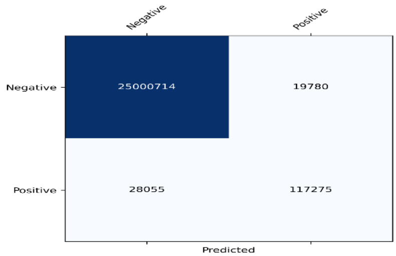


Fig. 4. Confusion matrix of DWI images by 3D U-Net model for validation data

Table 1 presents the formulas for the main metrics used to evaluate the performance of the model, calculated from the confusion matrix according to the values of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN):

Table 1. Metrics derived from confusion matrix

	Definition	*Formula*	*Unet-3D*
Precision (PR)	Proportion of correct positive predictions among all positive predictions.	$\frac{TP}{TP+FP}$	85,56 %
Sensitivity (SE)	Ability of the model to correctly identify positive instances.	$\frac{TP}{TP+FN}$	80,69 %
Specificity (SP)	Ability of the model to correctly identify negative instances.	$\frac{TN}{TN+FP}$	99,92 %
Accuracy (AC)	Proportion of correct predictions among all predictions.	$\frac{TP+TN}{TP+TN+FP+FN}$	99,80 %
F1-Score (FC)	Harmonic mean of accuracy and sensitivity.	$\frac{2*PR*SE}{PR+SE}$	83,06 %

4.1. Qualitative comparison

A qualitative comparison between simple thresholding, the 2D U-Net model and the 3D U-Net model for ischemic stroke segmentation reveals significant differences in terms of accuracy and robustness. Table 2 shows the 3D results of our tests with the chosen patient, where we applied the thresholding method, the 2D U-Net deep learning method and the 3D U-Net method. Lesions are marked in red in the images.

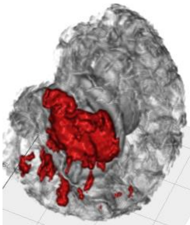
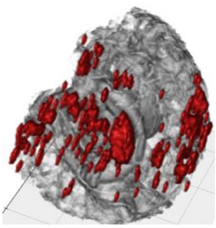
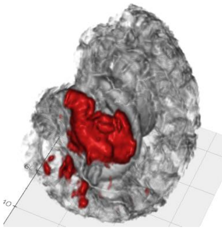
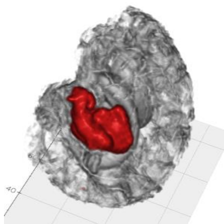
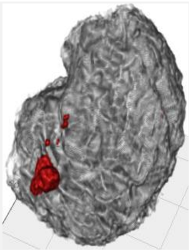
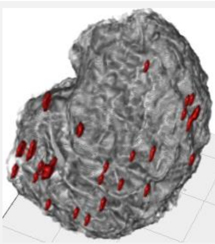
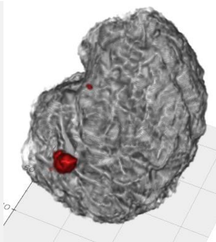
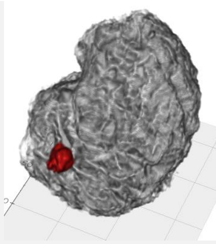
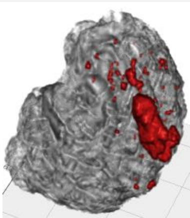
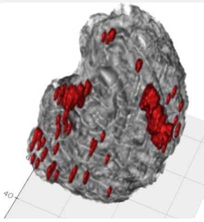
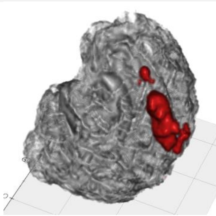
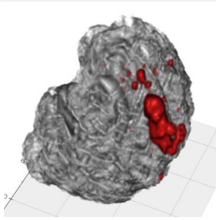
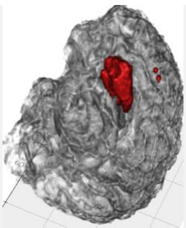
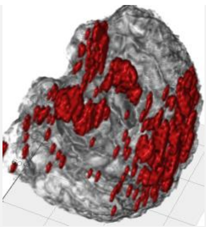
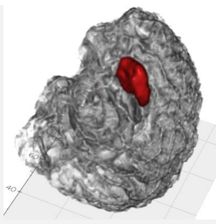
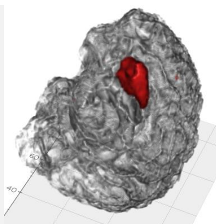
4.2. Quantitative comparison

In this section, we present the results obtained for each patient. Table 3 presents the lesion volumes calculated using three distinct methods: the optimized thresholding method (ADC segmentation), the 2D U-Net deep learning method (DWI segmentation), and the 3D U-Net deep learning method (DWI segmentation).

The lesion volume is calculated using the following formula:

$$Lesion\ Volume = Number\ of\ Lesion\ Pixels\ in\ all\ slices \times Pixel\ Width \times Pixel\ Height \times Slice\ Thickness \times 10^{-3} \quad \textbf{(2)}$$

Table 2. 3D qualitative study of the three applied segmentation methods

Patients	Actual Mask	Applied segmentation methods		
		Thresholding	***2D U-Net***	***3D U-Net***
N°168 (66F)				
N°172 (87F)				
N°178 (83F)				
N°190 (60M)				

N°211
(78M)

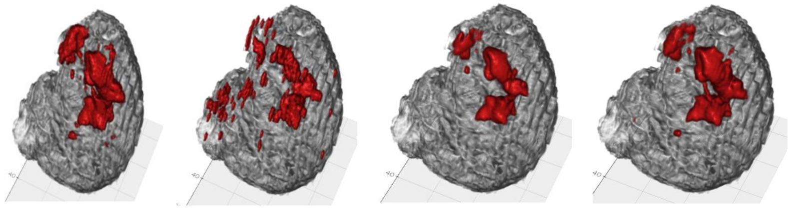


Table 3. Quantitative results obtained using the three methods: ADC Thresholding, 2D U-Net, 3D U-Net and the calculation of similarity coefficient

Patients	Number of lesion pixels				Lesion volume (ml)						
	Actual mask	***ADC Thresholding***	***2D U-Net***	***3D U-Net***	***Réf***	***ADC Thresholding***	***DS1***	***2D U-Net***	***DS 2***	***3D U-Net***	***DS 3***
P 168 (MCA-R)	3817	4583	3449	3681	30.536	36.664	0.213	27.880	0.859	22.088	0.788
P 172 (MCA-L)	728	540	379	593	5.824	4.320	0.104	3.032	0.651	4.744	0.780
P 178 (MCA-L)	1920	2145	1423	1474	15.360	17.160	0.492	11.384	0.809	11.792	0.816
P 180 (MCA-R)	2471	1433	1787	2217	19.768	11.464	0.175	14.296	0.826	17.736	0.882
P 190 (MCA-L)	1361	8006	1772	1310	10.888	64.048	0.143	10.176	0.815	10.480	0.867
P 211 (MCA-L)	7494	4773	3879	6589	59.952	38.184	0.549	31.032	0.673	52.712	0.875
P 221 (VB)	11862	18434	12637	11469	94.896	147.472	0.739	101.096	0.906	91.752	0.922
P 223 (MCA-L)	17435	14481	18150	15803	139.48	115.848	0.627	145.200	0.876	126.424	0.915
P 231 (VB)	4085	8040	3274	3612	32.680	64.320	0.427	26.192	0.865	28.896	0.871
P 247 (MCA-L)	2282	3208	1207	1696	18.256	25.664	0.310	9.656	0.682	13.568	0.805

The data in the table 3 will serve as a basis to determine whether the observed differences between the results of the simple thresholding method in ADC, the segmentation method by the 2D U-Net model and the 3D U-Net model compared to the reference values are statistically significant. The table contains data from 10 patients (8 patients with MCA (Middle Cerebral Artery) stroke in either the left or right hemisphere and 2 patients with Vertebro-Basilar (VB) stroke, for whom we calculated the similarity score (DS1, DS2, DS3) for each method.

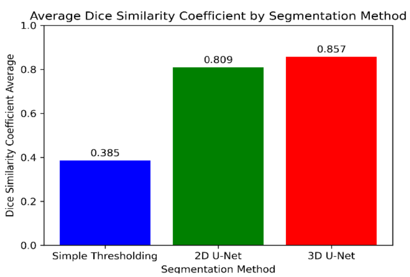


Fig. 5. Dice test results by the three segmentation methods

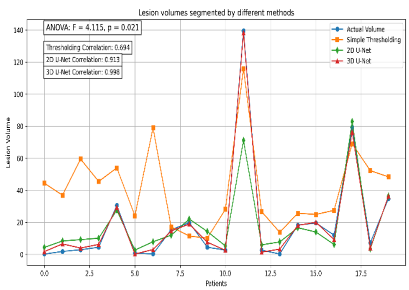


Fig. 6. ANOVA test results by the threes segmentation methods

The results of the ANOVA test for the segmentation of stroke lesions by the three methods (3D U-Net, 2D U-Net and Simple Thresholding) indicate an F-statistic of 4.115 which compares the variance between the three methods and a p-value of 0.021 lower than the common threshold of 0.05 shows that the differences observed between the segmentation methods are statistically significant.

Table 4 compares the proposed 3D UNet model with other prior models. The proposed model showed high performance compared to other models. However, we must note that the compared models applied different datasets to train and test their architectures. In spite of this, the given model demonstrates high performance in all three evaluation parameters, dice/f1 score similarity coefficient, precision and recall/sensitivity.

Table 4. Comparison of the proposed model test set with the prior models

Methods	Dice/f1 score	Precision	Recall/sensitivity
Proposed model	0.85	0.85	0.80
D-UNET[13]	0.53	0.63	0.52
X-NET[14]	0.48	0.60	0.47
V-NET[15]	0.43	0.50	0.49
DeepLabv3+[16]	0.46	0.34	0.44

Table 5 shows the spent time for the two methods, the proposed 3D UNet segmentation and the manual segmentation using ITK-SNAP. An evaluation was carried out and the results reveal significant differences between manual and automated segmentation. Indeed, the manual method requires an average time of 24.5 minutes to achieve segmentation for a patient. On the other hand, our proposed method has proven to be significantly more efficient in terms of speed. The application of this method allows segmentation to be achieved in an average time of 2.5 seconds for a patient.

Table 5. Comparison of spent time between manual and automated segmentation

Methods	Spent time
Proposed model	2.5 seconds
Manual segmentation using ITK-SNAP	24.5 minutes

5 Discussion

Following our qualitative and quantitative comparison of the results from the thresholding method and the 2D U-Net and 3D U-Net methods against the reference lesion masks in the study of detection and volumetric characterization of patients with ischemic stroke, we observe that, in most cases, the deep learning segmentation method is closest to the reference results in terms of lesion slices, lesion volumes, and the overall visual appearance of the lesion. These results highlight the increased efficiency and accuracy of deep learning-based approaches in lesion detection and analysis in the context of ischemic stroke.

5.1. Effect of Batch Size

In the case of the smallest batch size (i.e. 1) you update the weights and biases for every single input, which can be very good for accuracy but computationally very expensive.
In the other limit, where you use the largest batch size (i.e. "n" where n=all your data) you update your weights and biases only once. This method is computationally really good, but can be a very bad model.
Ideally you want to be somewhere in between to be computationally efficient and also get good accuracy.

Table 6 shows the Dice Similarity results for the four-batch size value 2,4,6 et 8 obtained during data validation of training using our first implementation of 2D U-Net and 3D U-Net trained on 50 epochs.

Table 6. Batch size effect on training

Batch Size	Dice Similarity	
	2D U-Net	**3D U-Net**
2	0.802	0.670
4	0.785	0.877
6	0.782	0.832
8	0.785	0.814

5.2. Impact of Lesion Size

In this subsection, we further divide our test set into large lesions set and small lesions set. A subject is categorized into the small lesions set only if all of the lesions are ≤5 mL. Otherwise, the subject is included in the large lesions set.

In our test set, the large lesions and the small lesions set include 11 subjects and 9 subjects respectively. Our proposed method achieves a DC of 0.85 on the large lesions set, while a DC of 0.54 on the small lesions set.

This limitation indicates that improvements are needed to optimize the detection and characterization of these small lesions. To enhance segmentation in these cases, more sophisticated pre-processing approaches to reduce artifacts and finer-grained network architectures specific to small lesions are needed. Additionally, data augmentation techniques and the integration of additional clinical data can also help improve segmentation accuracy.

6 Data Availability

For this study, we use ISLES 2022 open dataset provided by a medical image segmentation challenge [17] at the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2022 [18]. The dataset is available by the address http://www.isles-challenge.org/.

7 Conclusion

In conclusion, while deep learning-based AI techniques show promise in stroke diagnosis and lesion segmentation, challenges remain, particularly with small or chronic lesions. Current methods often rely on single MRI modalities and modest datasets, limiting their effectiveness. Proper preprocessing and multi-modal imaging are crucial for accurate segmentation. Future research should focus on refining network architectures and using diverse datasets to improve segmentation accuracy and clinical relevance.

References

1. Pu, L., et al. « Projected Global Trends in Ischemic Stroke Incidence, Deaths and Disability-Adjusted Life Years From 2020 to 2030, » *AHA Journals*, vol. 54, no. 15, p. 22, 24 avril 2023.
2. Wessels, T., et al. « Contribution of Diffusion-Weighted Imaging in Determination of Stroke Etiology, » *American Journal of Neuroradiology*, vol. 27, no. 11, pp. 35-39, janvier 2006.
3. Ryu, W., et al. « Acute Infarct Segmentation on Diffusion-Weighted Image Using Deep Learning Algorithm and RAPID DWI: A Comprehensive Stroke Center Clinical Validation Study, » *MedRxiv*, vol. 19, no. 11, pp. 1-18, 23 février 2023.
4. Subbanna, N., et al. « Stroke Lesion Segmentation in FLAIR MRI Datasets Using Customized Markov Random Fields, » *Frontiers in Neurology*, vol. 10, 24 mai 2019.
5. Wong, K., et al. « Automatic Segmentation in Acute Ischemic Stroke: Prognostic Significance of Topological Stroke Volumes on Stroke Outcome, » *AHA Journals*, vol. 53, no. 19, p. 22, 12 mai 2022.
6. Moon, H. S., Heffron, L., Mahzarnia, A., et al. « Automated Multimodal Segmentation of Acute Ischemic Stroke Lesions on Clinical MR Images, » *Elsevier*, vol. 92, pp. 45-57, 7 octobre 2022.
7. Soleimani, P., & Farezi, N. « Utilizing Deep Learning via the 3D U-Net Neural Network for the Delineation of Brain Stroke Lesions in MRI Images, » *Scientific Reports*, vol. 1, Iran, 13 novembre 2023.

8. Wang, W. C., Chien, S. Y., Tsai, S. T., et al. « Automated Delineation of Acute Ischemic Stroke Lesions on Non-Contrast CT, » *Elsevier*, vol. 11, no. 193, p. 11, 29 février 2024.
9. Moon, H. S., et al. « Automated Multimodal Segmentation of Acute Ischemic Stroke Lesions on Clinical MR Images, » *PubMed*, vol. 9, no. 110, pp. 45-57, 7 juin 2022.
10. Dice, L. R. « Measures of the Amount of Ecologic Association Between Species, » *Ecology*, vol. 26, pp. 297–302, 1945.
11. Kingma, D. P., & Ba, J. « Adam: A Method for Stochastic Optimization, » *arXiv Preprint arXiv:1412.6980*, 2014.
12. Lowekamp, B., Chen, D., Ibanez, L., & Blezek, D. « The Design of SimpleITK, » 2013.
13. Zhou, Y., Huang, W., Dong, P., Xia, Y., & Wang, S. « D-UNet: A Dimension-Fusion U-Shape Network for Chronic Stroke Lesion Segmentation, » *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 3, pp. 940–950, 2021.
14. Qi, K., Yang, H., Li, C., Liu, Z., Wang, M., et al. « X-Net: Brain Stroke Lesion Segmentation Based on Depthwise Separable Convolution and Long-Range Dependencies, » in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Shenzhen, China, pp. 247–255, 2019.
15. Pinheiro, G. R., Voltoline, R., Bento, M., & Rittner, L. « V-Net and U-Net for Ischemic Stroke Lesion Segmentation in a Small Dataset of Perfusion Data, » in *International MICCAI Brainlesion Workshop*, Granada, Spain, pp. 301–309, 2018.
16. Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. « Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation, » in *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, Germany, pp. 801–818, 2018.
17. ISLES'22 Challenge. https://isles22.grand-challenge.org/, 2022. [Online; accessed 06 December 2022].
18. The Medical Image Computing and Computer-Assisted Intervention Society. 2021. [Online]. Available: http://www.miccai2022.org/.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.