



HAL
open science

Comparative analysis of three advanced deep learning algorithms for Multiple Sclerosis lesion segmentation in FLAIR MRI

Yi Zhu, Thomas Grenier, C. Revol-Muller

► **To cite this version:**

Yi Zhu, Thomas Grenier, C. Revol-Muller. Comparative analysis of three advanced deep learning algorithms for Multiple Sclerosis lesion segmentation in FLAIR MRI. IEEE ICSP'24 & 5th Sino-French Workshop 2024 "Medical Image Analysis and AI (MAI)", IEEE Beijing Section, Beijing Jiaotong University, and Soochow University, Oct 2024, Suzhou, China, France. hal-04767216

HAL Id: hal-04767216

<https://hal.science/hal-04767216v1>

Submitted on 5 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Comparative analysis of three advanced deep learning algorithms for Multiple Sclerosis lesion segmentation in FLAIR MRI

Yi Zhu, Thomas Grenier, Chantal Revol-Muller

CREATIS

University of Lyon

Lyon, France

{yi.zhu, thomas.grenier, chantal.muller}@creatis.insa-lyon.fr

Abstract—This paper addresses the pressing need for enhanced tools in the diagnosis and management of Multiple Sclerosis (MS), particularly in the accurate detection and segmentation of MS lesions. Leveraging recent advances in deep learning, we evaluate the performance of three state-of-the-art algorithms, focusing on their potential to improve both precision and efficiency in MS lesion segmentation from medical images. Our study provides critical insights into the strengths and limitations of each model, offering valuable guidance for future applications of AI in MS diagnosis and treatment.

Keywords—Deep Learning, MS lesion segmentation.

I. INTRODUCTION

A. Multiple Sclerosis

Multiple Sclerosis (MS) is a chronic autoimmune disorder that affects the central nervous system (CNS), leading to a broad spectrum of neurological symptoms. The disease is characterized by the immune system attacking the myelin sheath, the protective covering of nerve fibers, causing inflammation, scarring (sclerosis), and eventually impairing nerve signal transmission. The progression and symptoms of MS are highly variable, manifesting in a range of physical, cognitive, and psychological impairments. This unpredictability and heterogeneity make both diagnosis and monitoring particularly challenging.

Magnetic Resonance Imaging (MRI) serves as the gold standard for diagnosing and monitoring MS. Among the various MRI modalities, 3D Fluid-Attenuated Inversion Recovery (FLAIR) images are particularly effective at highlighting MS lesions, as they suppress the cerebrospinal fluid signal, enhancing the visibility of white matter lesions. However, the manual interpretation of these images is labor-intensive and subject to inter-observer variability, leading to inconsistencies in both diagnosis and treatment planning.

In response to these challenges, there is a growing need for fully automated methods that can provide accurate and rapid segmentation of MS lesions from MRI scans. While numerous automatic segmentation methods have been proposed, recent advancements in deep learning have led to significant improvements in segmentation accuracy. However, despite these ad-

vances, current state-of-the-art methods still do not fully match the precision of expert manual segmentation, highlighting the need for further progress in developing reliable automated solutions for MS diagnosis and management.

B. Deep Learning-Based Lesion Segmentation Methods

Yang Ma *et al.* conducted a comprehensive study in 2022 on deep learning methods for MS lesion segmentation. They observed that traditional machine learning models often struggle with the Dice similarity coefficient (DSC), a standard metric for lesion segmentation. The DSC can be skewed by large lesions, potentially masking poor performance on cases with smaller lesions. Additionally, the heterogeneous spatial distribution and appearance of MS lesions, coupled with their small proportion relative to the total brain volume, create highly imbalanced datasets, further complicating model training.

To address these issues, more efficient inference techniques have emerged. Fully convolutional networks (FCNs) bypass fully connected layers, relying solely on convolutional layers for pixel-wise predictions. Architectures such as U-Net [1], which incorporate an encoder-decoder structure, balance network complexity with performance. Building on U-Net, Isensee *et al.* developed nnU-Net [2], a framework that automatically adjusts hyperparameters, preprocessing, and training schemes to suit different datasets, requiring no manual intervention. Another notable approach is Aslani *et al.*'s [3] fully convolutional CNN framework, which processes features from different 2D modalities independently and uses multi-modal fusion and multi-scale upsampling to generate precise lesion segmentations.

C. MS Dataset

This study leverages a publicly available dataset from the MICCAI 2016 MS lesion segmentation challenge. The dataset consists of high-quality scans from 53 MS patients, adhering to the OFSEP protocol, which is used in France for the national MS cohort. The MRI scans were acquired from three different French clinical sites, using four MRI scanners from Siemens,

Philips, and GE, with magnetic field strengths of 3T and 1.5T. For each patient, seven manual lesion segmentations were performed by trained experts across the three sites, and a consensus "ground truth" segmentation was generated using the LOP STAPLE algorithm. The dataset was divided into a training set of 15 patients and a testing set of 38 patients. The average age of the patients was 45.3 years, with a male-to-female ratio of 0.4. While there were minor demographic differences between the training and testing sets, the training set had an average age approximately 5 years younger than the testing set [4].

II. CUTTING-EDGE DEEP LEARNING METHODS

In recent years, deep learning models have revolutionized medical image analysis, particularly in segmentation tasks. Convolutional neural networks (CNNs), such as U-Net [1], have become the gold standard in biomedical image segmentation due to their ability to capture spatial hierarchies and preserve fine details through symmetric convolutional layers and skip connections.

While transformer-based networks have shown promise by capturing global dependencies across images [5], [6], [7], [8], their high computational demands and prolonged inference times make them less practical for real-world applications where resources are limited. These models often require high-performance hardware, such as GPUs, to offset their complexity, which can limit their scalability in clinical settings.

In contrast, MLP-based networks have emerged as a compelling alternative. Models like MLP-Mixer offer comparable performance to transformers but with significantly lower computational overhead. Inspired by these advancements, UNeXt—a hybrid model combining convolutional layers and MLP blocks—strikes an effective balance between parameter efficiency and strong segmentation accuracy, making it well-suited for practical medical applications with limited resources.

Additionally, the YOLO (You Only Look Once) family of models, originally designed for real-time object detection, offers high speed and efficiency, making it appealing for time-sensitive tasks such as real-time medical imaging.

This study evaluates three cutting-edge models—nnU-Net [2], UNeXt [9], and YOLOv9 [10]—for the segmentation of Multiple Sclerosis (MS) lesions from 3D FLAIR MRI images. By assessing the capabilities of these advanced models, we aim to push the boundaries of MS lesion segmentation and offer valuable insights into their practical applicability in clinical environments, where both accuracy and computational efficiency are paramount.

A. nnU-Net

nnU-Net (no-new-Net) [2] is a robust and flexible framework for medical image segmentation. It is known for its automatic adaptation to different segmentation tasks without the need for manual hyperparameter tuning. The framework dynamically configures itself based on the dataset's characteristics, making it a strong baseline for a wide range of medical

imaging tasks. The original implementation and methodology of nnU-Net can be found in the paper by Isensee et al. (2021)

B. UNeXt

UNeXt [9] is an architecture designed specifically for medical image segmentation, building on the foundational principles of U-Net with additional innovations tailored to enhance performance in specific tasks. Instead of incorporating transformer layers, UNeXt introduces elements of convolutional neural networks (CNNs) and MLP-based tokenization, enabling it to efficiently handle complex segmentation challenges. The detailed implementation and theoretical foundations of UNeXt are discussed extensively in the work by Jeya Maria Jose et al. (2022).

C. YOLO

YOLO (You Only Look Once [11]) is primarily known as an object detection framework, but it has been adapted for segmentation tasks in various research studies. YOLO's speed and accuracy make it a popular choice for real-time applications, including medical image segmentation. The segmentation capabilities of YOLO, particularly in the context of small object detection and segmentation, are covered extensively in research by Redmon et al. (2016) and subsequent studies that expanded its use in medical imaging.

III. IMPLEMENTATION

In this section, we detail the implementation of the three advanced deep learning models: nnU-Net, UNeXt, and YOLOv9. Each model was chosen for its unique architectural advantages and tailored to address the specific challenges of MS lesion segmentation in 3D FLAIR MRI images. The implementation process began with thorough preprocessing of the MRI data to ensure consistency and enhance model performance. We provide an overview of the specific versions of each model used in our study and describe how each was adapted to meet the demands of our segmentation task.

A. Pre-processing

The preprocessing of medical images is a critical step to ensure that the data is in the optimal form for further analysis and model training. In this study, we performed several preprocessing steps to standardize and enhance the MRI images before feeding them into our segmentation models. The following steps were applied:

- 1) **Image Cropping** : the original MRI images had a shape of 261 x 336 x 336. To standardize the input size and reduce computational load, we cropped the images to a size of 256 x 256 x 336. This step ensures that all images have consistent dimensions, which is essential for batch processing and model training.
- 2) **Skull Stripping** : to focus the analysis on brain tissues, we applied skull stripping using the HD-BET (Brain Extraction Tool [12]). Skull stripping removes non-brain tissues such as the skull, scalp, and other extraneous structures from the MRI images. This step is crucial for reducing noise and

improving the accuracy of subsequent image processing and analysis.

- 3) **N4 Bias Field Correction** : MRI images often suffer from intensity inhomogeneities, which can affect the accuracy of segmentation algorithms. To correct this, we applied the N4 ITK algorithm for bias field correction.
- 4) **Normalization and Slicing**: for methods using 2D segmentation such as UNeXt and YOLO, additional preprocessing steps were performed:
 - **Outlier removal** : The top 5% of the highest intensity values were removed to minimize the impact of outliers and artifacts.
 - **Normalization** : after bias correction, the intensity values of the MRI images were normalized to ensure that the pixel values are within a consistent range across all images.
 - **Slicing** : the 3D NIFTI files were sliced into 2D PNG images, adapting the volumetric data into a format that can be processed by 2D convolutional neural networks.

It is important to note that the nnU-Net model, being a 3D segmentation method, only utilized the first three preprocessing steps: image cropping, skull stripping, and N4 bias field correction.

B. nnU-Net 3D

For our study, we used the nnU-Net 3D variant due to the volumetric nature of MRI data. The MRI images were preprocessed and standardized as described earlier, and then fed into the nnU-Net model. The model was configured to handle 3D input data with an initial resolution of 128x128x128, and it automatically adjusted the network architecture and training parameters to fit this specific dataset. The preprocessing steps included resizing the input volumes to fit within the 128x128x128 voxel space, followed by applying intensity normalization and data augmentation techniques tailored to the dataset. The training process was carried out using nnU-Net's automated configuration for hyperparameters, which includes determining the appropriate learning rate, batch size, and data augmentation parameters. We implemented nnU-Net using the resources available from the official GitHub repository, which provided the necessary tools and code for our experiments [13].

C. UNeXt

UNeXt builds upon the well-established encoder-decoder structure of U-Net, integrating innovative features to enhance segmentation performance. It retains the familiar architecture but introduces a novel Tokenized MLP (Tok-MLP) block in the bottleneck. This Tok-MLP block tokenizes convolutional features and employs multilayer perceptrons (MLPs) to efficiently model and extract meaningful segmentation information. By incorporating a channel-shifting operation within the MLP layers, UNeXt enhances its ability to capture local dependencies, effectively reducing both the number of parameters and computational complexity. Additionally, skip connections between different levels of the encoder and decoder further

bolster the network's performance. Our implementation of UNeXt utilized the code and tools available from the UNeXt GitHub repository [14], which supported the development and execution of our experiments [9].

D. YOLO

YOLO (You Only Look Once) is a highly efficient object real-time object detection system. It applies a single neural network to the full image, dividing the image into regions and predicting bounding boxes and probabilities for each region. YOLO's architecture is highly efficient, allowing it to achieve high speed and accuracy. For medical image segmentation, we adapted YOLO to handle the specific requirements of lesion segmentation. This involved converting the segmentation masks into the format required by YOLO, where each object in the image is described by a set of coordinates that outline its contour. The most recent version, YOLOv10, represents the cutting-edge in this technology, but it has not yet been released for segmentation tasks. Therefore, we utilized YOLOv9 for our lesion segmentation purposes [10].

YOLOv9 has different versions tailored for specific tasks, such as YOLOv9e and YOLOv9c. The main differences between these versions lie in their design and application focus [15] :

- **YOLOv9e**: This version emphasizes enhanced accuracy and robust performance in complex detection scenarios. It integrates advanced techniques such as Programmable Gradient Information (PGI) and the Generalized Efficient Layer Aggregation Network (GELAN) to handle challenging environments with high precision.
- **YOLOv9c**: This version is optimized for computational efficiency, making it suitable for applications requiring real-time performance with limited computational resources. YOLOv9c balances accuracy and speed by employing a streamlined architecture.

We utilized the YOLOv9 variant from the ultralytics GitHub repository for our segmentation tasks [16].

IV. RESULTS AND DISCUSSION

Our evaluation of deep learning methods is based on the dataset described in I-C. We used only the training set of 15 patients, as the testing set of 38 patients was not available for download, to evaluate the performance of the algorithms. To optimize the use of the dataset, we employed cross-validation techniques. We implemented 5-fold cross-validation to ensure the robustness of our results. The dataset was split into five subsets, and in each fold, four subsets were used for training while one was reserved for validation. This process was repeated five times, ensuring that each subset served as the validation set once. This dataset, adhering to the OFSEP protocol, was used to train and test the algorithms in a manner that is representative of current medical imaging standards.

A. Performance comparison

The performance of the nnU-Net, UNeXt, YOLOv9e, and YOLOv9c models was evaluated using Dice scores for each

patient, as detailed in Table I. The results show that nnU-Net and UNeXt show superior performance, especially in more challenging segmentation scenarios, demonstrating its robustness in handling complex medical image segmentation tasks. In contrast, the YOLOv9e and YOLOv9c models exhibit similar performance, with comparable Dice scores. This suggests that while the YOLOv9 models differ in specific design features, their segmentation accuracy is notably consistent.

TABLE I: Dice comparison for each patient using different models, including average and standard deviation.

Patient ID	YOLOv9e	YOLOv9c	UNeXt	nnU-Net
01042GULE	0.75	0.76	0.84	0.84
07003SATH	0.57	0.48	0.67	0.78
08031SEVE	0.62	0.60	0.85	0.87
01016SACH	0.79	0.79	0.78	0.90
08002CHJE	0.51	0.51	0.77	0.71
08037ROGU	0.63	0.63	0.79	0.75
07010NABO	0.36	0.47	0.84	0.71
07040DORE	0.42	0.47	0.75	0.76
08027SYBR	0.32	0.35	0.75	0.70
01038PAGU	0.61	0.61	0.76	0.66
01039VITE	0.73	0.73	0.80	0.89
01040VANE	0.41	0.42	0.90	0.74
07001MOEL	0.24	0.24	0.79	0.64
07043SEME	0.65	0.64	0.83	0.76
08029IVDI	0.69	0.62	0.76	0.84
Average Dice	0.57	0.57	0.79	0.76
Standard Deviation	0.17	0.17	0.06	0.08

The following images illustrate the segmentation results for a sample patient, showcasing the differences in performance between nnU-Net, UNeXt, and YOLOv9e models. As seen in the images, nnU-Net captures the lesion boundaries with high precision (Fig.1a and Fig 1d), while UNeXt provides slightly more conservative estimates (Fig.1b and Fig.1e), and YOLOv9e tends to expand the lesion area (Fig.1c and Fig.1f).

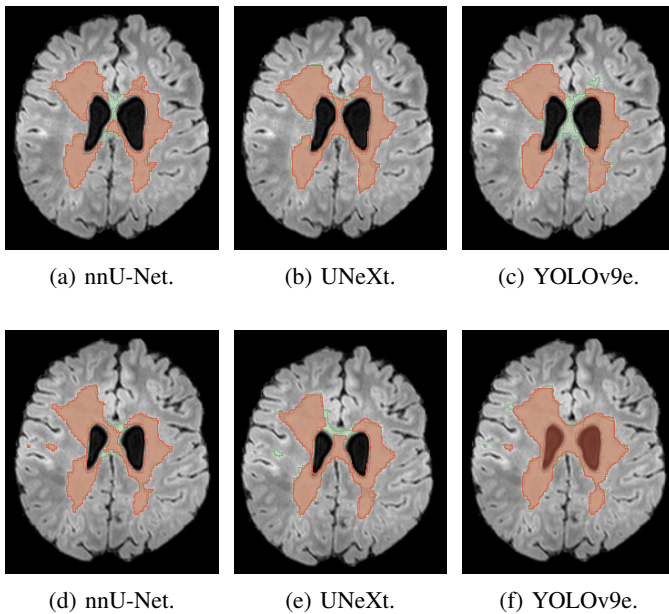


Fig. 1: Segmentation results in comparison by using nnU-Net, UNeXt, and YOLOv9e.

The comparison of segmentation performance across the nnU-Net, UNeXt, and YOLOv9 models reveals significant insights into the capabilities and limitations of each approach, particularly in the context of medical image analysis for multiple sclerosis (MS) lesion segmentation.

1) YOLOv9e's real-time capabilities with trade-offs:

- While the YOLOv9e model shows lower overall accuracy compared to nnU-Net and UNeXt, with an average Dice score of 0.57, its performance is notable given the model's design focus on real-time processing. YOLOv9e's architecture, optimized for speed, tends to expand the segmented area, leading to overestimation of lesion sizes. This behavior, while potentially useful in scenarios requiring rapid screening, may not be ideal for precise diagnostic applications.
- The higher standard deviation (0.17) observed in YOLOv9e's performance indicates less consistency, which can be a drawback in medical contexts where reliability is crucial. However, YOLOv9e's strengths lie in its potential application in environments where computational resources are limited, or where quick, approximate segmentations are acceptable as a first pass before more detailed analysis.
- The less favorable performance of YOLOv9e can partly be attributed to challenges in the preprocessing stage, particularly in handling complex lesion structures such as those with holes or perforations. Improving preprocessing methods to better address these complex structures could enhance the model's overall accuracy and reliability.

2) UNeXt's innovation and competitiveness:

- UNeXt, with an average Dice score of 0.79, marginally outperforms nnU-Net overall, demonstrating the efficacy of its novel architecture that combines convolutional layers with tokenized MLPs. This design enables UNeXt to capture both local and global features within the medical images, which is essential for detecting and accurately segmenting MS lesions that vary in size, shape, and intensity.
- The model's low standard deviation (0.06) across patient cases further emphasizes its consistency, making it a strong candidate for clinical applications where reliability is paramount. UNeXt's performance also underscores the potential of integrating MLP-based architectures into medical imaging, paving the way for future research into even more sophisticated hybrid models.
- UNeXt's performance suggests that it is particularly adept at handling datasets with a high degree of variability, such as those encountered in multi-center studies where imaging protocols and patient demographics differ. This adaptability makes it a promising tool for wide-scale deployment in diverse clinical settings.

3) nnU-Net's superior performance:

- The nnU-Net model consistently achieves the highest Dice scores across most patient cases, with an average Dice score of 0.76. This performance highlights the model's ability to adapt effectively to the complex and variable nature of medical images, particularly in cases where the anatomical structures are intricate or the lesions are small and dispersed.

- The architecture of nnU-Net, which includes extensive use of skip connections and a well-tuned 3D convolutional framework, allows for the preservation of spatial information across different scales. This ability is crucial in medical imaging, where accurate delineation of lesion boundaries can significantly impact diagnosis and treatment planning.
- The relatively low standard deviation (0.08) in nnU-Net’s performance across different patients indicates its robustness and reliability, suggesting that it can be effectively generalized to a broader range of cases, potentially improving clinical workflows and outcomes.

B. Comparison with Previous Results

To provide a comprehensive benchmark for our current study, we compared the performance of our models with the results from the 2016 MS lesion segmentation challenge dataset. This challenge involved various teams and experts who attempted to segment MS lesions, offering a valuable reference point for evaluating the effectiveness of our models. It is important to note that our results are not directly comparable to those from the MICCAI 2016 challenge teams because we had access only to the training set of 15 patients, whereas the challenge results are based on the testing set of 38 patients, which we do not have access to. Despite this limitation, our comparison provides valuable insights into the performance of our models relative to established benchmarks.

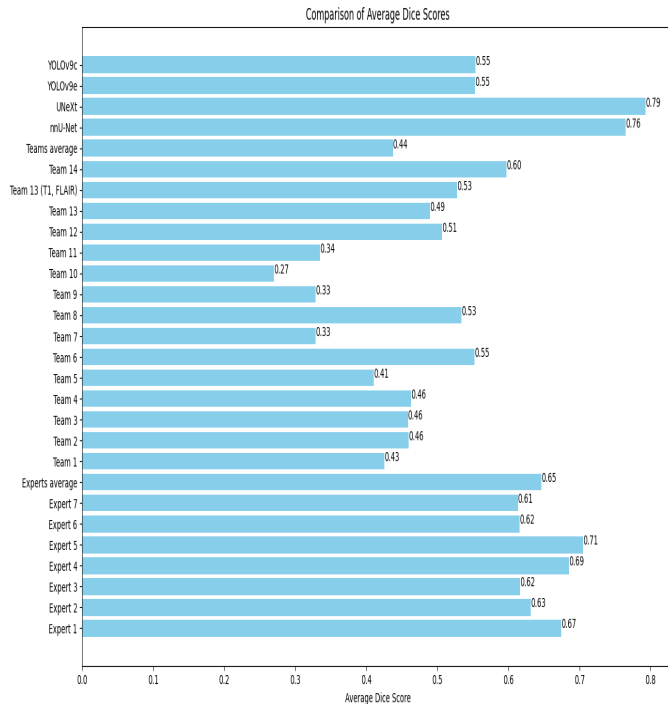


Fig. 2: Performance comparison of various teams and experts on the 2016 MS lesion segmentation challenge dataset, along with our models’ results.

The table below summarizes the average Dice scores for different groups, offering a clear benchmark for comparison.

Group	Average Dice
Teams Average	0.47
Best Team	0.64
Experts Average	0.71
Best Expert	0.78
nnU-Net	0.76
UNeXt	0.79
YOLOv9e	0.57

TABLE II: Average Dice scores for different groups in the 2016 MS lesion segmentation challenge, along with our models’ performance.

From the table, it is evident that both nnU-Net and UNeXt models significantly outperform the average scores of the teams and experts from the 2016 challenge. Specifically, UNeXt achieves the highest average Dice score of 0.79, which not only surpasses the best expert’s score from the challenge but also highlights its capability to handle the complex task of MS lesion segmentation with high accuracy. nnU-Net also demonstrates strong performance, with an average Dice score of 0.76, closely trailing UNeXt and outmatching the average expert performance.

The YOLOv9e model, while not surpassing the expert average, still demonstrates competitive performance with a score of 0.57. This result is particularly noteworthy given that YOLOv9e is designed for real-time processing, indicating its potential for fast segmentation tasks where speed is crucial, albeit at the cost of some accuracy.

Notably, two teams from the 2016 challenge [4] [17], Team 6 and Team 12, both employed models similar to U-Net and achieved average Dice scores of 0.57 and 0.54, respectively. These scores are comparable to the performance of our YOLOv9e model but fall short of the results achieved by UNeXt and nnU-Net. This comparison underscores the advancement in segmentation techniques and the effectiveness of our models in handling complex segmentation tasks.

C. Future work

In future work, we plan to extend our evaluation by applying our algorithms to the test set of the MICCAI 2026 challenge. Currently, our analysis is based on the training set of 15 patients, as the test set of 38 patients was not available for download. To ensure the robustness of our findings, we employed 5-fold cross-validation, where the dataset was divided into five subsets. In each fold, four subsets were used for training and one for validation, ensuring that every subset was used as the validation set once. Engaging with the dataset providers to gain access to additional data will be crucial for further validating and refining our approaches. It is also essential to compare our performance metrics with those of other teams.

Future advancements will focus on enhancing the 3D segmentation capabilities of our models. Although nnU-Net has demonstrated strong performance in 3D segmentation, we aim to develop specialized 3D Transformer models or improve existing architectures to better capture the volumetric context of MRI data. This enhancement could significantly advance

the segmentation of small and dispersed lesions, which current models often struggle with.

Additionally, we will explore self-supervised learning techniques to tackle the challenge of limited labeled data. By pre-training models on large amounts of unlabeled data and subsequently fine-tuning them on smaller labeled datasets, we aim to improve the robustness and accuracy of our segmentation models. This approach will enhance our models' adaptability to diverse datasets and clinical scenarios, further advancing the field of automated MS lesion segmentation.

V. CONCLUSION

This study has provided valuable insights into the performance of different deep learning models for MS lesion segmentation, highlighting both the strengths and limitations of each approach. nnU-Net and UNeXt demonstrate high potential for accurate and reliable segmentation, making them strong candidates for integration into clinical practice. YOLOv9e, while less accurate, shows promise for real-time applications where speed is essential. The findings underscore the importance of selecting the appropriate model based on the specific clinical requirements and computational resources available. Moving forward, future research should focus on improving the generalizability of these models, reducing computational demands, and incorporating more comprehensive evaluation metrics to better align with clinical needs. By addressing these limitations, the integration of deep learning into medical imaging can be further optimized, ultimately enhancing patient care and outcomes.

REFERENCES

- [1] 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*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham: Springer International Publishing, 2015, pp. 234–241.
- [2] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnu-net: a self-configuring method for deep learning-based biomedical image segmentation," *Nature methods*, vol. 18, no. 2, pp. 203–211, 2021.
- [3] S. Aslani, M. Dayan, L. Storelli, M. Filippi, V. Murino, M. A. Rocca, and D. Sona, "Multi-branch convolutional neural network for multiple sclerosis lesion segmentation," *NeuroImage*, vol. 196, pp. 1–15, 2019.
- [4] O. Commowick, A. Istace, M. Kain *et al.*, "Objective evaluation of multiple sclerosis lesion segmentation using a data management and processing infrastructure," *Scientific Reports*, vol. 8, p. 13650, 2018. [Online]. Available: <https://doi.org/10.1038/s41598-018-31911-7>
- [5] A. Dosovitskiy, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.
- [6] A. Hatamizadeh, Y. Tang, V. Nath, D. Yang, A. Myronenko, B. Landman, H. R. Roth, and D. Xu, "Unetr: Transformers for 3d medical image segmentation," in *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, 2022, pp. 574–584.
- [7] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. L. Yuille, and Y. Zhou, "Transunet: Transformers make strong encoders for medical image segmentation," *arXiv preprint arXiv:2102.04306*, 2021.
- [8] H. Cao, Y. Wang, J. Chen, D. Jiang, X. Zhang, Q. Tian, and M. Wang, "Swin-unet: Unet-like pure transformer for medical image segmentation," in *European conference on computer vision*. Springer, 2022, pp. 205–218.
- [9] J. M. J. Valanarasu and V. M. Patel, "Unext: Mlp-based rapid medical image segmentation network," in *International conference on medical image computing and computer-assisted intervention*. Springer, 2022, pp. 23–33.
- [10] C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, "Yolov9: Learning what you want to learn using programmable gradient information," *arXiv preprint arXiv:2402.13616*, 2024.
- [11] R. J., D. S., G. R., and F. A., "You only look once: unified, real-time object detection," in *proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.
- [12] <https://github.com/MIC-DKFZ/HD-BET>.
- [13] <https://github.com/MIC-DKFZ/nnUNet>.
- [14] <https://github.com/jeya-maria-jose/UNeXt-pytorch>.
- [15] Ultralytics, "The latest version of yolo for real-time object detection and segmentation," <https://docs.ultralytics.com/models/yolov9/>, 2023, accessed: Août 27, 2024.
- [16] <https://github.com/ultralytics/ultralytics>.
- [17] O. Commowick, A. Istace, M. Kain, B. Laurent, F. Leray, M. Simon, S. Camarasu-Pop, P. Girard, R. Ameli, J.-C. Ferré, A. Kerbrat, T. Tourdias, F. Cervenansky, T. Glatard, J. Beaumont, S. Doyle, F. Forbes, J. Knight, A. Khademi, and C. Barillot, "Miccai 2016 ms lesion segmentation challenge: supplementary results [data set]," 2018. [Online]. Available: <https://doi.org/10.5281/zenodo.1307653>