

Article

ZechariahNet: A Novel Method of MS Lesion Diagnosis Through MRI Images by the Combination of C-LSTM and 3D CNN Algorithms

Mahshid Dehghanpour ¹, Mansoor Fateh ^{1,*} , Zeynab Mohammadpoory ² and Saideh Ferdowsi ^{3,*} 

¹ Faculty of Computer Engineering, Shahrood University of Technology, Shahrood 3619995161, Iran; mahshid.dehghanpour@shahroodut.ac.ir

² Faculty of Electrical Engineering, Shahrood University of Technology, Shahrood 3619995161, Iran; z.mohammadpoory@shahroodut.ac.ir

³ School of Mathematics, Statistics and Actuarial Science, University of Essex, Colchester CO4 3SQ, UK

* Correspondence: mansoor_fateh@shahroodut.ac.ir (M.F.); s.ferdowsi@essex.ac.uk (S.F.)

Abstract

In light of the growing prevalence of the autoimmune disease multiple sclerosis (MS), accurate detection of MS lesions in brain magnetic resonance imaging (MRI) images plays a critical role in assisting neurologists with timely diagnosis. The high similarity between MS lesions and normal brain tissues, however, makes this task particularly challenging. Although numerous deep-learning-based approaches have been proposed for the automatic segmentation of MS lesions, the method presented in this study has achieved superior results. ZechariahNet is a U-Net-based architecture that integrates transition down blocks, squeeze-attention (SA) blocks, dense blocks, and Convolutional LSTM (C-LSTM) blocks within a 3D CNN framework. By jointly exploiting spatial-temporal information from three consecutive MRI slices (previous, current, and subsequent) and strategically applying C-LSTM modules across the encoder and decoder paths, the proposed model effectively captures the neighborhood dependencies for enhanced feature extraction and reconstruction. These architectural innovations significantly improve segmentation accuracy, enabling ZechariahNet to achieve a dice similarity coefficient (DSC) of 84.72%, outperforming existing state-of-the-art methods.

Keywords: multiple sclerosis; MRI; 3D CNN; C-LSTM; deep learning

1. Introduction

Multiple sclerosis (MS) is a chronic, inflammatory, and neurodegenerative disease of the central nervous system (CNS), which is characterized by demyelinating plaques in the brain and spinal cord. Accurate segmentation of these lesions from magnetic resonance imaging (MRI) is crucial for a reliable quantification of lesion numbers, monitoring disease progression, and evaluating treatment response. Among MRI modalities, fluid-attenuated inversion recovery (FLAIR) imaging is particularly effective for this purpose, as it suppresses signals from cerebrospinal fluid and enhances the visibility of white matter lesions [1]. Figure 1 illustrates a sample of a FLAIR image from an MS patient in the ISBI2015 dataset, alongside the corresponding expert-annotated lesion mask.



Academic Editor: Iman Dehzangi

Received: 14 December 2025

Revised: 9 January 2026

Accepted: 13 January 2026

Published: 15 January 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

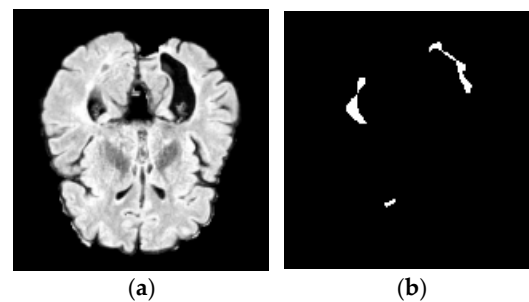


Figure 1. (a) A sample of an MRI FLAIR image from the ISBI2015 dataset. (b) The annotation of the lesion mask.

Traditionally, the identification of MS lesions has relied on a manual assessment by expert radiologists, who visually inspect MRI slices and detect lesion regions based on intensity differences. Although this approach can obtain precise results, it requires substantial time and is often inconsistent due to differences both between experts and within repeated assessments by the same expert. The subjectivity of manual interpretation often leads to inconsistencies across experts and sessions, particularly when lesions are small or have unclear boundaries. These drawbacks have motivated a growing interest in automated methods for MS lesion segmentation. Early automated techniques were mainly based on classical machine learning algorithms, such as k-means clustering, Gaussian mixture models, and random forests. These methods extracted handcrafted features related to voxel intensity, spatial context, or texture, which were used for lesion classification. Despite their initial success, such approaches heavily depended on expert knowledge and dataset-specific parameter tuning, limiting their generalization across scanners, imaging protocols, and clinical centers. Moreover, they often are not able to distinguish true lesions from noise or artifacts, especially in heterogeneous or low-contrast MRI data [2].

More advanced artificial intelligence methods, such as deep learning methods, revolutionized medical image analysis by allowing models to learn directly from the data [3,4]. Convolutional neural networks (CNNs) became the dominant method for automated segmentation tasks, due to their ability to learn hierarchical and task-specific representations without requiring handcrafted features. Architectures such as U-Net and its numerous variants have achieved promising performance in biomedical image segmentation by combining global contextual information with fine-grained spatial detail through encoder-decoder designs and skip connections. However, despite these advances, the existing CNN-based models still face significant challenges in MS lesion segmentation, including difficulty in detecting small or low-contrast lesions, sensitivity to intensity variations, and limited generalization across multi-center datasets [5]. The existing approaches face significant challenges such as insufficient accuracy in detecting multiple sclerosis (MS) plaques, even when evaluated using relatively favorable Dice score criteria, and the absence of segmentation algorithms that achieve an acceptable Dice score. In this work, we aim to address these limitations by proposing a novel segmentation framework called ZechariahNet. Given the limitations outlined above, and considering the substantial similarity between MS plaques and normal brain tissue, as well as the tendency of these plaques to adhere to one another, accurately determining changes in the size and volume of each white plaque is extremely challenging. As a result, reliably determining changes in the size, volume, and number of white matter plaques across successive follow-up scans is nearly impossible for neurologists by using visual inspection alone.

The innovation of this research lies in the segmentation of MS plaques of three consecutive MRI slices (concluding previous, current, and subsequent) with higher accuracy than previous works, resulting in a more accurate diagnosis of the disease and a Dice score

of 84.72%. This increase was achieved by combining the set of transition down blocks, dense blocks, and SA blocks with the C-LSTM block, using 3D CNN algorithms in the proposed network (ZechariahNet). In this model, the C-LSTM blocks are exactly located after each SA block (vertically and horizontally in the architecture). We apply C-LSTM modules vertically across the encoder path exactly after each SA block, to capture the neighborhood dependencies to reach advanced feature extraction, and we apply C-LSTM blocks across the decoder paths horizontally to capture neighborhood dependencies for better reconstruction.

The rest of this paper is organized as follows. In the next section, we review the most recent studies in this area. Section 3 describes the materials and methods. Sections 4 and 5 present the experiments and results, and the discussion, respectively. Finally, the last section provides the conclusion.

2. Literature Review

This section presents a comprehensive review of deep learning approaches proposed for the segmentation of MS lesions. Aslani et al. [6] proposed an automated method for MS lesion segmentation from multi-modal brain MRI using a deep end-to-end 2D CNN designed for slice-based processing of 3D data. The network features a multi-branch down-sampling path to independently encode modality-specific information, along with multi-scale feature fusion blocks that integrate features from different modalities at various levels. In the upsampling path, multi-scale feature upsampling blocks are used to recover the spatial and shape information of lesions. The model was trained using axial, sagittal, and coronal images to exploit contextual information in all directions. Evaluations on the ISBI 2015 dataset achieved a Dice value of 76%.

Zhang et al. [7] proposed ALL-Net, a cascaded 3D U-Net followed by 2D anatomical convolutional modules that aggregate 2D predictions into a 3D segmentation mask. The model achieved an overall Dice score of 63.9% on the ISBI 2015 dataset. Notably, it introduced a lesion-wise loss function and distance-map channels to enhance the detection of small lesions, which are often overlooked by traditional voxel-wise loss functions.

Hashemi et al. [8] proposed the segmentation of MRI FLAIR and T2 images by using a modified U-Net and Attention U-Net, proposing the fusion of the masks obtained from a better segmentation of FLAIR and T2, and could obtain a Dice score of 81%.

Sarica et al. [9] developed a Dense-Residual U-Net enhanced with Attention Gates (AG), Efficient Channel Attention (ECA), and an Atrous Spatial Pyramid Pooling (ASPP) module. Using a three-plane 2D fusion strategy, their method achieved a mean Dice of 0.6688 on the ISBI 2015 dataset. This architecture effectively combines dense and residual blocks to increase the network depth while attention mechanisms and the ASPP bottleneck refine feature maps and suppress false positives.

Rondinella et al. [10] proposed a framework that combines U-Net with convolutional long short-term memory and attention modules to segment and quantify lesions in MRI. The method outperforms the existing approaches, achieving an average 84% DSC and demonstrating robustness on the ISBI2015 dataset.

Wahlig et al. [11] examined the effectiveness of transfer learning for improving multiple sclerosis (MS) lesion segmentation, using limited training data. The authors employed a private MRI dataset of 149 patients (2014–2021) and trained 3D convolutional neural networks for three segmentation tasks: (1) single-timepoint FLAIR lesion segmentation, (2) new lesion detection between timepoints, and (3) enhancing lesion segmentation on post-contrast T1 images. The best performance was achieved in the enhancing lesion segmentation task when using transfer learning, with a lesion-wise sensitivity of 0.74 and PPV of 0.69, significantly outperforming the de novo model (sensitivity = 0.44, PPV = 0.37).

Krishnan et al. [12] developed a 3D multi-arm U-Net for T2 lesion segmentation, trained on a large dataset from relapsing MS clinical trials. Its generalization was tested on various datasets, achieving a mean DSC ≥ 0.66 and sensitivity of ≥ 0.72 internally, and 0.62 and 0.68 externally. Their results indicated that performance decreased for smaller lesions.

Zhang et al. [13] introduced a generalizable MS lesion segmentation model based on the standard U-Net architecture without structural modifications. The authors proposed a test-time self-ensembled lesion fusion (SELF) strategy, which achieved a Dice score of 68.2% on the ISBI 2015 MS segmentation challenge. The method demonstrated robustness across different ensemble parameters and superior generalization performance on clinical datasets acquired from various scanners. Additionally, the study showed that instance normalization outperformed batch normalization in terms of cross-dataset generalization.

Zhu et al. [14] conducted a comparative study on MS lesion segmentation, using four advanced deep learning architectures: YOLOv9e, YOLOv9c, UNeXt, and nnU-Net. The models were trained and evaluated, using the publicly available MICCAI 2016 MS lesion segmentation challenge dataset to ensure standardized benchmarking. Among the tested methods, YOLOv9e and YOLOv9c achieved Dice scores of 0.57, indicating moderate performance for coarse lesion localization. In contrast, UNeXt and nnU-Net achieved higher accuracies, with Dice scores of 0.79 and 0.76, respectively, demonstrating their superior ability to capture detailed lesion boundaries and spatial characteristics.

Pishvai et al. [15] integrated brain parcellation into a diffusion-based MS lesion segmentation framework and evaluated it on the ISBI 2015 dataset. While parcellation did not substantially improve the average segmentation accuracy, it enhanced robustness by reducing variability across brain regions and improved sensitivity to small lesions. The method achieved a Dice score of 74.14%.

Hindawi et al. [16] conducted a multi-center study across three clinical sites, in which a self-configuring nnU-Net was trained within a federated learning (FL) framework. The resulting cross-site ensemble achieved Dice scores ranging from 0.66 to 0.80 on held-out datasets, demonstrating that privacy-preserving FL can attain state-of-the-art performance despite scanner heterogeneity and the absence of data sharing.

Andishgar et al. [17] proposed R2AUNet, a 3D U-Net architecture enhanced with recurrent residual blocks and dual attention gates for FLAIR-only lesion segmentation. Trained on 112 MRI scans from 95 patients, the network achieved a Dice Similarity Coefficient of 0.768 on the test set, with precision and recall scores of 0.825 and 0.765, respectively. These results highlight the model's effectiveness in capturing spatial dependencies and refining lesion boundaries, particularly in cases of small or low-contrast lesions.

Gaj et al. [18] used transfer learning techniques by fine-tuning a pre-trained model with a limited number of labeled scans; the proposed approach aimed to improve segmentation accuracy. The methodology involved subject-specific fine-tuning, where the model was adapted using the first scan of each subject to achieve better segmentation results in subsequent scans. The proposed method could achieve a DSC of 79.2% on the ISBI 2015 dataset.

Belwal et al. [19] proposed a modified U-Net architecture for MS lesion segmentation using the ISBI 2015 dataset. The model achieved high accuracy and robust performance, with a Dice score of 77.56% on the training set and 73.98% on the test set, surpassing previous benchmarks. Additional metrics included accuracy $\sim 99.3\%$, precision 73–74%, and recall 91–94%, demonstrating the model's effectiveness in accurately detecting and delineating MS lesions. The results highlight the potential of the modified U-Net for improving MS diagnosis and management.

Zhang et al. [20] proposed UNISELF, a deep learning method for automated MS lesion segmentation that balances high in-domain accuracy with strong out-of-domain

generalization. Trained on the ISBI 2015 longitudinal MS dataset, UNISELF uses test-time self-ensembled lesion fusion and instance normalization to handle domain shifts and missing contrasts. The model achieved a Dice score of 66.4% and PPV of 88.2%, outperforming benchmark methods on both the ISBI test set and multiple out-of-domain datasets, including MICCAI 2016, UMCL, and a private multisite dataset.

Table 1 provides a concise overview of the representative MS lesion segmentation methods reported in the literature, including studies evaluated on different datasets.

Table 1. Summary of some MS lesion segmentation studies.

Study	Year	Network	Dataset
Aslani et al. [6]	2019	2D CNN multi-branch fusion	ISBI 2015
Zhang et al. [7]	2021	ALL-Net (3D U-Net + 2D AC modules)	ISBI 2015
Hashemi et al. [8]	2022	Modified U-Net + Attention U-Net (FLAIR + T2 fusion)	Custom dataset
Sarica et al. [9]	2023	Dense-Residual U-Net + AG + ECA + ASPP	ISBI 2015
Rondinella et al. [10]	2023	U-Net + ConvLSTM + Attention	ISBI 2015
Wahlig et al. [11]	2023	3D CNN + Transfer Learning	Private dataset (149 subjects)
Krishnan et al. [12]	2023	Multi-arm U-Net	ISBI 2015
Zhang et al. [13]	2024	SELF + U-Net (no structural modification)	ISBI 2015
Zhu et al. [14]	2024	YOLOv9e/YOLOv9c/UNeXt/nnU-Net	MICCAI 2016
Pishvai et al. [15]	2024	Brain parcellation	ISBI 2015
Hindawi et al. [16]	2025	Federated Learning nnU-Net	Multi-center
Andishgar et al. [17]	2025	R2AUNet (Rec. Residual + Dual Attention)	112 scans
Gaj et al. [18]	2025	Transfer Learning (subject-specific fine-tuning)	ISBI 2015
Belwal et al. [19]	2025	Modified U-Net	ISBI 2015
Zhang et al. [20]	2025	UNISELF (self-ensembled, inst. norm.)	ISBI 2015 + external sets

Overall, the recent work on MS lesion segmentation shows a movement from standard U-Net models toward architectures that use attention, 3D information, transfer learning, and even federated learning. Although these methods have improved performance, challenges such as the ongoing need for higher segmentation accuracy, scanner variability, limited annotated data, and poor cross-center generalization are still major challenges. With these gaps in mind, the next section describes our proposed approach and explains how it aims to improve the robustness and reliability of lesion segmentation.

3. Materials and Methods

In this section, the dataset and proposed method are described in Sections 3.1 and 3.2, respectively. Section 3.2 comprises two subsections, Sections 3.2.1 and 3.2.2, presenting our preprocessing approach and the proposed ZechariahNet network architecture.

3.1. Image Dataset

In this study, we utilized a subset of the publicly available ISBI2015 Challenge dataset. ISBI2015 was presented at the Longitudinal MS Lesion Segmentation Challenge [21], which was arranged in association with the 2015 ISBI conference. The complete dataset comprises 19 patients' MR images, which were collected at multiple timepoints using a 3.0 T scanner; however, only 5 patients with the appropriate segmentation were available. The five chosen patients participated in data collection at various timepoints. Each acquisition consists of

the original MR images, together with corresponding segmentation masks indicating the MS lesions. Each scan includes 4 distinct image sequences: T1-weighted, T2-weighted, PD-weighted, and FLAIR. Since MS lesions in white matter appear clearer and more hyperintense on FLAIR images than on other imaging sequences, we only employed FLAIR images in our experiments. There are 181 images in each sequence of Flair, with a size of 181×217 .

3.2. Proposed Method

In this study, we first preprocess the data, and then these data are fed to the proposed network for training. After training, the performance of the model is evaluated by using test images. Figure 2 presents the block diagram of the proposed method.

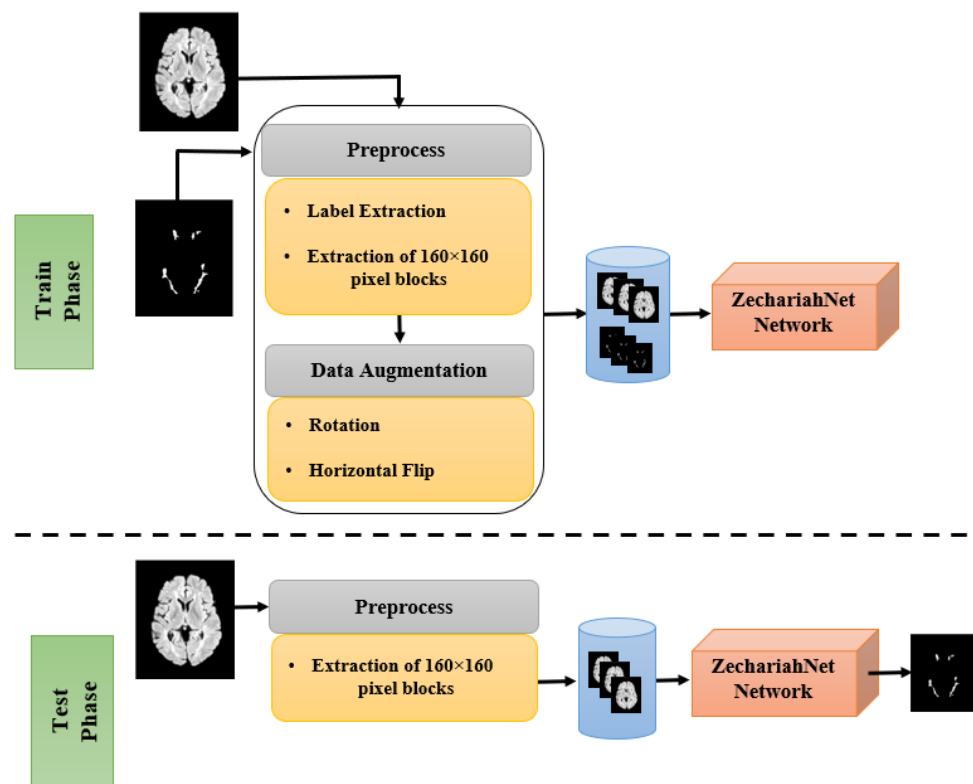


Figure 2. The block diagram of the proposed method (ZechariahNet).

3.2.1. Preprocessing

At this stage, we performed several data augmentation approaches to strengthen robustness and increase the generalization of the model. The augmentation includes rotation and a horizontal flip to add variability while maintaining important spatial relationships by changing the scale of images [22]. After augmentation, non-clinically irrelevant regions were removed by cropping areas outside of the brain, and all images and their corresponding masks were then resized to 160×160 pixels to ensure consistency across the dataset. This approach is consistent with that adopted by Hashemi et al. [8] and Rondinella et al. [10]. Figure 3 shows two MRI slices before and after preprocessing.

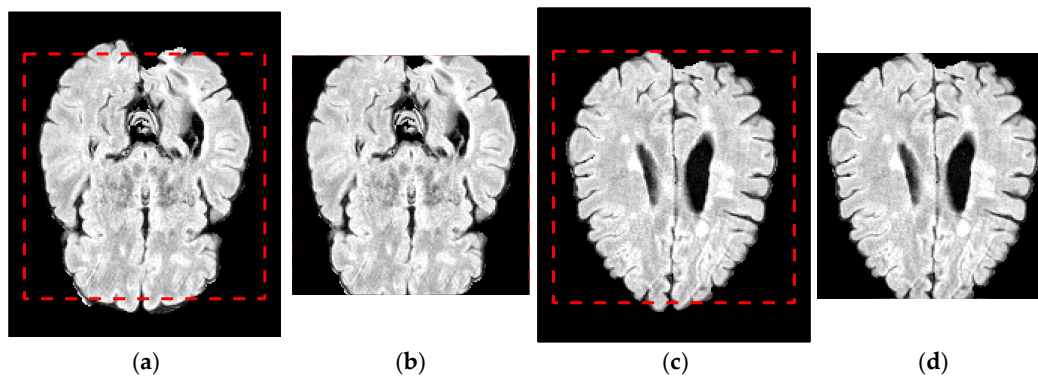


Figure 3. A phase of preprocessing: trimming the area of the black border. (a,c) are original slices, while (b,d) are preprocessed ones. The dashed red box indicates the trimming border.

3.2.2. Proposed Network Architecture

Reducing image dimensionality may result in the loss of important information. To minimize this loss, improve accuracy, and enhance processing robustness, the proposed algorithm considers both the preceding and subsequent slices when evaluating each image simultaneously. The traditional U-Net architecture is strengthened and expanded upon in our proposed model. It is customized for MS lesion detection through FLAIR MRI images, which are reinforced and expanded upon in our proposed model, ZechariahNet.

In this study, the proposed method is derived from an improvement of the classical version of the U-Net architecture [23,24], with a special focus on segmentation of MS lesions from MRI FLAIR images to achieve higher accuracy in MS plaque detection. As shown in Figure 4, the proposed architecture utilizes an enhanced encoder–bottleneck–decoder structure. This customized architecture consists of a $3 \times 3 \times 3$ convolution block on the input side and a 1×1 convolution block on the output side. The following combination is used on the encoder side. In the first stage, we have a convolutional block, dense blocks, and SA modules [25,26]. From the second stage, each step has been promoted by integrating transition down blocks, dense blocks, and SA modules.

- **The dense blocks** extract hierarchical, multi-scale spatial features, promoting feature reuse and mitigating the vanishing gradient issues. This enables the network to learn complex and discriminative representations of lesion boundaries.
- **The SA blocks** apply both channel-wise and spatial attention mechanisms to emphasize the most informative features and suppress irrelevant ones. In particular, the attention mechanism introduces a pixel-group attention process, enabling the model to focus on groups of neighboring pixels that likely belong to the same lesion region. This is particularly important in MS lesion segmentation, where the lesions often appear as small, irregular, and poorly contrasted regions.
- **The transition down blocks** reduce the spatial resolution of the feature maps while increasing their depth, allowing deeper layers to process more abstract information.

Among the stages of the encoder side, we have added C-LSTM [27,28] units after each SA block (in both vertical and horizontal paths) to further enhance feature modeling. These recurrent convolutional units allow the network to integrate contextual dependencies along both spatial dimensions, leading to a more robust and feature-aware representation of lesion patterns. Although the primary C-LSTM unit is located in the bottleneck to capture inter-slice correlations, similar units can also be added in the decoding path to reinforce contextual recovery during upsampling, potentially improving spatial consistency in the reconstructed masks.

- **The C-LSTM blocks** are used to model inter-slice spatial dependencies across consecutive MRI slices. Unlike traditional 2D U-Net architectures that process each slice independently, the integration of C-LSTM allows the model to exploit contextual information from neighboring slices, improving the coherence and continuity of lesion detection. This approach leverages the sequential nature of MRI scans and leads to better generalization and higher segmentation accuracy. These additions enable the network to capture rich spatial patterns as well as sequential and directional dependencies, which are essential for accurately detecting small and irregular MS lesions.

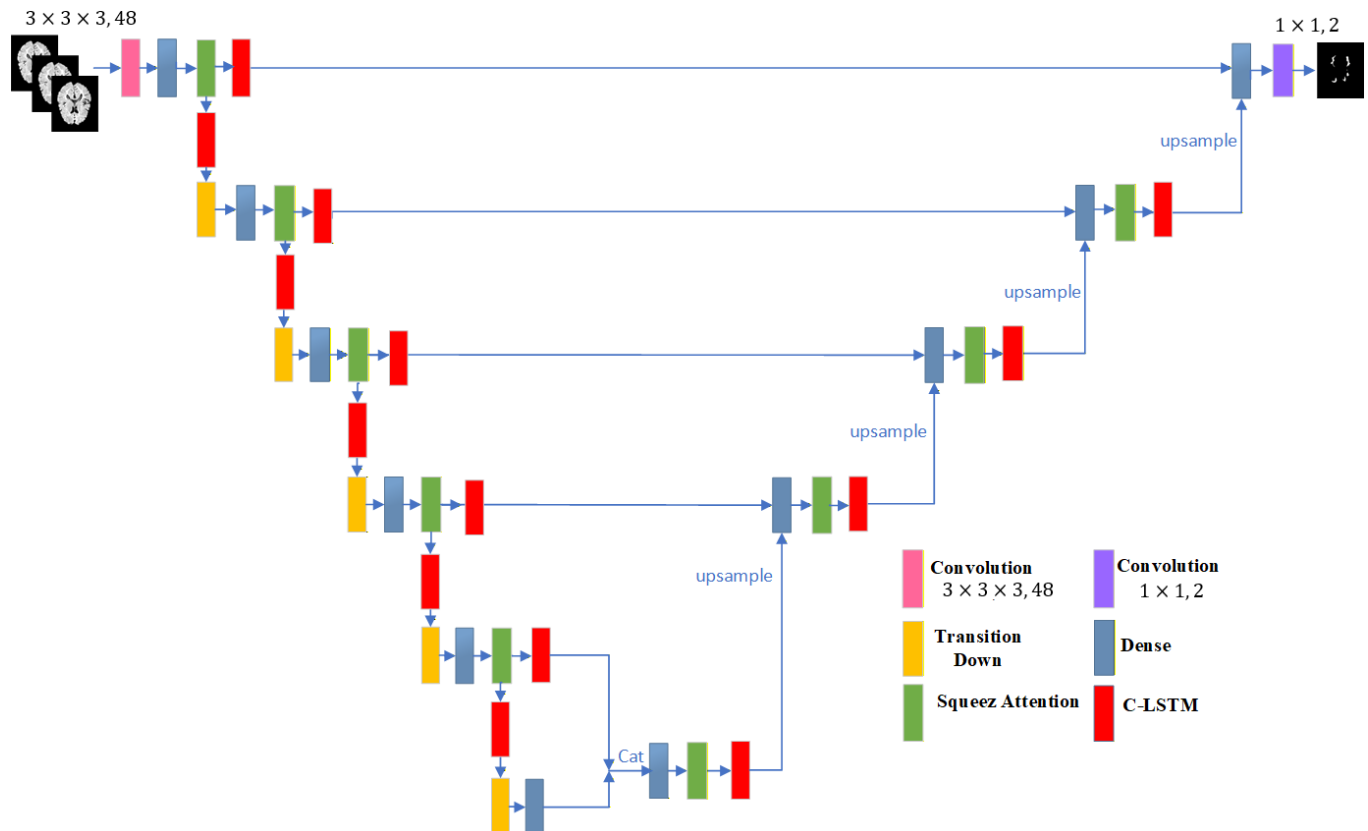


Figure 4. The architectural view of ZechariahNet. This hybrid model combines transition down blocks, dense blocks, SA blocks, and C-LSTM blocks.

In the proposed architecture, C-LSTM modules are deployed at multiple network levels, with two distinct modes of information propagation: horizontal and vertical. Horizontally propagated C-LSTM modules transfer information across parallel network paths, enabling enhanced neighborhood-aware feature interaction. Vertically propagated C-LSTM modules forward information to subsequent network stages, reinforcing spatial-temporal neighborhood dependencies across successive representations.

These two propagation modes serve different functional purposes and exhibit different learning behaviors. Horizontal C-LSTMs primarily focus on strengthening neighborhood-aware feature extraction by facilitating contextual interactions across parallel representations, while vertical C-LSTMs aim to preserve and propagate spatial-temporal contextual information across stages. As a result, separate C-LSTM modules are employed to effectively model these distinct behaviors. Importantly, this design choice does not introduce a significant increase in model complexity; the dominant computational overhead is mainly attributed to the use of the 3D CNN backbone.

Although the network processes consecutive frames using slice-based representations, three-dimensional context is explicitly captured through 3D convolutional kernels that operate across both spatial and inter-slice dimensions. Additionally, C-LSTM modules further reinforce spatiotemporal dependencies by modeling neighborhood relationships between successive feature representations.

At the input stage, three consecutive FLAIR slices are provided in grayscale. A conventional layer expands the number of channels to extract low-level visual information such as edges, textures, and intensity transitions. The resulting feature maps are then passed through the encoder, which consists of five hierarchical levels. Each encoder level consists of a dense block that extracts multi-scale spatial features, followed by an SA module that highlights the most informative regions while suppressing irrelevant background responses. Immediately after each SA block, the network applies two directional C-LSTM units: one processing features horizontally and the other vertically. These recurrent units allow the model to integrate the spatial context along both axes, helping it to better understand the elongated or irregular shapes of MS lesions. The transition down block then reduces the spatial resolution while increasing the feature depth, allowing deeper layers to learn more abstract representations.

The bottleneck represents the most compressed and high-level stage of the network. It includes a dense block for deep feature extraction, followed by a C-LSTM that models inter-slice dependencies across the neighboring MRI slices. Unlike a standard 2D U-Net that treats each slice independently, this sequential modeling improves the coherence and continuity of lesion appearance across slices, leading to more stable segmentation results.

The decoder mirrors the encoder in structure, progressively restoring spatial resolution through transition up blocks. At each level, skip connections bring in the corresponding encoder features, helping the decoder to recover fine-grained lesion boundaries that may have been lost during down-sampling. Similarly to the encoder, each decoder level also contains a dense block followed by an SA module, and again, two directional C-LSTM units are placed immediately after each SA block. These C-LSTM units reinforce spatial consistency during reconstruction and help the network to maintain the structural coherence of lesions when upsampling. Finally, a convolutional layer followed by a Softmax activation produces the pixel-wise segmentation map of the central slice.

Compared to the U-Net, the proposed architecture provides a richer and more context-aware feature representation. The addition of SA modules allows for the model to focus more effectively on lesion-relevant regions. Dense blocks promote feature reuse and improve gradient flow, making the network easier to train and more expressive. The use of C-LSTM units, both directional ones at every level and the inter-slice unit at the bottleneck, enables the model to learn patterns that extend across space and across slices: something that traditional U-Net cannot achieve. Overall, the model becomes more robust in identifying subtle, low-contrast MS lesions and more accurate in reconstructing their shapes and boundaries.

This combined design creates a sequentially enhanced, attention-guided U-Net variant that captures both local texture details and broader spatial relationships, leading to precise and reliable MS lesion segmentation in FLAIR MRI scans.

4. Experiments and Results

Our experiments were implemented in Python (version 3.9.7) with the PyTorch 2.1.0 library, using a Windows-based operating system. For the implementation of our deep-learning-based architecture, six NVIDIA V100 GPUs, each equipped with 40 GB of memory, were used. A five-fold cross-validation strategy was adopted in this study. In Fold 1, 1119 images were used for training, 197 images for validation, and 70 images for testing. In

Fold 2, the dataset was split into 1119 training images, 183 validation images, and 84 test images. For Fold 3, 1119 images were assigned for training, 200 for validation, and 67 for testing. In Fold 4, 1136 images were used for training, 208 images for validation, and 42 images for testing. Finally, in Fold 5, 1111 images were allocated for training, 225 images for validation, and 50 images for testing. This fold-wise partitioning ensures a fair and reliable evaluation of the proposed model while preventing data leakage.

While training the proposed network, a dropout rate was applied to prevent overfitting and improve generalization. Also, the Adam optimizer was employed for parameter optimization due to its adaptive learning capability and efficient handling of sparse gradients [29]. The hyperparameters used in this study are presented in Table 2. As mentioned before, in the first stage, MRI images from the ISBI2015 dataset were preprocessed. Then, they were given to the proposed network for MS plaque segmentation. Due to the limited number of patients involved in the study (five individuals), as well as the small number of scans (totaling just 21), the evaluation process was executed similarly to the method employed by Hashemi et al. [8]. Specifically, a 5-fold cross-validation strategy was implemented. The model achieving the highest Dice score on the validation set for each fold was then selected for all subsequent tests. Additionally, data augmentation was applied to the training samples in each folder.

Table 2. List of hyperparameters used in this study.

Hyper-Parameters	Values
Batch size	12
Drop out	0.2, 0.1, 0
Learning rate	3×10^{-4}
Weight decay	3×10^{-4}
Learning rate decay	0.99
Optimizer	Adam
Number of parameters	13 M
Number of epochs	165
Training computation time	3.0 h

Since the network receives the target slice, along with its previous and next slices, as input, special care was taken to ensure that the augmentation process did not disrupt this ordering, and all three slices were augmented in a coordinated way, so their alignment and order stayed correct.

Model evaluation was carried out by comparing the predicted segmentation masks with the reference masks provided by a clinical expert. Dice score, sensitivity, specificity, accuracy, PPV, and NPV were used as evaluation metrics, as defined in Equations (1)–(6) [30].

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (1)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (2)$$

$$\text{Accuracy} = \frac{TN + TP}{TN + TP + FN + FP} \quad (3)$$

$$\text{DSC} = \frac{2TP}{2TP + FP + FN} \quad (4)$$

$$PPV = \frac{TP}{TP + FP} \quad (5)$$

$$NPV = \frac{TN}{TN + FN} \quad (6)$$

- True positive (TP) is the number of pixels that the model correctly labeled as an MS lesion.
- True negative (TN) is the number of pixels that the model correctly labeled as normal tissue.
- False positive (FP) is the number of pixels that the model incorrectly labeled as an MS lesion.
- False negative (FN) is the number of pixels that the model incorrectly labeled as normal tissue.

Figure 5 shows the training and validation Dice scores and loss values across different epochs. As can be seen, the training curves gradually improve, while the validation curves follow a very similar trend, indicating that the network is learning effectively without overfitting. The small gap between the training and validation metrics demonstrates that the model generalizes well to unseen data. In addition, the fluctuations in the early epochs decrease over time, reflecting the stabilization of the learning process. Overall, the number of chosen epochs, 165, is sufficient for convergence, as both the training and validation metrics show minimal oscillations towards the end of training. These observations confirm that the selected training procedure and the number of epochs is appropriate for the task.

All metrics for the five test folds, along with their mean, are presented in Figure 6. Evaluation of the performance metrics across the different folds shows that the model achieves the best Dice score of 91.37% in Fold 4. Additionally, using this fold, the model demonstrates sensitivity, specificity, PPV, NPV, and accuracy of 91.37%, 89.32%, 99.98%, 96.64%, 99.98%, and 99.95%, respectively. The lowest Dice score is obtained by Fold 5, which is recorded at 81.18%. However, this score is higher than those reported in many previous works. In general, the performance measures remain robust and comparable across all folds.

Furthermore, the proposed method obtains average $\pm$ standard deviation values for the Dice score, sensitivity, specificity, PPV, NPV, and accuracy of $84.72\% \pm 4.3$, $81.17\% \pm 4.64$, $99.91\% \pm 0.1$, $90.41\% \pm 4.86$, $99.91\% \pm 0.09$, and $99.88\% \pm 0.11$, respectively, across all folds. Metrics with values close to 100% naturally show very low variance, indicating stable and consistent performance across the folds. In contrast, lesion-level metrics exhibit a higher standard deviation, which is expected due to the heterogeneous distribution of MS lesions across different patients and slices. Variations in lesion size, number, and spatial location introduce a natural variability between the folds. This behavior is commonly observed in MS lesion segmentation tasks and reflects the intrinsic difficulty of accurately segmenting small and sparse lesions, rather than the instability of the model.

Figure 7 provides a visual representation of the segmentation results of the proposed approach on the test set. This figure presents five axial slices from five patients and ground truth segmentations alongside the segmentations produced by our proposed architecture. False positive and false negative pixels are highlighted in red and green, respectively. As can be seen, the predicted lesion masks closely match the ground truth, demonstrating that the proposed method accurately segments the majority of lesions. The visualization of false positive and false negative pixels further confirms that the model effectively detects most MS lesions. Additionally, Figure 7 shows that our approach can accurately segment even very small MS plaques, highlighting its high precision in lesion segmentation.

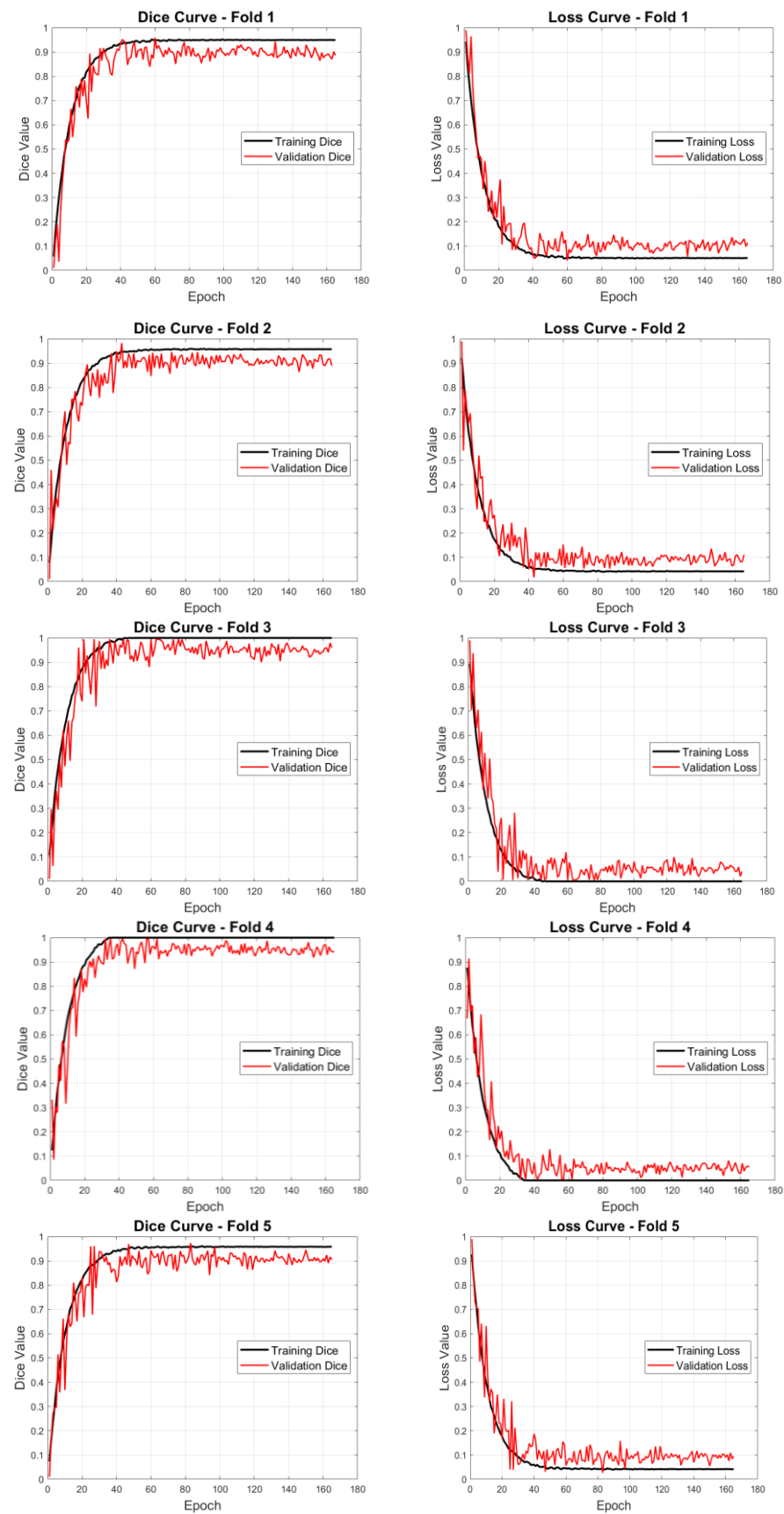


Figure 5. Dice (first column) and loss (second column) values obtained from training and validation phases for Fold 1 to Fold 5, arranged from top to bottom, respectively. These results were obtained after 165 epochs.

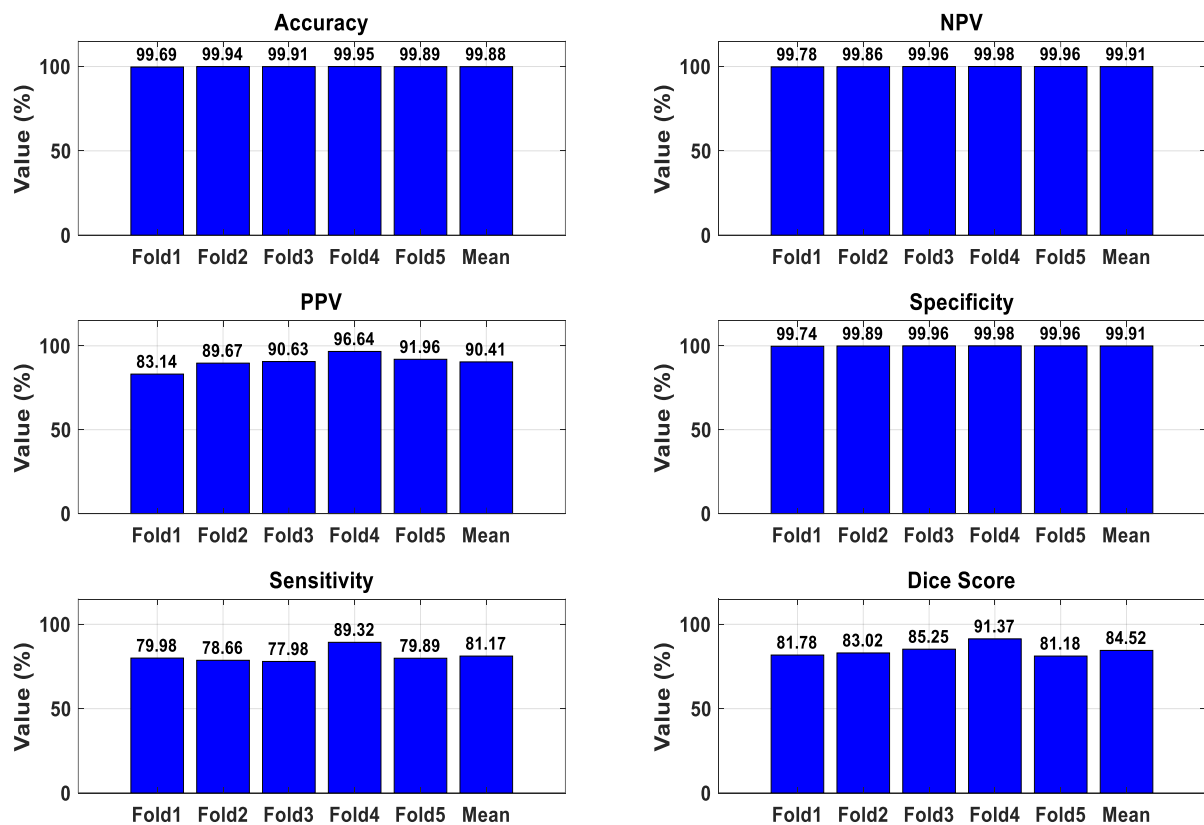


Figure 6. Average values of the evaluation metrics across five folds.

Ablation Study

To evaluate the contribution of each component in the proposed ZechariahNet architecture, a comprehensive ablation study was conducted. We started by training a baseline model that combines U-Net with a 3D CNN, allowing the network to capture the volumetric context across adjacent MRI slices. This improved the model's ability to detect plaques, highlighting the importance of 3D information for MS lesion segmentation. The Dice score obtained at this stage was 0.61. Then, we added dense blocks to the network. These blocks improve feature interaction and help information flow more smoothly across layers. With them, the model became even better at detecting small and confluent plaques. The Dice score obtained at this stage was 0.64. After that, we introduced the SA block. This module helps the network to focus on the lesion while ignoring irrelevant background noise, which improves lesion localization. By adding the SA modules, the Dice score increased to 0.84. Finally, we incorporated a C-LSTM module after each SA block. This allowed the network to better understand spatial dependencies and contextual relationships across feature maps. As a result, the model could distinguish MS plaques from the surrounding white matter more accurately. By incorporating the C-LSTM modules, the Dice score improved and reached 0.8472.

Overall, our ablation study shows that every component we added contributed positively to the model's performance and confirms that the ZechariahNet architecture is highly effective for accurate MS plaque segmentation. Table 3 summarizes these results and shows the Dice scores at each stage of the ablation study.

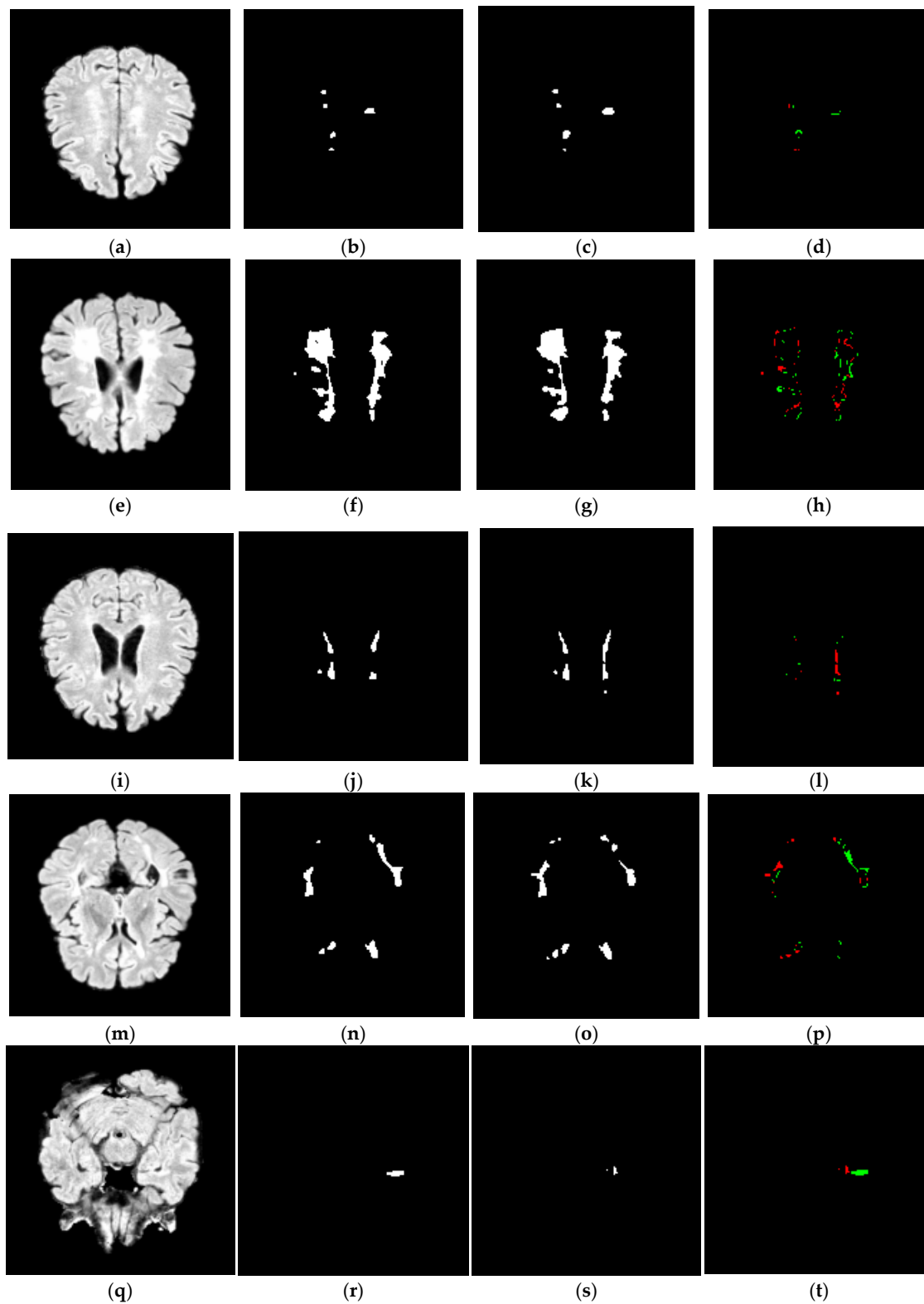


Figure 7. (a,e,i,m,q) Five axial slices of five subject, (b,f,j,n,r) ground-truths for the corresponding slices, (c,g,k,o,s) results of the proposed architecture, and (d,h,l,p,t) false-positive (in red) and false-negative (in green) pixels of the proposed model, respectively.

Table 3. Results of the ablation study in terms of Dice score.

Stage	Dice Score
U-Net + 3D CNN	61%
Dense blocks + U-Net + 3D CNN	64%
SA + dense blocks + U-Net + 3D CNN	84%
ZechariahNet (C-LSTM + SA + Dense blocks + U-Net + 3D CNN)	84.72%

5. Discussion

The proposed architecture, ZechariahNet, is an improved version of the classical U-Net, designed specifically for segmenting MS lesions in FLAIR MRI images. The network incorporates dense blocks, SA modules, and C-LSTM layers within an encoder–bottleneck–decoder framework. These components help the model capture both spatial and sequential patterns, retain fine details, and focus on the most relevant features, leading to accurate and context-aware segmentation. Performance evaluation across different folds shows that the model achieves a maximum Dice score of 91.37% in Fold 4, along with high sensitivity, specificity, PPV, NPV, and accuracy. Also, average metrics across all folds are promising with a Dice score of 84.72%, indicating the model’s reliability and ability to segment MS plaques.

Our results show that ZechariahNet provides robust and accurate segmentation of MS lesions, using the ISBI2015 dataset. The model’s promising performance is due to several key factors. First, the dense connections allow for effective feature reuse across layers, enabling the network to capture subtle intensity differences. Second, the use of C-LSTM layers helps the model to learn spatial dependencies between consecutive slices, which is crucial for maintaining lesion continuity in 3D space. Third, attention mechanisms enable the network to focus on relevant regions while suppressing background noise and reducing false positives. These combined design choices are highly effective in detecting small and low-contrast lesions, as the network can highlight subtle features while preserving spatial context and inter-slice continuity.

Compared with classical U-Net and recent CNN- and U-Net-based approaches on the ISBI2015 dataset, ZechariahNet achieves superior segmentation performance.

Table 4 presents a comparison between the results of state-of-the-art techniques and those of the proposed method, including both the overall average and the results for Fold 4. For a fair and consistent quantitative comparison, this table only reports performance metrics for methods evaluated on the ISBI 2015 dataset. Sensitivity, PPV, and Dice score are reported to provide a comprehensive evaluation of lesion detection performance.

The results shown in Table 4 highlight that our framework improves performance metrics compared to existing approaches. Notably, our method achieves a Dice score increase of approximately 0.72% from the best results reported in previous studies.

Although our proposed method achieved promising results, a few limitations remain. MRI scans can vary depending on the scanner, imaging settings, and noise levels. Such variations can reduce the accuracy of plaque detection and affect the reliability of results in clinical practice. Additionally, most previous studies have only utilized image features and investigated how plaque patterns relate to a patient’s symptoms or disease progression. By integrating imaging and clinical data, our system can help clinicians to improve their decision-making and treatment planning.

Table 4. The comparison of the proposed method with other state-of-the-art techniques.

Method	Year	Method	Dice Score	PPV	Sensitivity
Aslani et al. [6]	2019	CNN	76%	--	--
Zhang et al. [7]	2021	All Net	63.90%	--	--
Hashemi et al. [8]	2022	U-Net	81%	84%	79%
Sarica et al. [9]	2023	Dense-Residual U-Net	66%	86%	--
Rondinella et al. [10]	2023	FC-DenseNet + SA + C-LSTM	84%	85%	84%
Krishnan et al. [12]	2023	Multi-arm U-Net	67.70%	86.50%	--
Zhang et al. [13]	2024	U-Net	68.2	--	--
Pishvai et al. [15]	2024	Brain parcellation	74.14%	--	75.95%
Gaj et al. [18]	2025	transfer learning techniques	79.20%	--	79%
Bewal et al. [19]	2025	U-Net	73.98%	--	94%
Zhang et al. [20]	2025	UNISELF	66.40%	88.2%	--
Proposed method (Mean)	2025	ZechariahNet	84.72%	90.41%	81.17%
Proposed method (Fold 4)	2025	ZechariahNet	91.37%	96.64%	89.32%

Another important requirement is the availability of larger and more diverse datasets to ensure the model generalizes across different patient populations. We acknowledge that the current dataset is relatively small, which may increase the risk of overfitting despite using data augmentation and regularization techniques during training. We also plan to expand ZechariahNet for application to other brain disorders that present similar biomarkers in MRI. Owing to its flexible design, including dense blocks, SA modules, and C-LSTM layers, the model can be adapted to different medical imaging tasks and may be useful for detecting brain tumors, Alzheimer’s disease, and other conditions that require accurate MRI analysis.

Our trained system can be installed as a software tool in clinics, where clinicians can upload patient MRI scans, have them automatically analyzed, and receive segmented plaques to support quicker and more reliable decisions.

6. Conclusions

MS diagnosis depends mainly on correctly identifying brain plaques in MRI scans and evaluating the patient’s clinical symptoms. However, this process can be difficult because MS lesions often resemble nearby brain tissue. In this study, we introduced ZechariahNet, an enhanced U-Net-based architecture for MS lesion segmentation using axial FLAIR MRI images. ZechariahNet demonstrated strong and reliable performance across all five folds, using the ISBI2015 dataset. The model achieved its best performance in Fold 4, with a Dice score of 91.37%, and corresponding sensitivity, specificity, PPV, NPV, and accuracy of 89.32%, 99.98%, 96.64%, 99.98%, and 99.95%, respectively. Overall, the model achieved average metrics of 84.72% for Dice score, 81.17% for sensitivity, 99.91% for specificity, 90.41% for PPV, 99.91% for NPV, and 99.88% for accuracy across all folds, confirming its robustness and general segmentation capability. Furthermore, our architecture improved lesion segmentation compared with classical U-Net and several recent CNN- and U-Net-based approaches on the same dataset. In particular, the method achieved a 0.72% improvement in Dice score, relative to the best previously published model on ISBI2015, highlighting the effectiveness

of combining dense connections, attention mechanisms, and sequential learning. Overall, our findings demonstrate that deep-learning-based systems such as ZechariahNet can play an important role in assisting clinicians by providing accurate, automated MS lesion segmentation, supporting more reliable diagnosis and clinical decision-making. For future work, we plan to evaluate ZechariahNet on additional MRI datasets, integrate clinical and patient-specific data to enhance diagnostic relevance, and extend the model to other neurological conditions.

Author Contributions: M.D.: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing—original draft, Visualization. M.F.: Conceptualization, Validation, Writing—review and editing, Project administration, Supervision. Z.M.: Conceptualization, Validation, Writing—review and editing, Project administration, Supervision, Visualization. S.F.: Conceptualization, Writing—review and editing, Validation, Supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The dataset and the generated code can be obtained from the corresponding authors upon request without any additional restrictions.

Acknowledgments: During the preparation of this manuscript, the authors used (ChatGPT-5.2), only to improve grammar and readability. After that, all authors reviewed the contents, and they take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kamraoui, R.A.; Mansencal, B.; Manjón, J.V.; Coupé, P. Longitudinal detection of new multiple sclerosis lesions using deep learning. *Front. Neuroimaging* **2022**, *1*, 948235. [\[CrossRef\]](#)
2. Alsahanova, N.; Bartenev, P.; Sharaev, M.; Ljubisavljevic, M.; Mansoori, T.A.; Statsenko, Y. Integrating radiomics with deep learning enhances multiple sclerosis lesion delineation. *arXiv* **2025**, arXiv:2506.14524. [\[CrossRef\]](#)
3. Khajeha, H.R.; Fateh, M.; Abolghasemi, V. Diagnosis of glaucoma using multi-scale attention block in convolutional neural networks and data augmentation techniques. *Eng. Rep.* **2024**, *6*, e12866. [\[CrossRef\]](#)
4. Saber, A.; Fateh, A.; Parhami, P.; Siahkarzadeh, A.; Fateh, M.; Ferdowsi, S. Efficient and accurate pneumonia detection using a novel multi-scale transformer approach. *Sensors* **2025**, *25*, 7233. [\[CrossRef\]](#)
5. Alrabai, A.; Echtioui, A.; Hamida, A.B. Multiple sclerosis segmentation using deep learning models: Comparative study. In Proceedings of the 6th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Sfax, Tunisia, 24–27 May 2022; IEEE: New York, NY, USA, 2022; pp. 1–6. [\[CrossRef\]](#)
6. Aslani, S.; Dayan, M.; Storelli, L.; Filippi, M.; Murino, V.; Rocca, M.A.; Sona, D. Multi-branch convolutional neural network for multiple sclerosis lesion segmentation. *NeuroImage* **2019**, *196*, 1–15. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Zhang, H.; Zhang, J.; Li, C.; Sweeney, E.M.; Spincemaille, P.; Nguyen, T.D.; Gauthier, S.A.; Wang, Y.; Marcille, M. ALL-Net: Anatomical information lesion-wise loss function integrated into neural network for multiple sclerosis lesion segmentation. *NeuroImage Clin.* **2021**, *32*, 102854. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Hashemi, M.; Akhbari, M.; Jutten, C. Delve into multiple sclerosis lesion exploration: A modified attention U-Net for MS lesion segmentation in brain MRI. *Comput. Biol. Med.* **2022**, *145*, 105402. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Sarica, B.; Seker, D.Z.; Bayram, B. A dense residual U-Net for multiple sclerosis lesion segmentation from multi-sequence 3D MR images. *Int. J. Med. Inform.* **2023**, *170*, 104965. [\[CrossRef\]](#)
10. Rondinella, A.; Crispino, E.; Guarnera, F.; Giudice, O.; Ortis, A.; Russo, G.; Di Lorenzo, C.; Maimone, D.; Pappalardo, F.; Battiato, S. Boosting multiple sclerosis lesion segmentation through an attention mechanism. *Comput. Biol. Med.* **2023**, *161*, 107021. [\[CrossRef\]](#)
11. Wahlig, S.G.; Nedelec, P.; Weiss, D.A.; Rudie, J.D.; Sugrue, L.P.; Rauschecker, A.M. 3D U-Net for automated detection of multiple sclerosis lesions: Utility of transfer learning from other pathologies. *Front. Neurosci.* **2023**, *17*, 1188336. [\[CrossRef\]](#)
12. Krishnan, A.P.; Song, Z.; Clayton, D.; Jia, X.; De Crespigny, A.; Carano, R.A. Multi-arm U-Net with dense input and skip connectivity for T2 lesion segmentation in clinical trials of multiple sclerosis. *Sci. Rep.* **2023**, *13*, 4102. [\[CrossRef\]](#)
13. Zhang, J.; Zuo, L.; Dewey, B.E.; Remedios, S.W.; Pham, D.L.; Carass, A.; Prince, J.L. Towards an accurate and generalizable multiple sclerosis lesion segmentation model using self-ensembled lesion fusion. In Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI), Athens, Greece, 27–30 May 2024; IEEE: New York, NY, USA, 2024; pp. 1–5. [\[CrossRef\]](#)

14. Zhu, Y.; Grenier, T.; Revol-Muller, C. Comparative analysis of advanced deep learning algorithms for multiple sclerosis lesion segmentation in FLAIR MRI. In Proceedings of the IEEE International Conference on Signal Processing (ICSP), Suzhou, China, 28–31 October 2024; IEEE: New York, NY, USA, 2024; pp. 697–702.
15. Pishvai, D.S.; Rondinella, A.; Guarnera, F.; Battiato, S. Advantages of brain parcellation in multiple sclerosis lesion segmentation. In Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Lisbon, Portugal, 3–6 December 2024; IEEE: New York, NY, USA, 2024; pp. 6646–6653. [\[CrossRef\]](#)
16. Hindawi, S.; Szubstarski, B.; Boernert, E.; Wuerfel, J.; Tackenberg, B. Federated learning for lesion segmentation in multiple sclerosis: A real-world multi-center feasibility study. *Front. Neurol.* **2025**, *16*, 1620469. [\[CrossRef\]](#)
17. Andishgar, A.; Rismani, M.; Mohammadi, Z.; Bazmi, S.; Fathian, M.; Shafaei, M.E.; Tajali, M.; Khazforoosh, S.; Sajjadi, L.; Goodarzi, E.; et al. Automated multiple sclerosis lesion segmentation from 3D-FLAIR MRI using R2AUNet. *Mult. Scler. Relat. Disord.* **2025**, *102*, 106620. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Gaj, S.; Thoomukuntla, B.; Ontaneda, D.; Nakamura, K. Subject-based transfer learning in longitudinal multiple sclerosis lesion segmentation. *J. Neuroimaging* **2025**, *35*, e70024. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Belwal, P.; Singh, S. Multiple sclerosis lesion segmentation using MRI: A U-Net approach. *Procedia Comput. Sci.* **2025**, *260*, 513–522. [\[CrossRef\]](#)
20. Zhang, J.; Zuo, L.; Dewey, B.E.; Remedios, S.W.; Liu, Y.; Hays, S.P.; Pham, D.L.; Mowry, E.M.; Newsome, S.D.; Calabresi, P.A.; et al. UNISELF: A unified network with instance normalization and self-ensembled lesion fusion for multiple sclerosis lesion segmentation. *arXiv* **2025**, arXiv:2508.03982. [\[CrossRef\]](#)
21. Carass, A.; Roy, S.; Jog, A.; Cuzzocreo, J.L.; Magrath, E.; Gherman, A.; Button, J.; Nguyen, J.; Prados, F.; Sudre, C.H.; et al. Longitudinal multiple sclerosis lesion segmentation: Resource and challenge. *NeuroImage* **2017**, *148*, 77–102. [\[CrossRef\]](#)
22. Xu, M.; Yoon, S.; Fuentes, A.; Park, D.S. A comprehensive survey of image augmentation techniques for deep learning. *Pattern Recognit.* **2023**, *137*, 109347. [\[CrossRef\]](#)
23. Moeskops, P.; Viergever, M.A.; Mendrik, A.M.; De Vries, L.S.; Benders, M.J.; Išgum, I. Automatic segmentation of MR brain images with a convolutional neural network. *IEEE Trans. Med. Imaging* **2016**, *35*, 1252–1261. [\[CrossRef\]](#)
24. Du, G.; Cao, X.; Liang, J.; Chen, X.; Zhan, Y. Medical image segmentation based on U-Net: A review. *J. Imaging Sci. Technol.* **2020**, *64*, 20508-1–20508-11. [\[CrossRef\]](#)
25. Hasan, N.; Bao, Y.; Shawon, A.; Huang, Y. DenseNet convolutional neural networks application for predicting COVID-19 using CT images. *SN Comput. Sci.* **2021**, *2*, 389. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; IEEE: New York, NY, USA, 2017; pp. 1492–1500.
27. Nasrolahzadeh, M.; Mohammadpoory, Z.; Haddadnia, J. Weighted visibility graph-based deep complex network features: New diagnostic spontaneous speech markers of Alzheimer’s disease. *Phys. D Nonlinear Phenom.* **2025**, *476*, 134681. [\[CrossRef\]](#)
28. Mohammadpoory, Z.; Nasrolahzadeh, M.; Amiri, S.A.; Haddadnia, J. A non-invasive approach for early Alzheimer’s detection through spontaneous speech analysis using deep visibility graphs. *Cogn. Comput.* **2025**, *17*, 42–58. [\[CrossRef\]](#)
29. Ghonchi, H.; Foulsham, T.; Ferdowsi, S. Assessing neural patterns of anxiety using deep learning: An EEG study. In Proceedings of the 32nd European Signal Processing Conference (EUSIPCO), Lyon, France, 26–30 August 2024; IEEE: New York, NY, USA, 2024; pp. 1596–1600.
30. Ye, J.; Dao, S.D.; Wu, Y.; George, Y.; Nguyen-Duc, T.; Schmidt, D.F.; Shi, H.; Chong, W.; Cai, J. New multiple sclerosis lesion segmentation via calibrated inter-patch blending. In *Medical Image Computing and Computer Assisted Intervention—MICCAI 2025*; Springer: Berlin/Heidelberg, Germany, 2025; Part XVI; pp. 365–375. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.