Deep Learning Methods for Miltiple Sclerosis Lesions Segmentation on MRI Data

Timur Kabulov
Department of Cell Biology and Plant Bioengineering
Belarusian State University
Minsk, Belarus
kabulovtimur64@gmail.com

Abstract. **Multiple sclerosis is a dangerous human disease that has no etiological cure. Diagnosis of the disease in its early stages involves MRI and analyzing the resulting images for the presence of lesion foci to identify the stage of the disease. The aim of the work described in the paper is to train a model using modern deep learning and data preprocessing methods for automatic segmentation of multiple sclerosis lesions on MRI data.**

Keywords: **multiple sclerosis, MRI, image segmentation, CNN**

I. INTRODUCTION

Multiple sclerosis is a human autoimmune disease, often seen in young adults, in which the immune system attacks neurons, contributing to the destruction of the cells myelin sheaths. The destruction of myelin sheaths leads to vision problems, impaired walking or balance, impaired mental clarity, numbness or weakness, especially in the legs and arms, and depression [1].

Diagnosis of this disease necessarily includes a stage of examination with MRI. The images obtained using this method must localize the foci of multiple sclerosis lesions in the brain in order to make a diagnosis and determine the severity of the disease [2]. This process is difficult and time-consuming, as the foci are often small areas that are difficult to distinguish from normal tissue.

II. TRAINING

A. Convilution Neural Netwroks with 3D convolution

To automate the segmentation of multiple sclerosis (MS) foci, we employ deep learning-based convolutional neural networks (CNNs), which have become the standard for medical image analysis due to their ability to learn hierarchical features directly from data [3]. Among CNN architectures, U-Net has emerged as a leading solution for segmentation tasks, particularly in medical imaging, owing to its symmetric encoder-decoder structure and skip connections that preserve spatial context while recovering fine-grained details [4].

In this work, we utilize a 3D U-Net architecture, which extends the original 2D U-Net by replacing 2D convolutions with volumetric (3D) convolutions. This modification allows the model to process 3D MRI scans as holistic volumetric structures, capturing spatial relationships along all three axes (height, width, depth). For MS segmentation, this is critical: 3D convolutions explicitly model the spatial distribution of lesions across adjacent slices, improving detection of small or irregularly shaped foci that might be overlooked in slice-by-slice 2D analysis [5]. Additionally, 3D context helps reduce false positives by leveraging anatomical consistency across planes.

However, 3D CNNs introduce significant computational challenges. The number of trainable parameters increases dramatically compared to 2D counterparts [6].

B. Dataset and pre-processing methods

The training data were MRI images that contained multiple sclerosis foci as well as masks for the corresponding lesion areas. The dataset incorporated images acquired through three distinct MRI sequences, each providing unique contrast mechanisms to highlight different tissue properties:

- T1-weighted imaging: optimal for visualizing anatomical structures and detecting gray matter lesions;
- T2-weighted imaging: highly sensitive to fluid-filled lesions, making it indispensable for MS diagnosis;
- FLAIR (Fluid-Attenuated Inversion Recovery): Particularly effective in suppressing cerebrospinal fluid (CSF) signals, thereby improving the visibility of periventricular and cortical lesions.

All images were initially categorized into groups, where each group was a series of images of a single patient taken at intervals. 256 images were used for dataset creation:

- training dataset is composed of 204 images;
- 32 were used for validation;
- remaining images were used for 20.

The images for one patient were exclusively found in one of the datasets.

Pattern Recognition and Information Processing (PRIP2025) :
Proceedings of the 17th International Conference, 16–18 Sept. 2025, Minsk, Belarus. – Minsk : UIIP NASB, 2025. – 496 p. – ISBN 978-985-7198-20-7.

To mitigate slow model training and standardize data, we implement the following optimizations all MRI volumes were resampled and cropped to a uniform size of 160×160×96 voxels. The transverse (axial) plane consisted of 96 slices, ensuring sufficient coverage of the brain along the inferior-superior axis. The sagittal and coronal planes were standardized to 160 slices each, preserving anatomical alignment and resolution.

To avoid overfitting additional data augmentations was used. It includes vertical and horizontal flipping of the image, random changes in brightness and contrast and adding gaussian noise.

Prior to training, the data underwent normalization to translate pixel values into the range $[0; 1]$. After normalization gaussian blur and contrast enhancing were applied for denoising and better distinction between the lesion and normal brain nerve tissue respectively. Contrast enhancing was performed with specific 3D kernel for each image in our dataset.

C. Training and evalutation

The model was trained on an NVIDIA GPU 5070 Ti (with CUDA 12.9). Training required ~48 hours for 50,000 iterations (batch size=2), with inference time of ~1 second per 3D scan (160×160×96 voxels) on the same hardware.

Training was performed on 120 epochs using the Adam optimizer [6]. A combined loss function was used:

$$L_{comb} = \lambda L_{BCE} + (1 - \lambda) L_{Dice}, \quad (1)$$

where L_{comb} is the combined loss function, λ is the coefficient that determines the contribution of binary cross entropy to the combined loss function, L_{BCE} is the binary cross entropy, and L_{Dice} is the Dice loss function.

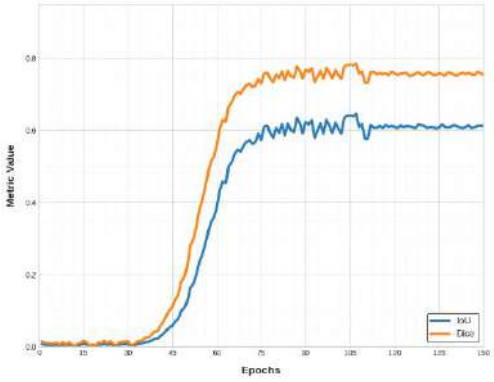

Fig. 1. IoU and Dice metrics progression during training

To train the model, a search for suitable hyperparameters using the grid search method. The most appropriate value for the learning coefficient was 0.0001, followed by a 10-fold decrease in this value with no growth of the model metrics over 5 epochs.

Fig. 1 shows the charts of changes in model's learning metrics, where maximum score was gained. Due to data size, it's complexity and difference between MRI scan types in first epochs model struggles to learn patterns from data. Only after 30-40 epochs metrics are starting to grow significantly.

III. Results and Discussion

After training, the model was evaluated on the validation and test samples where it performed at 0.78 and 0.72 on the Dice metric respectively. Fig. 2 represents, that the lowest score model commonly gets on FLAIR images, which are the smallest type of images in size in dataset.

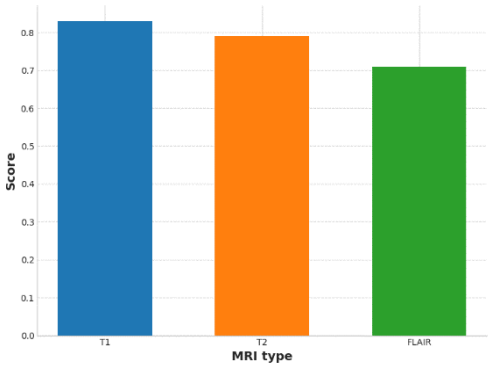

Fig. 2. Dice metrics across main types of MRI images

Thus, the paper presents a method to automate the recognition of multiple sclerosis on MRI images. With the current results, the model can already be applied in medical institutions to assist physicians in making a diagnosis.

References

[1] World Health Organization, Multiple Sclerosis [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/multiple-sclerosis. Accessed: Apr. 26, 2025.

[2] J. H. Simon et al., "Standardized MR imaging protocol for multiple sclerosis: Consortium of MS centers consensus guidelines," Amer. J. Neuroradiol., vol. 27, no. 2, 2006, pp. 455–461.

[3] Z. Li et al., "A survey of convolutional neural networks: Analysis, applications, and prospects," IEEE Trans. Neural Netw. Learn. Syst., vol. 33, no. 12, 2022, pp. 6999–7019.

[4] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," arXiv:1505.04597 [cs.CV], 2015.

[5] Ö. Çiçek et al., "3D U-Net: Learning dense volumetric segmentation from sparse annotation," arXiv:1606.06650 [cs.CV], 2016.

[6] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv:1412.6980 [cs.LG], 2017.

Pattern Recognition and Information Processing (PRIP2025) :
Proceedings of the 17th International Conference, 16–18 Sept. 2025, Minsk, Belarus. – Minsk : UIIP NASB, 2025. – 496 p. – ISBN 978-985-7198-20-7.