Automation of the Study of Radiologically Isolated Syndrome in Multiple Sclerosis

Ivan Kosik
FAMCS of
Belarusian State University
LICT of Belarusian State
Medical University
Minsk, Belarus
ivankosik91@gmail.com

Alexander Nedzved
FAMCS of
Belarusian State University
UIIP of NAS Belarus
Minsk, Belarus
ORCID 0000-0001-6367-5900

Ryhor Karapetsian
Laboratory of Information and
Computer Technology of
Belarusian State Medical
University
Minsk, Belarus
greg.itlab@gmail.com

Vera Yashina
Federal Research Center
"Computer Science and
Control" of RAS
Moscow, Russia
werayashina@gmail.com

Igor Gurevich
Federal Research Center
"Computer Science and
Control" of RAS
Moscow, Russia
igourevi@ccas.ru

Abstract. **In this paper the UNet 3+ model is used for detection regions of multiple sclerosis on radiological images. For increase quality the specific image preprocessing improves quality of dataset and results of detection. The proposed solution for the automatic identification of pathological areas using artificial neural networks has significantly increased the speed of analyzing the state of the pathological pattern.**

Keywords: **medical image analysis, UNet 3+, regions detection, segmentation, dataset preprocessing**

I. Introduction

Multiple sclerosis is a serious disease of the central nervous system that leads to disability among people, including young people of working age. Insufficient knowledge of the pathogenesis of this disease and an increase in the frequency of its occurrence require the intensification of studies of this pathology [1].

The most important symptom of multiple sclerosis is a focal lesion of the central nervous system caused by auto aggression against myelin proteins in the brain and spinal cord. In this case sites of demyelination occur and sclerotic plaques appear in the small veins of the brain, in the cerebellum, in the spinal cord, in the optic and other cranial nerves. They are an important diagnostic feature of multiple sclerosis.

Magnetic resonance imaging (MRI) is the most effective tool for visualization of demyelination sites nowadays. However, for an objective description of the state and dynamics of the pathological process, visualization of foci on MRI sections should be supplemented with data on the size, intensity and localization [2].

Detection of lesions is the first step towards obtaining additional quantitative information from MRI images. In most cases, their boundaries are blurred and have poor contrast against the background of the brain tissue. Therefore, the detection of pathological sites is the most time consuming and at the same time the most important stage. The accuracy of the diagnostics largely depends on the accuracy of the detection. Meanwhile, in medical practice, segmentation of foci is carried out either by manual methods (contouring), or by semi-automatic methods (for example, by algorithms of area growth or "smart brush"). The diagnostician must first visually assess the information content of the image area in terms of the level of brightness and localization, and then carry out the recognition procedure. Considering that the radiologist has to analyze at least 120 slices (50 + 40 + 30 in three orthogonal projections), it is easy to understand how difficult this work is. This affects the accuracy of the results, especially when examining multifocal patterns [3].

The use of advanced methods of volumetric reconstruction of MRI images, based on fully automatic segmentation of informative objects by neural networks, contributes to an increase in the productivity and accuracy of the study.

II. Preparation of Training Images

We used MRI-series of 50 patients obtained in different modes (T1, T2, Flair, etc.) for training the neural network. In addition, data augmentation was

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept 2021, Minsk, Belarus – Minsk : UIIP NASB, 2021 – 246 p – ISBN 978-985-7198-07-8

used. It is the technique for generation additional training data from the initially available initial set of images [4,5]. In our case, the initial training set consisted of 455 3D-series. It corresponds to different modes and different studies of 50 patients of the 9th City Clinical Hospital of Minsk.

For the augmentation the Albumentations library was used. We chose the following parameters for augmentation:

- rotations up to 20 degrees;
- shift;
- scaling;
- horizontal reflection;
- vertical reflection;
- sharpening;
- spectrum change;
- optical distortion.

To increase the volume of initial data, we used methods based on geometric and brightness variability for informative objects in the original images. Methods of the first type increas the data volume by reorienting and scaling the available input images. This allowed the network to learn invariance to this kind of distortion, even if the distortion was absent in the original images.

III. DEFINITION OF CNN MODEL

For the most complete automation of the procedure for segmentation of pathological areas on MRI scans, a selection of the architecture of a convolutional neural network (CNN) and a strategy for its training was performed, which would allow obtaining a high return with a limited initial sample. The learning strategy is based on the most complete extraction and effective generalization of meaningful information concentrated in the initial data. The effect is achieved due to the automatic generation of additional data from this data. The network architecture consists of a tapered section for capturing context, and a symmetrical expanding section for more accurate localization and contouring of objects. This network organization is called U-Net 3+ [6] like as Fig 1.

The UNet 3+ gives simplified overviews of UNet, UNet++. Compared with UNet and UNet++, UNet 3+ combines the multi-scale features by re-designing skip connections as well as utilizing a full-scale deep supervision, which provides fewer parameters but yields a more accurate position-aware and boundary-enhanced segmentation map [7, 8].

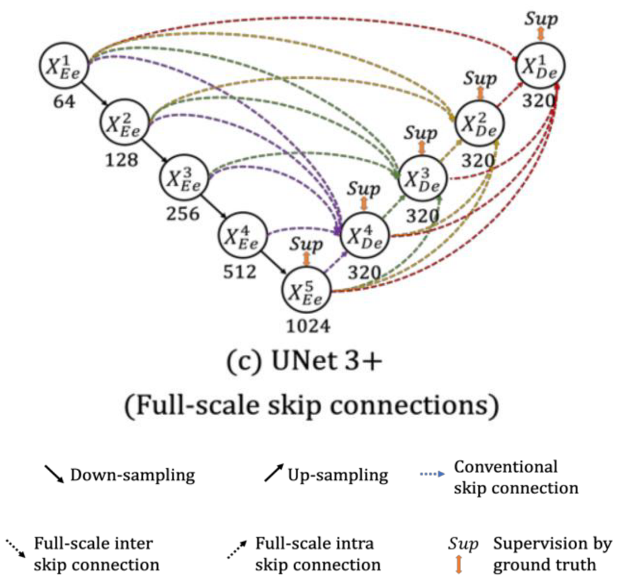


Fig .1. The model of UNet 3+ [1]. The depth of each node is presented below the circle

IV. NEURAL NETWORK TRAINING

The series were divided into training set and validation one in a ratio of 75% to 25%, respectively. The formation of sets of training and validation images by series allows you to include all MRI slices of the series in only one sample (training or validation). This is justified by the fact that similar regions belonging to neighboring slices will lead to overfitting if they get in different samples. An additional set of test images for a separate small group of patients is formed during the testing phase of the final version of the neural network.

The next step in the training procedure is to obtain training images (MRI slices) from the 3D series, divided into training set and validation one. Sections without pathological regions are not used in training, since preliminary tests showed that training on all sections gives worse results. Examples of the formation of training sets are given below. It is clear that the number of series selected for the same MRI scan mode exceeds the number of unique patients, since several series (taken at different times) can belong to the same patient.

Two neural networks were trained: one on MRI series obtained in the T2 TSE scan mode; the second on all images (all modes were used). The amount of training data is shown in Table I.

TABLE I. AMOUNT OF TRAINING DATA FOR EACH MODEL

Model	Number of 3D-series		Number of slices	
	Trainin g	***Validation***	***Training***	***Validation***
Mode T2 TSE	119	40	1500	506
All MRI-modes	341	114	4789	1452

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept 2021, Minsk, Belarus – Minsk : UIIP NASB, 2021 – 246 p – ISBN 978-985-7198-07-8

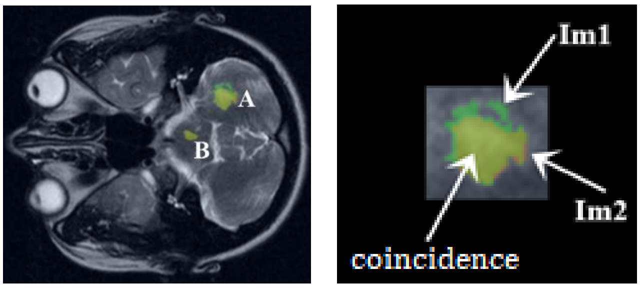
V. Quality Assessment of Network

To assess the quality, the Intersection Over Union metric (or Jaccard index) was used, given by the formula (1): (1):

$$IoU = \frac{TP}{TP+FP+FN} = \frac{Im_1 \cap Im_2}{Im_1 \cup Im_2} = \frac{Im_1 \cap Im_2}{Im_1 + Im_2 - Im_1 \cap Im_2}, \quad (1),$$

where *TP* (true positive) is selection of a pixel that actually belongs to the focus of demyelination; *FP* (false positive) is false selection of a pixel that does not belong to the focus; *FN* (false negative) is false marked pixel as not belonging to the focus; *Im1* is an area identified by an expert as a focus of demyelination; *Im2* is area identified by the neural network as a focus of demyelination.

The metric can take values from 0 to 1. It assesses how closely the area of the real focus Im1 (selected interactively by the expert) matches the area Im2, segmented automatically by the neural network. The higher the metric value, the greater the coincidence, and, therefore, the more reliable the neural network model works. In Figure 1, the same focus is highlighted by an expert (green) and a neural network (brown).



a) b)

Fig. 2. Jind metric demonstration: total image of MRI-section, with highlighted foci A and B (a); fragment of the MRI image with focus A (b)

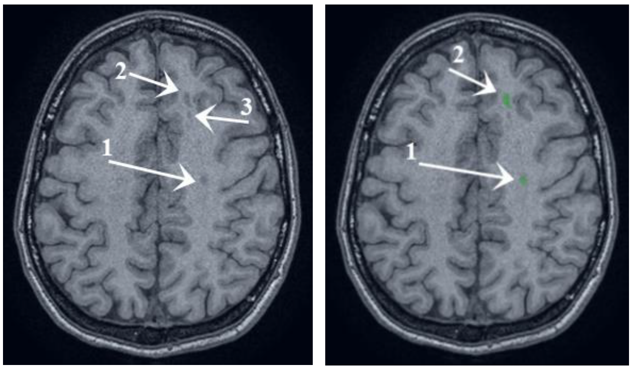
There is an overlap of areas of green and brown in Figure 1b. In Figure 1b, the overlapping areas of green and brown correspond to the area selected by both the expert and the neural network.

VI. Discussion and Conclusion

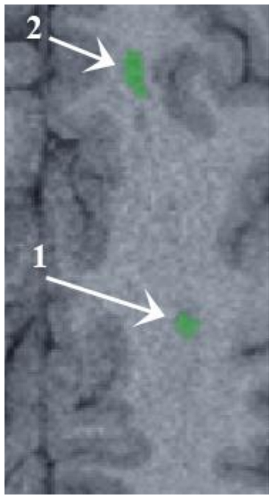
The best accuracy was obtained for the trained model with the input image size 352x352 and the batch-size 9. It is equal to 0.62 (IoU, Jaccard index). combination of the binary cross-entropy and the Sørensen loss function (Dice loss). This model was trained for 100 epochs using a GeForce GTX 1080 Ti graphics card. The training lasted 4 hours.

Examples of the results of segmentation of areas of destruction of the myelin sheath are shown in Fig. 3, 4.

The color of individual areas corresponds to the color in Fig. 2: green corresponds to the selection of the focus by the expert, brown to the selection of the focus by the neural network. The overlapping areas of green and brown correspond to the area selected by both the expert and the neural network.



a) b)



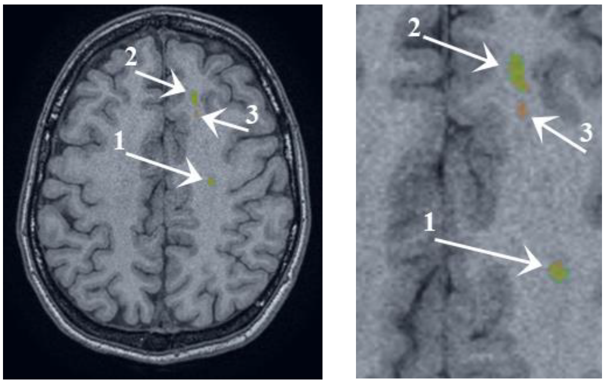
c)

Fig. 3. Myelin sheath lesions: original T1W image (a); areas identified by an expert (b); part of an enlarged image (c)

Fig. 4 shows that, the network found all the affected areas. For two of them, the highlighting coincided with the expert's opinion, one is highlighted contrary to the expert's opinion. In such cases, additional validation is required.

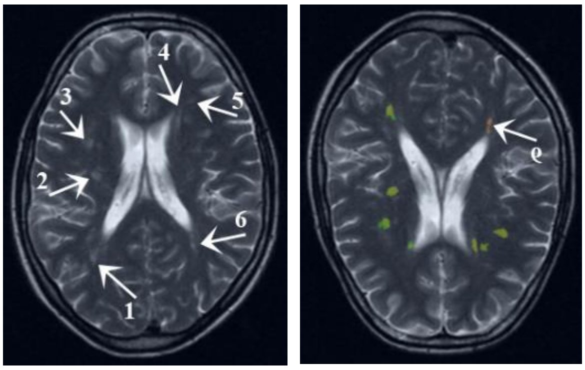
In Fig. 5, the network also highlighted all the foci (the areas highlighted by the expert and the network changed color when superpositioned). In addition, the network highlighted a problem area, which the expert did not mark as a focus of demyelination (it remained brown). In this case, additional validation of the results is required.

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept 2021, Minsk, Belarus – Minsk : UIIP NASB, 2021 – 246 p – ISBN 978-985-7198-07-8



a) b)

Fig. 4. Myelin sheath lesions: areas identified by neural network (a); a enlarged part of the image with the result of comparing the lesions identified by the expert and the network (b)



a) b)

Fig. 5. Myelin sheath lesions: marked on the original T2W image (a); overlay of results of the selection by the expert and the neural network (b). The arrow indicates area 6, identified as an outbreak only by the network

In Fig. 6, the neural network did not identify area 3 as a focus of demyelination, but at the same time found areas 1, 2, 4, which for some reason were not noted by the expert. Additional analysis of the results is required, but neural network extraction looks more promising.

In general, it can be seen that the network successfully detects problem areas on MRI-scans obtained in different modes. The most accurate results are obtained with T2 TSE images. It should be noted that the sizes of the areas selected by the network do not always coincide exactly with the sizes of the same areas identified by the expert. This is due to the ambiguity of the focus boundaries, and the result of network segmentation is not always less accurate, especially when there are many lesions on different sections. Working under severe stress, the expert gets tired and over time his attention weakens.

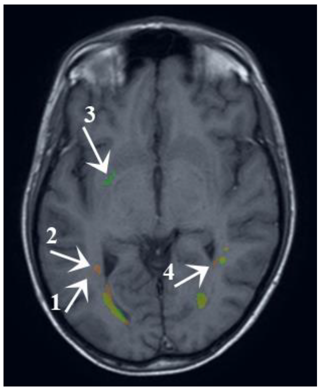


Fig. 6. Myelin sheath lesions: the result of the superposition of the results of the selection of foci by the expert and the neural network. The arrow indicates the areas of non-coincidence

The proposed solution for the automatic identification of pathological areas using artificial neural networks has significantly increased the speed of analyzing the state of the pathological pattern. Traditional manual isolation of demyelination lesions requires at least 65 minutes per patient, semi-automatic isolation takes about 23 minutes, whilea automatic segmentation is performed within 1 minute. In addition, automatic segmentation using a neural network allows to get hundred-percent repeatability of the analysis results and significantly reduce the workload for neurologists and radiologists.

ACKNOWLEDGMENT

The reported study was funded by RFBR and BRFBR, project number 20-57-00025/BRFFI F20R-134.

REFERENCES

[1] C. M. Poser, V. V. Brinar, "Diagnostic criteria for multiple sclerosis: an historical review". Clinical Neurology and Neurosurgery, June 2004, 106 (3), pp. 147–58.

[2] Massimo Filippi and Maria A. Rocca MR Imaging of Multiple Sclerosis. Radiology, 2011, 259:3, pp.659-681.

[3] C. Shorten, T. M. Khoshgoftaar, A survey on Image Data Augmentation for Deep Learning, J Big Data 6, 2019, pp. 60.

[4] J. Nalepa, M. Marcinkiewicz, M. Kawulok, Data Augmentation for Brain-Tumor Segmentation: A Review. Frontiers in Computational Neuroscience, no. 83, 2019, pp. 1662-5188.

[5] Ankur Biswas, Paritosh Bhattacharya, Santi P. Maity & Rita Banik Data Augmentation for Improved Brain Tumor Segmentation, IETE J. of Research, 2021, pp. 0377-2063.

[6] Huang, Huimin, Lanfen Lin, Ruofeng Tong, Hongjie Hu, Qiaowei Zhang, Y. Iwamoto, Xianhua Han, Yen-Wei Chen and Jian Wu, "UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation." ICASSP 2020 – 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020, pp.1055-1059.

[7] J. Duan and X. Liu, "Online Monitoring of Green Pellet Size Distribution in Haze-Degraded Images Based on VGG16-LU-Net and Haze Judgment," in IEEE Transactions on Instrumentation and Measurement, vol. 70, 2021, pp. 1-16.

S. -T. Tran, C. -H. Cheng and D. -G. Liu, "A Multiple Layer U-Net, Un-Net, for Liver and Liver Tumor Segmentation in CT," in IEEE Access, vol. 9, 2021, pp. 3752-3764.

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept 2021, Minsk, Belarus – Minsk : UIIP NASB, 2021 – 246 p – ISBN 978-985-7198-07-8