DOI 10.18699/SBB-2022-466

Artificial intelligence (AI) of 3D MRI images for neurooncology

Letyagin A.Yu.[1, 3]*, Amelina E.V.[1], Tuchinov B.N.[1], Tolstokulakov N.Yu.[1], Amelin M.E.[1, 2], Pavlovsky E.N.[1], Groza V.V.[1, 4], Golushko S.K.[1]

[1] *Novosibirsk State University, Novosibirsk, Russia*

[2] *Federal State Budgetary Institution "Federal Neurosurgical Center", Novosibirsk, Russia*

[3] *Research Institute of Clinical and Experimental Lymphology – Branch of the Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia*

[4] *Median Technologies, Valbonne, France*

* *letyagin-andrey@yandex.ru*

In structural and statistical analysis of neurooncological lesions, there are problems of detection, classification, contouring of tumors, and segmentation of its subregions (Gd-contrast-accumulating, Gd-contrast-non-accumulating, necrotic part of the tumor, perifocal edema), treatment prognosis. Almost all of these tasks are usually solved diagnosis neuro-oncology diseases with use the Artificial Intelligence (AI) applications. AI models training is based on the "Federal Neurosurgical Center" in Novosibirsk, Russia data on patients, who underwent MRI imaging of the brain on a 1.5 T (Siemens Magnetom Avanto) and 3.0 T (Ingenia, Philips) magnetic resonance tomographs: astrocytomas (70; 39 % male; median = 37, IQR (31; 56) years), glioblastomas (140; 53 % male; median = 57, IQR (48; 63) years), meningiomas (160; 26 % male; median = = 58, IQR (50; 64) years), neurinomas (130; 32 % male; median = 53, IQR (40; 59) years). MRI sequences: T1-WI, T1-WI-Gd+, T2-WI, T2-FLAIR and DWI; scans in a single isovoxel anatomical template with a resolution of 1 mm3. The data were processed in the open-source 3D Slicer (www.slicer.org) cross-platform system with manual tumor segmentation and post-correction by two expert neuroradiologists.

Firstly, we elaborated 2D approach for segmentation and classification tasks. As part of the PyTorch machine learning framework, two convolutional neural networks have been developed based on the se-resnext50-32x4d, se-resnext101-32x4d networks [1]. Segmentation of tomograms is implemented according to a multilayer model of a brain tumor (layer = disease) of logically connected 4 blocks (= variant or stage of the disease), consisting of a series of subblocks of the 2D convolution layer, batch normalization, and activation, as well as channel attention blocks, compression and excitation blocks, followed by a Pooling layer, a DropOut layer, and a Linear Fully Connected layer to enhance tumor-specific features. Sections including only perifocal edema are not informative and were excluded from the test set.

The task of tumor type classification was solved on the basis of our dataset. We performed tumor classification using two different CNN developed in the PyTorch framework, that are slightly changed variants of the well-known se- resnext50-32x4d, se-resnext101-32x4d networks. We obtained results using a test set of 40 independent patients with balanced representation of tumor classes. There were only 3 misclassified patients among given 40 in the test set, so it was possible to analyze them individually as well as some other difficult cases that were tricky to resolve properly. We analyzed the observed classification errors from both machine learning and medical perspectives. We also developed recommendations how to improve the results of the algorithm.

Now we expand our SBT dataset and extend our research work to 3D models for account the all context along the axial, sagittal and coronal planes. We have used a well-proven (won on BraTS 20) the nnU-net framework as a 3D baseline. This is a symmetric encoder-decoder architecture using skip-connection. Downsampling is performed with stridden convolutions and upsampling is performed with convolution transposed. Each decoder output is used for dip supervision. A total of five downsampling operations are performed. The initial number of convolution kernels is set to 32, which is doubled with each downsampling up to a maximum of 320. The non-linearity function is Leaky ReLU and normalization is instance normalization

The second 3D approach is DMFnet (Dilated Multi-fiber network). Their main idea is to divide the convolution operation into several paths, called Fiber. Information between the fibers is distributed by means of multiplexer, consisting of 1x1 reapplication of convolution. The difference between DMF and MF units is that the first one applies not only conventional convolutions, but also dilated convolution with d = 1, 2, 3 and different weights. This is done to increase the receptive field and capture spatial correlations at different scales.

The main body of the network is composed of the MF/DMF units, excluding the first and last convolution layers. In the feature encoding stage, we apply the DMF unit in the first six encoding units to achieve multi-scale representation. In the decoding stage, the high-resolution features from the encoder are concatenated with the upsampled features. All upsampling operations are performed using 3-linear interpolation. The nonlinearity function is standard ReLU and normalization in the form of Batch Normalization.

TransBTS is a hybrid architecture of CNN-Transformers type encoder-decoder with skip-connection and bottle-neck transformer sequence. The standard encoder block consists of a 3x3x3 convolution, activation and normalization functions in the form. Downsampling done with strided convolutions. After the encoder feature map contains 128 channels. Then, the 3x3x3 convex representation of each volume is applied again and the number of channels is increased to 512. Since transformers work with 2D patches, the feature map turns into a vector. After that, 8 layers of Vision Transformer (ViT) \cite{vit} are followed to find global dependencies in the obtained features. The ViT output is also a sequence, so it is recorded into a volumetric matrix to obtain an image. The resulting 3-dimensional weighted matrix goes to the input of the decoder. The building block of the decoder consists of 3x3x3 usual and transposed convolutions

Table 1. Results on BraTS dataset. All metrics are average results for 5 folds

Models	Dice ET	Dice TC	Dice WT
LinkNet	0,614	0,758	0,871
nnU-net	0,771	0,836	0,916
DMFnet	0,742	0,812	0,893
TransBTS	0,764	0,817	0,897

Table 2. Results on SBT testing dataset

Models	Dice ET	Dice TC	Dice WT
LinkNet	0.826	0,826	0,911
nnU-net	0,846	0,867	0,917
DMFnet	0,796	0,827	0,900
TransBTS	0,756	0,753	0,858

Based on nnU-net and DMFnet we developed the new techniques – Brain Tumor Segmentation with Self-supervised Enhance Region Post-processing. It is based on the clinical radiology hypothesis and presents an efficient approach to combining and matching 3D methods to search for areas of comprised the GD-enhancing tumor in order to significantly improve the model's performance of the particular applied numerical problem of brain tumor segmentation.

The obtained stability and robustness results of this method was confirmed by the Brats 2021. The proposed method also demonstrates significant improvement on the segmentation problem with respect to Dice and Hausdorff metrics compared to similar training / validation procedures based on any regardless of the chosen neural network architecture. These particular results are very important from the clinical point of view and can strongly increase the quality of the computer-aided systems where similar solutions are deployed.

Acknowledgements: The study was supported by the Russian Foundation for Basic Research project No. 19-29-01103.

References

1. Hu J., Shen L., Sun. G. Squeeze-and-Excitation Networks. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018. doi: 10.1109/CVPR.2018.00745.