

# ANALYSIS OF THE TRAINING QUALITY OF BRAIN TUMOUR SEGMENTATION IN DEEP LEARNING THROUGH SIMILARITY

**Usevalad Ustsinau**

Master Degree Programme (2), FEEC BUT

E-mail: xustsi00@stud.feec.vutbr.cz

Supervised by: Jiri Chmelik

E-mail: chmelikj@feec.vutbr.cz

**Abstract:** Manual segmentation of brain tumours in MR images is a time-consuming process, which increases the required time for the research of tumour development and its lesion on the cognitive functions of human. Recently there were developed modern solutions for this problem by using a fully automatic segmentation algorithm. As far as segmentation quality plays a highly important role for doctors, we have to train such a model with a significant amount of care to quality. In this paper, it is provided with an analysis of the training quality using state-of-art technology - convolutional neural network U-Net and with training on manually segmented data. The experiment has shown the effectiveness of the provided model and performed 50 training cases with the following analysis through the similarity. The results were put on the similarity matrix and dendrogram. The proposed outcome gives us certain ideas for future improving the quality of image segmentation.

**Keywords:** segmentation, convolutional neural networks, deep learning, U-Net

## 1 INTRODUCTION

The rise in deep learning performance through convolutional neural networks, due to their abstractions of different levels of features, motivated researchers to transfer their knowledge acquired by neural networks, when trained on millions of images into new tasks such as medical image segmentation, to benefit from their learned parameters, in particular, weights. By now it has already been invented several successful deep learning methods to solve the segmentation and detection task and to show the importance of medical automatic image recognition. The most common and the most used convolutional neural network for image segmentation is called U-Net, which was developed especially for biomedical image segmentation at the Computer Science Department of the University of Freiburg, Germany.

U-Net at the present moment is the state-of-art technique for image segmentation. However, this fully convolutional network is constantly modifying by international groups of scientists with upgrades for each specific aim. In the case of brain imaging sphere, great achievements can be discovered and shown in the Multimodal Brain Tumor Segmentation Challenge (BRATS) provided by Center for Biomedical Image Computing and Analytics (CBICA) and supported by non-profit corporation MICCAI. Due to the sheer number of such variants, it becomes increasingly difficult for researchers to keep track of which modifications extend their usefulness over the few datasets they are typically demonstrated on. In this article, it was chosen option of No New-Net by MIC-DKFZ team. Unlike other segmentation methods published recently, nnU-Net does not use complicated architectural modifications and instead revolves around the popular U-Net architecture [1]. MIC-DKFZ team have implemented a number of these variants and found that they provide no additional benefit if integrated into a well-trained U-Net. In this context, contribution to the BRATS challenges was intended to demonstrate that such a U-Net, without using significant architectural alterations, is capable of generating competitive state-of-the-art segmentations.

## 1.1 DATASET

Currently, one of the biggest open-source brain image datasets is provided by BRATS competitions. The vast majority of nnU-Net model was developed in the context of the Medical Segmentation Decathlon (<http://medicaldecathlon.com>) [2], where among different tasks the first one is brain tumour with based on the BRATS datasets. The main attributes of provided dataset are collected in Table 1.

**Table 1:** Brain Tumours Dataset Description

|           |                                                        |
|-----------|--------------------------------------------------------|
| Target    | Gliomas segmentation necrotic/active tumour and oedema |
| Modality  | Multimodal multisite MRI data (FLAIR, T1w, T1gd, T2w)  |
| Size      | 750 4D volumes (484 Training + 266 Testing)            |
| Source    | BRATS 2016 and 2017 datasets                           |
| Challenge | Complex and heterogeneously-located targets            |

## 2 METHOD

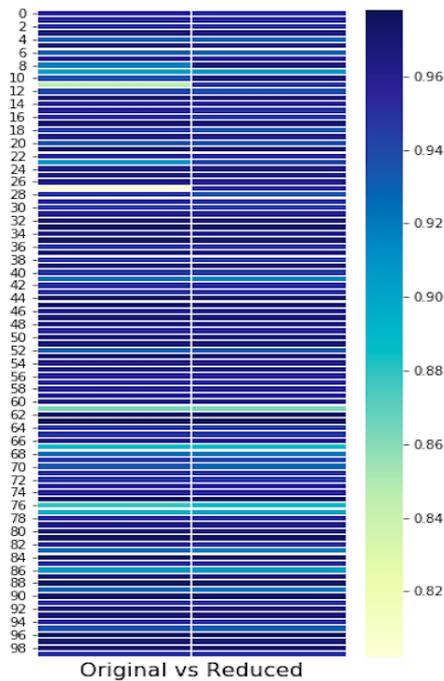
The final and ideal task of this experiment is to verify and improve the quality of image segmentation in convolutional neural networks and optimize the training dataset through its similarity. For analysis of the training quality between several cases, we can use record modalities of one patient for each training case consequently. Then we compare predicted segmentations with hand-marked ground truths by doctors from the original dataset, evaluate them and cluster or classify them. The presented similarity method can be used only, if the ground truth is known, which makes it unusable in real circumstances. However, if the final aim is to achieve a better quality of predictions in image segmentation, it can be implemented via optimized training dataset. Optimizing of the dataset will be organized with similarity classifier that needs ground truths only on the creating stage. To generate the automatic classifier, which can evaluate and optimize image datasets and also will be universal for all images is a generous task, which cannot be solved instantly and requires extensive discussion.

## 3 EXPERIMENT

For training of the models, it was used free GPU source - Google Colab. It offers conditions with 37 GB on disc space and Tesla K80 GPU plus intuitively understandable environment based on Jupyter Notebook. These characteristics do not allow us to provide an experiment on the whole dataset because the size is massive and not enough of available disc space. To solve this problem the dataset was reduced to 100 images, which should be sufficient for training of the networks and some experiments. NnU-Net gives us a variety of set-up for tumour segmentation: 2D, 3D and 3D cascade [3]. For increasing the speed of computation it was chosen the least-dependent capacity model type - 2D.

The experiment includes four steps. In the first step, it was a preparation of training weights. For such purpose, the network has been trained by 30 epochs on 100 images from the provided dataset. The obtained model was saved with its training weights. The saved weights will be used as initial conditions in the next steps. There was made due to the reason for preventing fail and also better visualizing the further training results with the possibility to give them assessment as upgrade or downgrade. Besides, from the received model, we can predict segmentation (Figure 1) and it can give us comparison before and after implementing the classifier.

The second step consisted of preparation and setting up the new pieces of training. Unfortunately, nnU-Net does not allow training the model on only one image. To complete the requirements it was used a data augmentation process - batchgenerators by MIC, DKFZ [4], wherefrom one original



**Figure 1:** Results of segmentation with full dataset and after removing the cluster

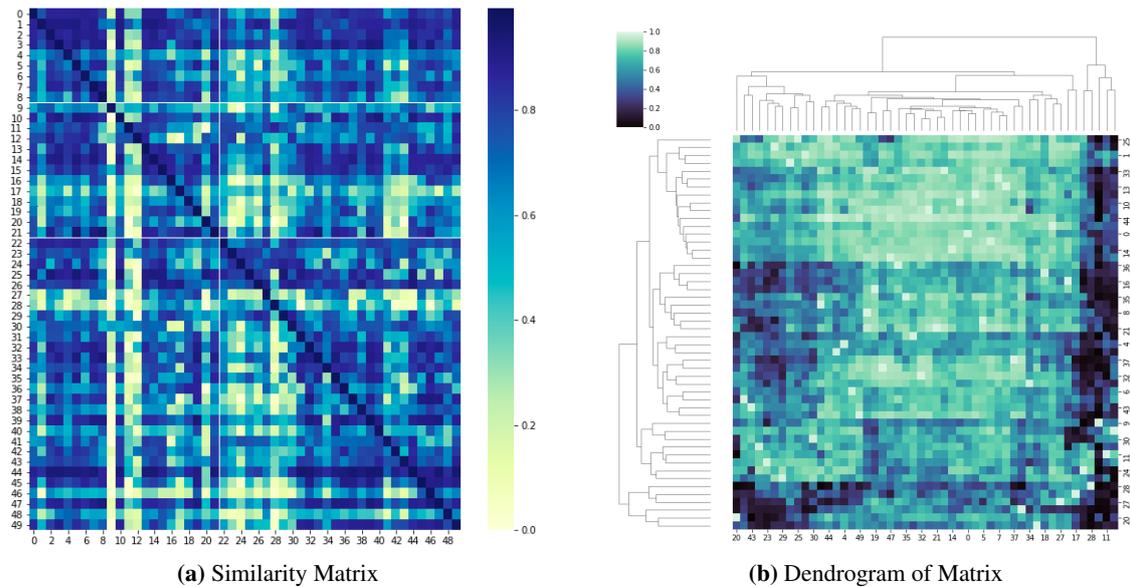
image converted into a set of five with random mirroring by X, Y, Z axes and spatial deformation with the low scale of the deformation parameters.

In the next step, we were training 50 separate models with obtained previously sets during 100 extra epochs. When the training process is done, we can perform inference of brain tumour segmentation for original dataset - 100 images. The calculation of segmentation quality will be provided through the similarity between predicted ground truth and the original one. For comparison of segmentation usually, it is used Dice Coefficient (F1 Score) which calculates as the Area of Overlap multiplied by 2 and divided by the Total Number of Pixels in both images. As a result, we have a value in the range from 0 to 1, where 1 signifies the greatest similarity between predicted and truth. Collecting all the values in a table and visualizing them on the heatmap can get us a similarity matrix (Figure 2a), i.e. one cell of the matrix corresponds to the dice score, that is trained on the corresponding training image, is used to segment the corresponding test image [5]. The resulting matrix indicates how similar each patient training case are to each other.

In the last step, from the heatmap, we built a dendrogram with the average method (also called the UPGMA algorithm) for analysing the results. The following dendrogram (Figure 2b) gives us several clusters of images, where each of them shows us a selection sample for training certain case. To prove this statement, we can take the worst cluster, which has the shadiest colours in a dendrogram and remove these images from the training. There are images 9, 11, 12, 24 and 28. When we will train the model on 95 images and inference the predictions, we can compare them with the predictions, which were made on the beginning of experiment (Figure 1). Compare results we can see how cases 11, 23 and 27 significantly improved their segmentation quality and others remain steadily.

#### 4 CONCLUSION

In this paper, it was conducted a brief analysis of the training quality of brain tumour segmentation with state-of-art convolutional neural network nnU-Net. It shows us effectiveness and reliability in segmentation task, however, the quality can be improved. During the experiment it was trained 50



**Figure 2:** Similarity Matrix and its Dendrogram of 50 Training Cases

convolutional neural networks to assess the quality prediction and based on them it was built similarity matrix (Figure 2a). From the following matrix, we constructed dendrogram (Figure 2b) and displayed clusters of images. Deleting images from the worst cluster from the dataset we can achieve an increase of the segmentation quality more than 10 pct. in several particular cases (Figure 1). The discussion about training setup, an algorithm of clustering and similarity classifier, which unites data in a certain cluster is behind the scope of this paper and will be a future scientific spotlight.

## ACKNOWLEDGEMENT

This work was carried out with scientific support from Dr. Michael Goetz from the German Cancer Research Center (DKFZ), Heidelberg, Germany.

## REFERENCES

- [1] Isensee F., Kickingereder P., Wick W, Bendszus M. and Maier-Hein K. *No New-Net* Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, Springer International Publishing, Cham, p. 234–244, ISBN 978-3-030-11726-9 (2019)
- [2] Isensee F, Petersen J., Klein A., Zimmerer D., Jaeger P. F., Kohl S., Wasserthal J., Kohler G., Norajitra T., Wirkert S., and Maier-Hein K. *nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation*. Division of Medical Image Computing, German Cancer Research Center (DKFZ), Heidelberg, Germany arXiv:1809.10486 (2018)
- [3] Isensee F., et al. *nnU-Net: Breaking the Spell on Successful Medical Image Segmentation*. arXiv:1904.08128 (2019)
- [4] Isensee F., Jaeger P., Wasserthal J., Zimmerer D., Petersen J., Kohl S., Schock J., Klein A., Ross T., Wirkert S., Neher P., Dinkelacker S., Koehler G, Maier-Hein K. *batchgenerators - a python framework for data augmentation*. DOI:10.5281/zenodo.3632567 (2020)
- [5] Goetz, M., Weber C., Thiel C., Maier-Hein K. H. *Input Data Adaptive Learning (IDAL) for Sub-acute Ischemic Stroke Lesion Segmentation* Springer International Publishing Switzerland, A. Crimi et al. (Eds.): BrainLes 2015, p. 284-295. DOI: 10.1007/978-3-319-30858-6 25 (2016)