Golestan University of Medical Sciences

SEMANTIC SEGMENTATION FOR BRAIN VOLUMETRY: A DEEP LEARNING APPROACH USING U-NET ON T1- AND T2-WEIGHTED MAGNETIC RESONANCE IMAGING (MRI) SCANS

I.Attackpour Torik (1) , M.H.GHarib (2) , A.Zare (1) , G.Roshandel (3) , D.Amanian (2)

PURPOSE

In recent years, artificial intelligence has achieved significant metrics in the field of prediction, diagnosis, and prognosis. Unfortunately, these achievements in the field of clinical research have not been able to be used effectively in the clinic for various reasons including limitations in the volume of datasets used for training, the lack of differentiation of non-anatomical structures, and the lack of generalizability of the results. This research aims to produce a reliable and generalizable brain-specified deep learning model for future applications in the clinic.

METHODS & MATERIALS

We extracted the data of 186 patients in the form of DICOM files from the imaging center database, in regard to the principles of research ethics. Then, this data was converted into 80,000 2D images in PNG format with a single color channel using PyDiCOM and OpenCV libraries. Then, this dataset was divided into three groups: training, test, and validation in the proportions of 75-15-15. The U-Net was used as the deep learning architecture. After the first round of training, 300 images were extracted randomly to be annotated manually by 2 radiologists with the purpose of fine-tuning the model. We used DICE coefficient, accuracy, precision, recall, and F1 score as the measurements for our model performance.

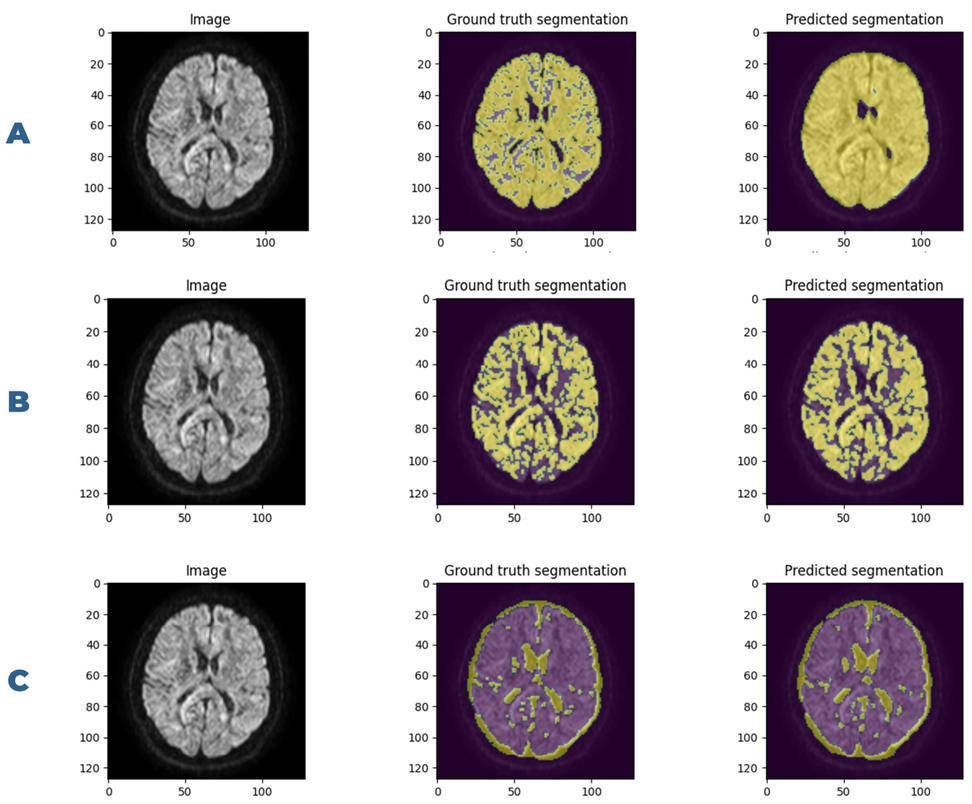


Figure A : Grey Matter Segmentation Post Fine-Tune. Figure B : White Matter Segmentation Post Fine-Tune . Figure C : CSF Segmentation Post Fine-Tune

RESULTS

The Dice score after the fine-tuning process was reported as 0.8645 in gray matter (GM), 0.8452 in white matter (WM), and 0.8593 in cerebrospinal fluid (CSF), all of which showed significant improvement after fine-tuning. The accuracy of the model was reported as 0.9470 for GM, 0.9505 for WM, and 0.9786 for CSF. The precision of the model was reported as 0.7959 for GM, 0.8044 for WM, and 0.8467 for CSF. The recall was reported as 0.9621 for GM, 0.9238 for WM, and 0.8725 for CSF. Finally, the F1 score was reported as 0.8711 for GM, 0.8600 for WM, and 0.8594 for CSF.

	CSF	WM	GM
Dice Score	0.8593	0.8452	0.8645
Pixel Accuracy	0.9786	0.9505	0.9470
Accuracy	0.9786	0.9505	0.9470
Precision	0.8467	0.8044	0.7959
Recall (Sensitivity)	0.8725	0.9238	0.9621
F1 Score	0.8594	0.8600	0.8711
Epochs for Training	281	159	89

Table 1 - Model Scores after Fine-Tune Process in Cerebrospinal Fluid (CSF), White Matter (WM) and Grey Matter (GM)

CONCLUSION

The results of the Fine-Tuning process indicate an improvement in the model measurements in all three cases of GM, WM, and CSF. This change in the model standards indicates the effectiveness of the proposed method in enhancing segmentation accuracy along with its generalizability for its use in clinic. Future work should focus on further validating the model on larger and more diverse datasets, incorporating advanced techniques to handle anatomical variations, and optimizing performance for real-time clinical use.

REFERENCES & AFFILIATIONS

1. Student Research Committee, Golestan University of Medical Sciences, Gorgan, Iran
2. Golestan Research Center of Radiology, Golestan University of Medical Sciences, Gorgan, Iran
3. Golestan Research Center of Gastroenterology and Hepatology, Golestan University of Medical Sciences, Gorgan, Iran