I.R.C.C.S Ospedale San Raffaele
Gruppo San Donato

Glioma Segmentation in PET/MRI studies: a preliminary comparative study between Swin Transformer and state-off-the-art CNN models

PSMR2024
10th Conference
20-23 May, Isola D'Elba

Nicolo Pecco[1,2], Pasquale Anthony Della Rosa [1], Ilario Tagliaferri[6], Gianluca Nocera [1,2,4], Michele Bailo[4], Matteo Canini[1], Paola Scifo[3], Antonella Castellano[1,2], Andrea Falini[1,2]

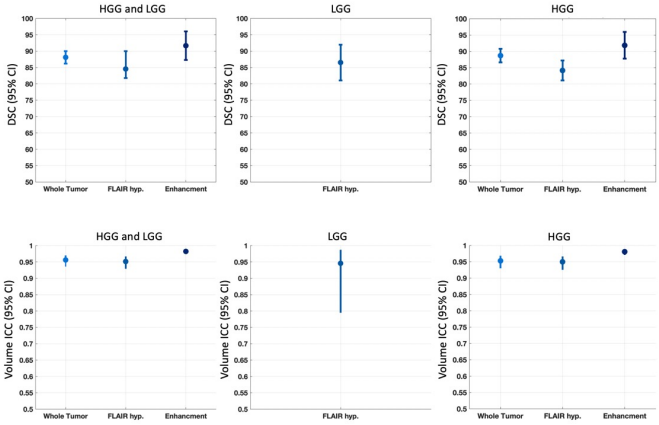
[1] Neuroradiology Unit and CERMAC, IRCCS Ospedale San Raffaele, Milan, Italy, [2] Vita-Salute San Raffaele University, Milan, Italy, [3] Department of Nuclear Medicine, IRCCS Ospedale San Raffaele, Milan, Italy , [4] Department of Neurosurgery, IRCCS Ospedale San Raffaele, Milan, Italy, [6] IRCCS San Raffaele Research Institute

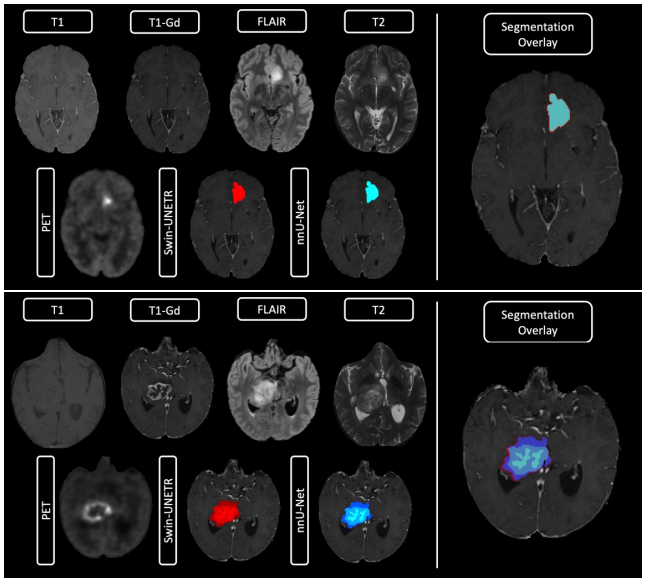
INTRODUCTION

Fully Hybrid PET/MRI is a recently introduced imaging technique enabling the simultaneous acquisition of PET and MR image in one single session. The potential of 11C-Methionine and 18F-FET PET/MRI in neuro-oncology relies on the multimodal investigation of brain tumors [1] but also in the optimization of radiomics workflows, as the MRI-based glioma segmentation step to delineate PET-based target volumes for treatment planning. Automated glioma segmentation plays a crucial role in radiomics and machine learning model development. While Convolutional Neural Networks (CNNs) have demonstrated success in MRI segmentation [2,3], attention-based transformers are emerging as a promising alternative. This study aims to compare the Swin Transformer-based model [4] and the current state-of-the-art CNN model nnUNET [5] on glioma segmentation task.

MATERIALS AND METHODS

Swin-UNETR was trained and validated on the BraTS dataset (n=1251), enabling the segmentation of contrast-enhancing tissue, FLAIR hyperintensity, and necrosis. In contrast, nnUNET was trained and validated on the Heidelberg University Hospital dataset (n=455) and focused on only segmenting FLAIR hyperintensity and contrast-enhancing tissue. A retrospective cohort of **105 patients** from the San Raffaele Hospital with histologically confirmed gliomas, encompassing WHO grade low-grade (n=26) and high-grade (n=79) gliomas was used as test set for the study. The dataset included T1-weighted pre- and post-contrast sequences, FLAIR, and T2-weighted. The T1 post-contrast sequence of each patient was resampled to 1mm isotropic voxel and used as the reference volume for rigid coregistration after skull stripping with the "HD-bet" algorithm . Dice Similarity Coefficient (DSC) and the volume interclass correlation coefficient (ICC) with 95% Confidence Interval (CI) were used for assessing segmentation accuracy.





RESULTS

The results demonstrate a high level of agreement between the two algorithms with a DSC of 88.09% (95% CI: 86.18% - 89.99%) for the entire tumor, 84.57% (95% CI: 81.75% - 87.38%) for FLAIR hyperintensity, and 91.66% (95% CI: 87.28% - 96.04%) for the enhancing tissue. Excellent reliability was observed regarding the ICC based on the number of voxels for segmentation compartments. The ICC values were 0.956 (95% CI: 0.935 – 0.970) for the entire tumor, 0.951% (95% CI: 0.929 – 0.966) for FLAIR hyperintensity, and 0.982 (95% CI: 0.973 - 0.987) for enhancing tissue, respectively. Results for high-grade gliomas report similar results with 88.69% (95% CI: 86.62% - 90.76%) for the entire tumor, 84.15% (95% CI: 81.11% - 87.19%) for FLAIR hyperintensity, and 91.86% (95% CI: 87.76% - 95.95%) for enhancing tissue, with excellent ICC score. Differently, low-grade gliomas report results for FLAIR hyperintensity with a DSC of 86.52% (95% CI: 81.06% - 91.98%) and an ICC of 0.945% (95% CI: 0.79 – 0.987).

DISCUSSION and CONCLUSION

Both models exhibited a high DSC for high-grade glioma in segmenting each glioma compartment, while showing an increase in variability in the case of FLAIR hyperintensity for low-grade glioma. One of the limitations of this study is the absence of manually segmented gliomas, which could serve as a benchmark for evaluating the performance of the Swin UNETR and nnU-NET. The next step will be to extend the workflow including PET images to optimize functional volume quantitation. These findings not only have the potential to refine the accuracy of glioma diagnosis and treatment planning but also serve as a stepping stone for future research in the domain of radiomics analyses.

REFERENCES

[1] Lohmann P, Werner JM, et al. Combined Amino Acid Positron Emission Tomography and Advanced Magnetic Resonance Imaging in Glioma Patients. Cancers (Basel). 2019 Jan 29;11(2):153; PMCID: PMC6406895. [2] Akkus Z, Galimzianova A, Hoogi A, Rubin DL, Erickson BJ. Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions. J Digit Imaging. 2017;30:449–459. [3] Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. Z Für Med Phys. 2019; 29:102–127. [4] Hatamizadeh A, Nath V, Tang Y, Yang D, Roth H, Xu D. Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images. arXiv; 2022. [5] Kickingereder, Philipp, et al. "Automated quantitative tumour response assessment of MRI in neuro-oncology with artificial neural networks: a multicentre, retrospective study." The Lancet Oncology 20.5 (2019): 728-740.

Acknowledgement: This research was funded by the Italian Ministry of Health, grant number NET-2018- 12366666 NeuroArtP3. The study was approved by the Ethics Committee of Ospedale San Raffaele on November 10th, 2021 (code 357/2021).