ONE COMMUNITY
ISMRM & SMRT
Virtual Conference & Exhibition
08-14 August 2020

The openGTN project is funded by the European Union in the Marie Curie ITN-EID program (project 764465)

Addressing the need for less MRI sequence dependent DL-based segmentation methods: model generalization to multi-site and multi-scanner data

Yasmina Al Khalil[1], Cristian Lorenz[2], Juergen Weese[2], Marcel Breeuwer[1,3]

 # 3560 [1]Biomedical Engineering Department, Eindhoven University of Technology, Eindhoven, The Netherlands. [2]Philips Research Laboratories, Hamburg, Germany. [3]Philips Healthcare, MR R&D - Clinical Science, Best, The Netherlands

INTRODUCTION

The versatility of MRI acquisition parameters and sequences can have a substantial impact on the performance of medical image segmentation algorithms. Even though recent studies report excellent results of deep-learning (DL) based algorithms, their generalization capability is rarely addressed[1,2]. This study attempts to demonstrate the lack of adaptation of such algorithms to unseen data from different sites and scanners.

METHODS

Data

- **BRaTS 2015[3] -** Skull-stripped and co-registered **multi-modal MRI (T1-2, post-contrast T1-w, T2-w and FLAIR per patient)** – Fig. 1
- **Heterogeneous** - different institutions (**4 sites, labeled as CS, DU, HT and FG**) acquired with **different protocols** and **scanners**
- Labels: edema, necrosis, non-enhancing and enhancing tumor

Approach

- State-of-the-art **3D U-Net**[4] (Fig.2) to assess its **adaptability to unseen data** acquired from a site not included in the training
- Split the data per site and train 3 models, with each site excluded per model

CONCLUSION

- The performance of **DL-based segmentation** models significantly depends on the **type of data available** for training
- Combining **multi-site and multi-vendor data** is a potential solution to improve **generalization**; however, it is unrealistic
- There is an evident need for **better** algorithm **evaluation** the development of better generalization methods to unseen data

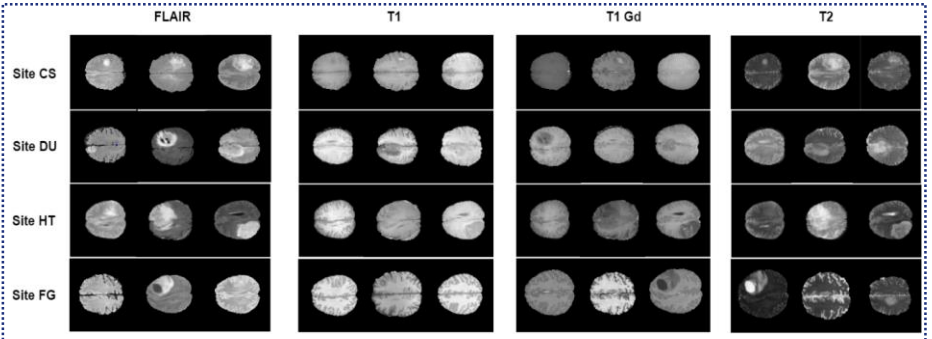

Figure 1: Representative depiction of contrast differences due to variation in scan parameters and vendors between the datasets from 4 different sites per each MRI modality. These are acquired for the same slice across all volumes.

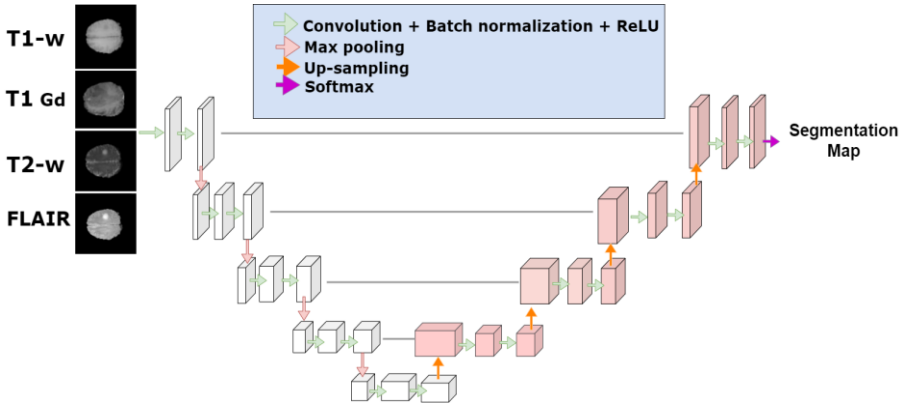

Figure 2: A 3D U-Net training strategy. All data is normalized using z-score normalization per patient. Data augmentation is applied on the fly, using elastic deformations, random scaling and rotation. The sum of cross-entropy and dice is used as a loss function, optimized utilizing Adam with an initial learning rate of $2*10^{-4}$ and an l2 decay of 10^{-5}.

[1] Bai W et al. *JCMR* 2018. [2] Chen C et al. *Front. Cardiovasc. Med.* 2020. [3] Bakas S et al. *MICCAI Brain lesion* 2017 [4] Isensee F et al. arXiv 2019.

RESULTS

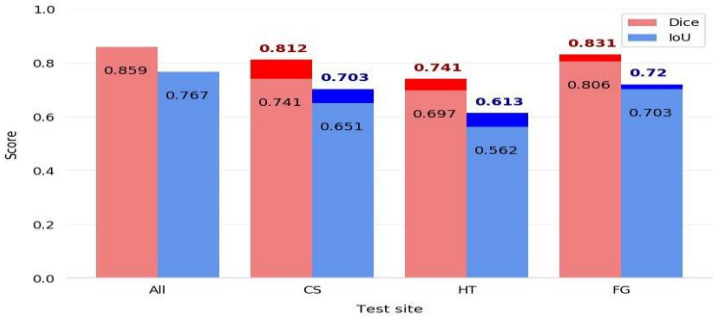

Figure 3: Segmentation performance of the 3D U-Net across different sites, with each site excluded from the training (columns 2-4, light red and blue). Column 1 represents the baseline performance of the network trained on all sites. Addition of a small subset of the testing set to the training (3 volumes from CS, 3 volumes from HT and 2 from FG) improves the segmentation performance (dark red and blue). All scores are the mean Dice and intersection over union (IoU) per site.

Train site	Test site	Train set size	Test set size	Whole tumor Dice	Whole tumor IoU
DU	CS	35	11	0.584	0.406
DU	HT	35	13	0.522	0.392
DU	FG	35	6	0.791	0.641
DU + CS	CS	38	8	0.682	0.517
DU + HT	HT	38	10	0.639	0.483
DU + FG	FG	37	4	0.827	0.662

Table 1: Segmentation performance of the network trained on site DU and tested across other sites (rows 1 to 3). Rows 4 to 6 demonstrate that with addition of a small subset of the test data, the segmentation performance and generalization capability of the network significantly increases.