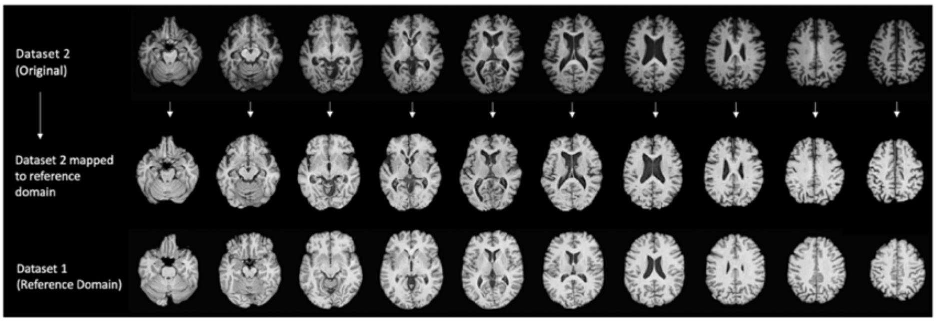
Federated harmonization of MRI volumes

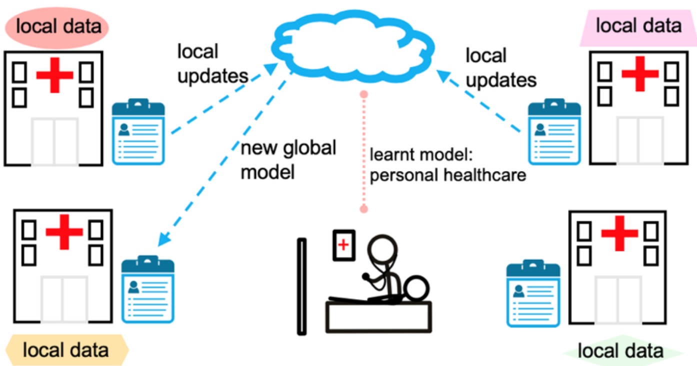
Background

MRI (magnetic resonance imaging) is used to study many diseases, such as brain tumors and breast cancer. While MRI provides excellent contrast for soft tissue, a general problem for developing deep learning tools is that images from different MR scanners can look very different, due to different scanner manufacturers (Philips, Siemens, …), different field strengths (1.5 T, 3 T, …) and different MR sequences. Harmonization of the MR volumes can improve performance of deep learning tools, when trained on images from one MR scanner and evaluated on images from another MR scanner. Bashyam et al. (2022) showed that brain age prediction improved from a mean error of 9.78 years without harmonization, to 7.74 years with histogram normalization and to 5.32 years with GAN-based harmonization.



An example of harmonization of MR volumes from different MR scanners using deep learning. The harmonization is much more complex than simply changing the intensity of each volume, as different MR sequences are used.

A general problem in medical imaging is that it is difficult to share data between hospitals or researchers. Federated learning can be a solution to using larger datasets; a computer (node) at each hospital joins a federation that together trains a network to for example perform segmentation. No images are sent between the nodes during the training, only updates of the segmentation network. We have previously had three master theses on federated learning for medical images, and now want to investigate if harmonization of MR volumes can be performed without having all data in one computer.



The main idea in federated learning is to not store all data in a single computer, but to instead store for example image data locally at each hospital. Instead of sending medical images and other medical data between the hospitals, the hospitals send updates, or parameters, of deep learning models. This process is then iterated to convergence.

Objectives

Investigate how different the MR volumes from different hospitals are, for example using metrics from radiomics.

Implement and evaluate centralized harmonization of the MR volumes, i.e. having all volumes in one computer, to see how it for example affects segmentation performance (for example using a U-Net) compared to using standard image augmentation during training. The harmonization can for example be done using GANs (generative adversarial networks) or diffusion models. The harmonization can be started in 2D, and be extended to 3D if time permits.

Implement and evaluate federated harmonization of the volumes, to see how it affects segmentation performance. How good is federated harmonization compared to centralized harmonization?

Investigate how big the difference is between centralized and federated harmonization when using different solutions for privacy, such as differential privacy.

Data

Several large open multi-site MRI datasets can be used, such as BraTS 2021 (1251 brain tumor patients) and MAMA-MIA (1506 breast cancer patients) which contain data from different MR scanners.

Required background

Machine learning, deep learning, Python programming

Computing resources

The student will have access to very good computing resources (graphics cards) for federated learning. It is also possible to use the supercomputer Berzelius (752 graphics cards) for simulations.

References

Bashyam, V. M., … Singh, A., (2022). Deep Generative Medical Image Harmonization for Improving Cross-Site Generalization in Deep Learning Predictors. Journal of Magnetic Resonance Imaging

Contact persons
Anders Eklund, anders.eklund@liu.se, Department of Biomedical Engineering, Department of Computer and Information Science, Center for Medical Image Science and Visualization