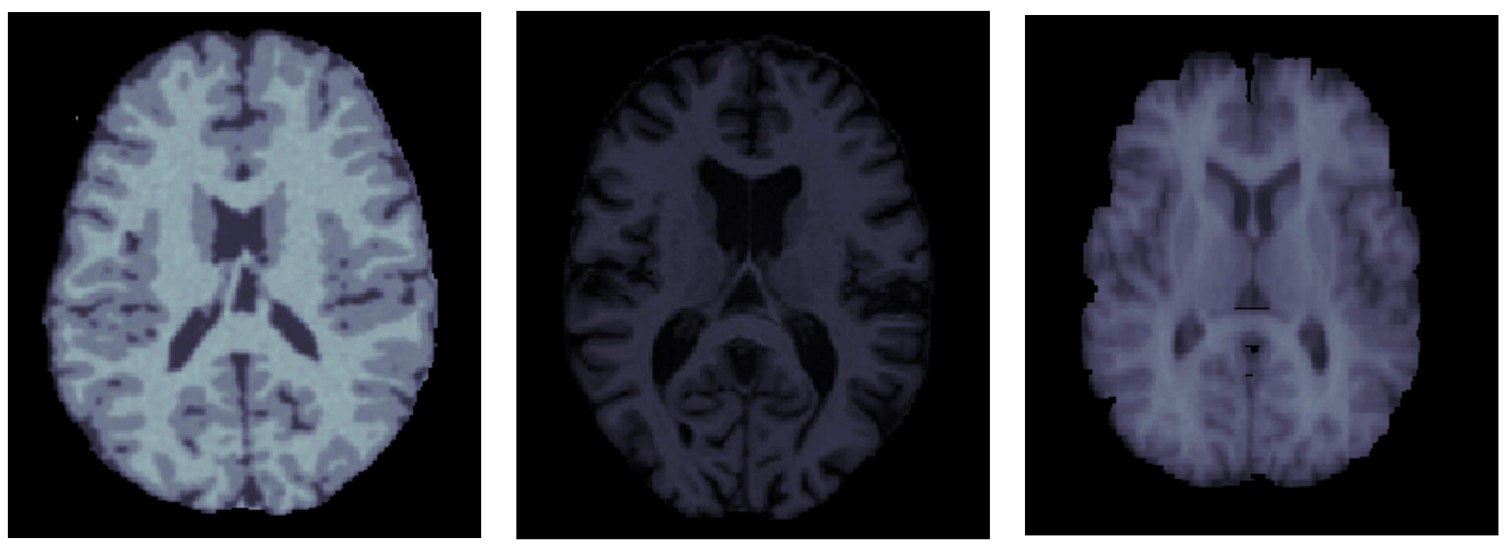
MR acquisition-invariant representation learning

WM Kouw M Loog LW Bartels AM Mendrik

Generalization of voxelwise classifiers is hampered by variation in acquisition protocols across medical centers. To address this limitation, we propose a Siamese neural network (MRAI-net) that extracts acquisition-invariant feature vectors. These can be used by task-specific methods, such as voxelwise classifiers for tissue segmentation.

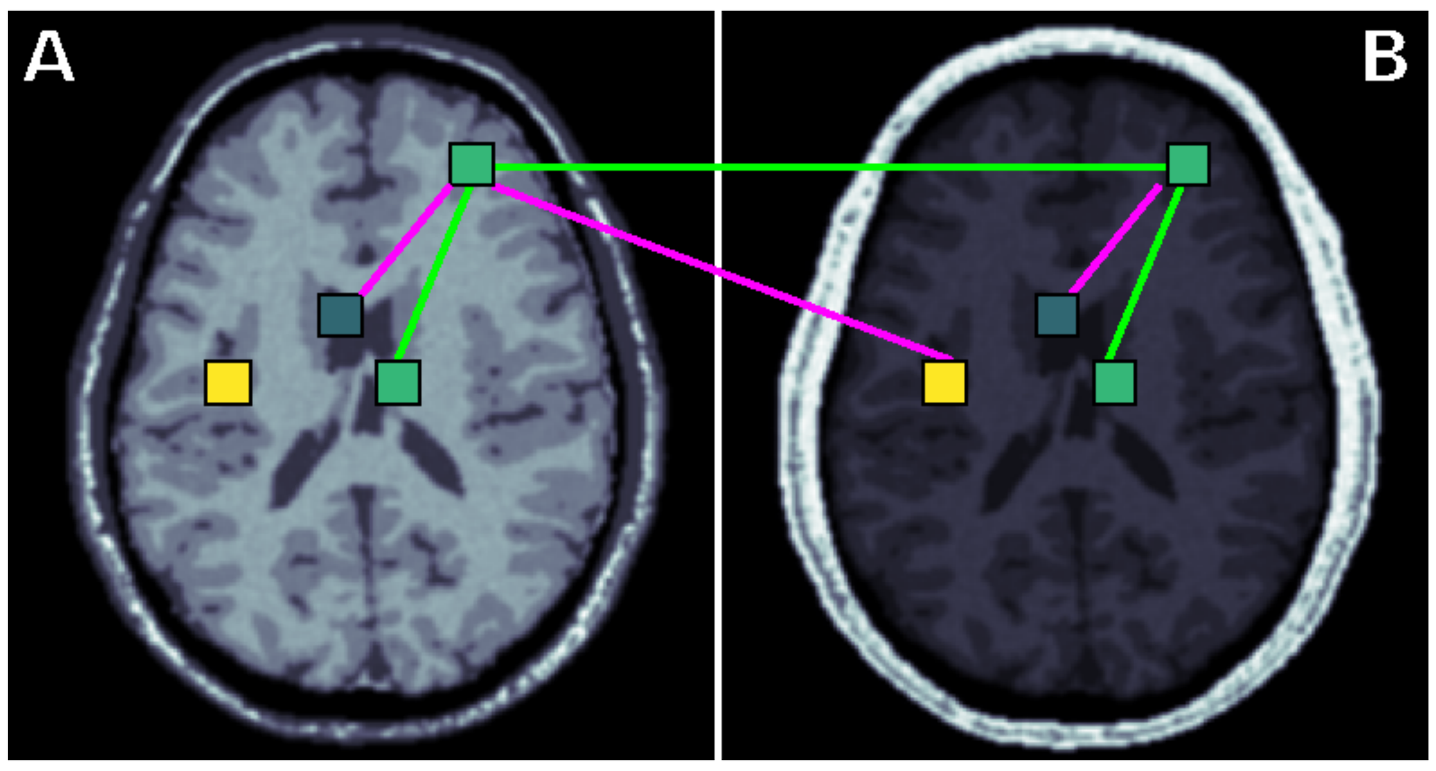
MR acquisition-variation

MR scans vary across medical centers due to different acquisition protocols.



Similarity-based pairing of patches

We collect pairs of patches from two different scanners, based on tissues.

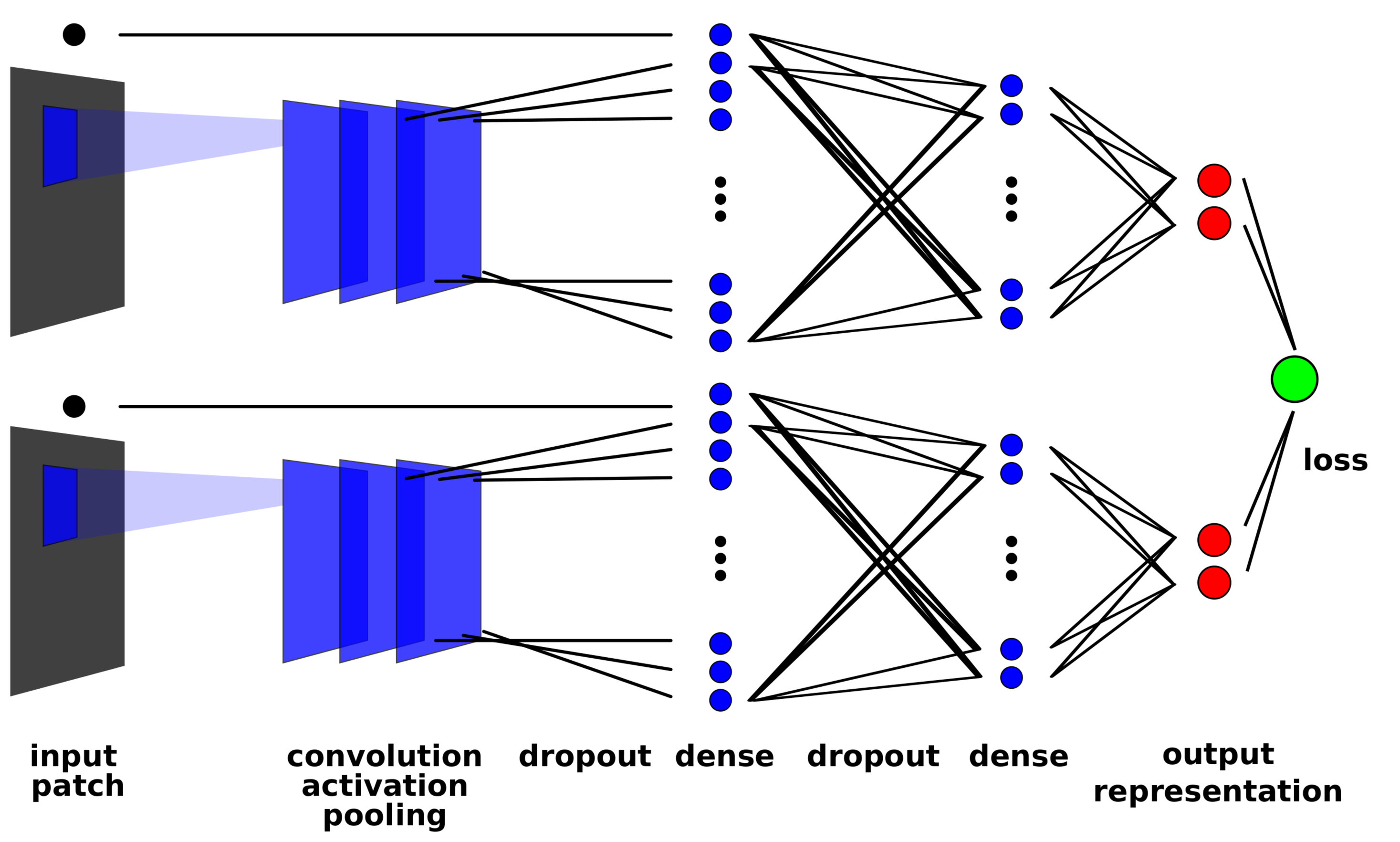


Siamese neural network

Pairwise similarity loss function:

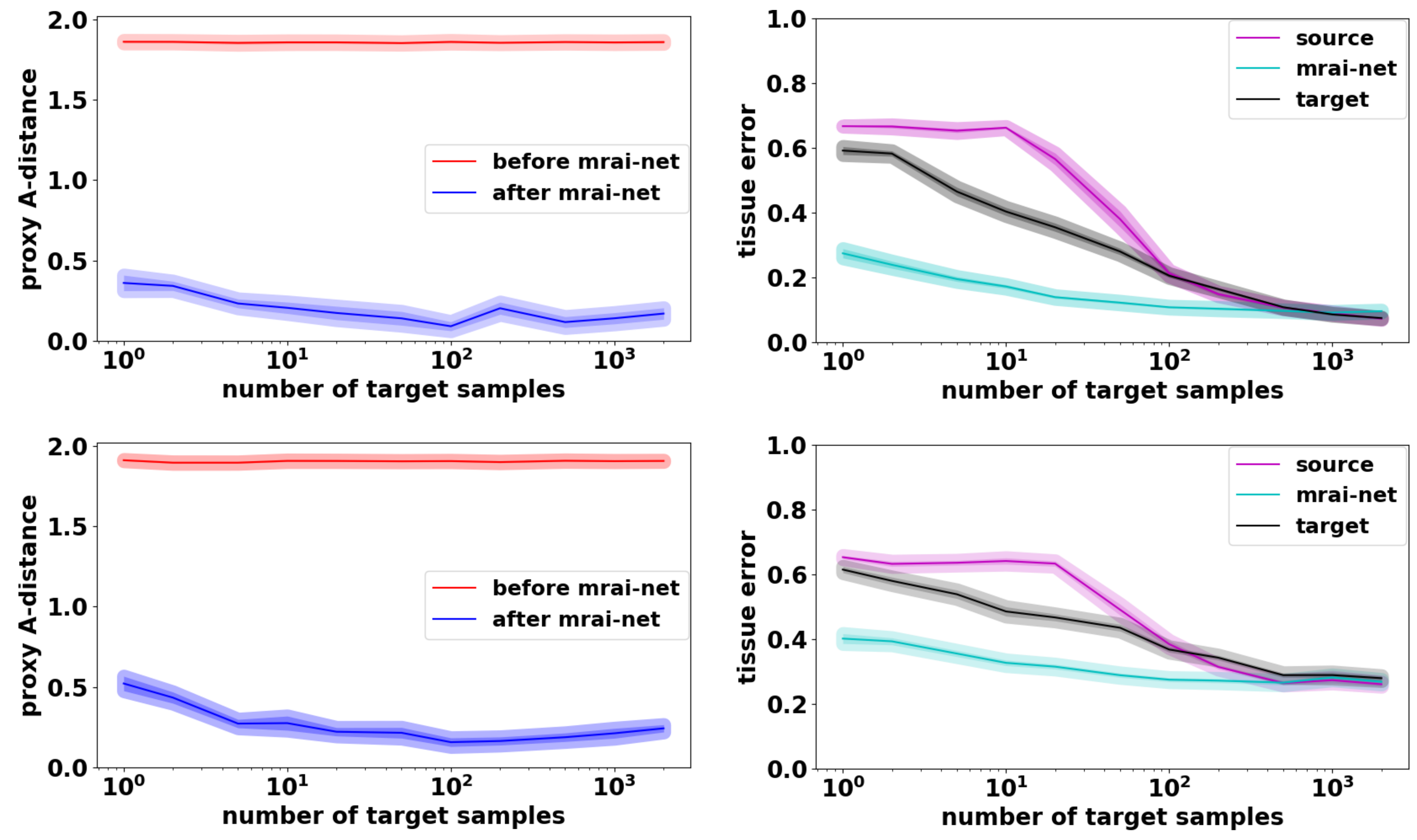
$$\ell(f) = \sum_i y_i \; d_f(s_i, t_i)^2 + (1 - y_i) \max\left[0, m - d_f(s_i, t_i)\right] \, .$$

We train the Siamese neural network to pull pairs labeled as *similar* closer together, while pairs labeled as *dissimilar* are pushed apart.



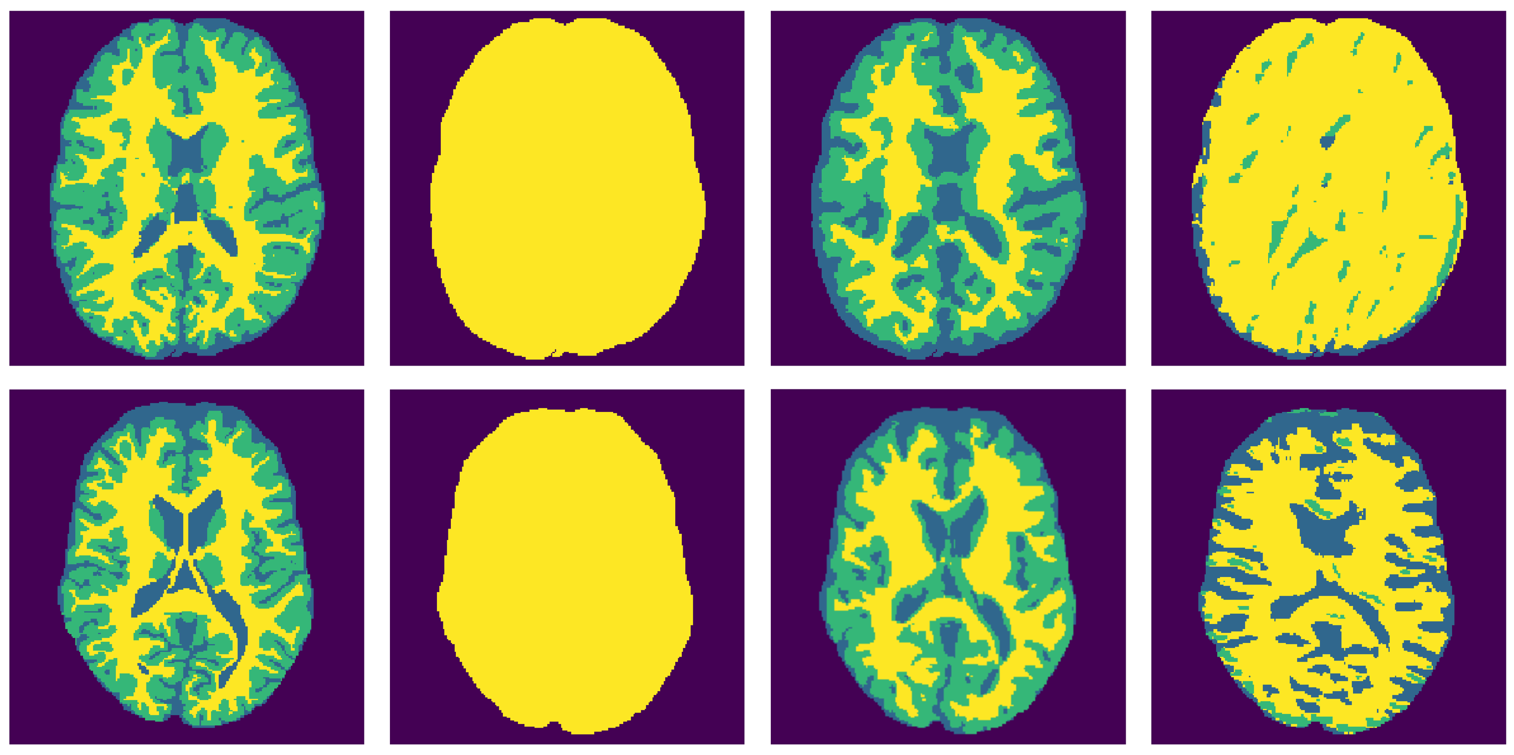
Experiments

We perform two experiments to show the effect of the learned representation. First, we measure the "closeness" between pairs of patches from two different scanners in terms of proxy A-distance. The network quickly matches the two data sets (see left figures below). Secondly, we trained a linear classifier on source patches mapped to the learned representation and classified the target patches (see right figures below). This approach outperforms both a CNN trained on source + target patches and one trained only on target patches. These experiments are performed for using Brainweb1.5T as the source with Brainweb3.0T as the target (top), and using Braintweb1.5T as source with MRBrainS13 data set.



Example segmentations

Below are shown example segmentations with 1 labeled target patch per tissue, produced for Brainweb1.5T as source with Brainweb3.0T as target (top row) and Brainweb1.5T as source with MRBrainS13 as target (bottom row). The leftmost figures show the ground-truth segmentations. The second-column shows the segmentations produced by a CNN trained on both source patches and the 1 target patch per tissue. The third column shows the performance of a classifier trained on source patches mapped to MRAI-net's representation and the last column shows the performance of a CNN trained on the 1 target patch per class. Note that the performance of source is worse than that of target (interference) and MRAI-net already produces good results with only 1 target patch per tissue.



The proposed representation learning method could be used to reduce any type of variation, through adjusting the definition of similar and dissimilar pairs. Key is to identify the factors of variation that should be preserved and the factors that should be reduced.

TU Delft UMC Utrecht DATALOGISK INSTITUT KØBENHAVNS UNIVERSITET