CONSTRAINED HARMONIZATION ALGORITHM FOR POOLING MULTI-SITE DATASETS

Background

Pooling datasets from multiple studies can significantly improve statistical power: larger sample sizes can enable the identification of otherwise weak disease-specific patterns. When modern learning methods are utilized (e.g., for predicting progression to dementia), differences in data acquisition-methods / scanner-protocols can enable the model to “cheat”, i.e. utilizes site-specific artifacts rather than disease-specific features. In this study, we develop a method to harmonize the performance of DNN classifiers across scanners/sites, via so-called fairness constraints, thereby encouraging consistent behavior while controlling for site-specific nuisance variables.

Methods

We conducted two studies: (a) to demonstrate feasibility of pooling across sites (Site-Pooling) and (b) to pool data across scanners (Scanner-Pooling). For Site-Pooling, our analysis included summaries from Freesurfer processed T1-weighted images of the Wisconsin Alzheimer's Disease Research Center (ADRC) and German Center for Neurodegenerative Diseases (DZNE). The Freesurfer summaries were used to train a two layer neural network classifier and five-fold cross-validation performance was assessed. For Scanner-Pooling experiments, Freesurfer processed MR images from the Alzheimer's Disease Neuroimaging Initiative (ADNI) were used to train a deep 3D convolutional network. Performance average on a held-out test dataset was evaluated. In both cases, a constraint to equalize the performance of the trained classifier across the domains (sites/scanners) was incorporated during training.

Results

Table 1 shows the results of AD/MCI classification for site-pooling analysis. Our proposed method is compared against a naive pooling approach which does not incorporate the “harmonization constraint”. As shown, the proposed method improves the “difference of errors” measure by 8% / 7% and with only a small drop in overall error rates. Figure 1 illustrates the results from our scanner-pooling analysis. The performance across the three scanners, GE, Siemens and Philips, is evaluated pair-wise. A consistent improvement in harmonization is observed and only ~2% drop in overall error rate is seen.

Conclusions

We provide a harmonization constraint based algorithm to mitigate site specific differences when performing analysis of pooled brain imaging datasets in AD studies. In contrast to a method which modifies the data, we achieve harmonization by constraining the classifier to perform similarly across sites/groups/scanners, improving reproducibility.

Learning Objective

Discuss novel datasets pooling methods for the analysis of AD imaging datasets using harmonization constraint techniques.

	TASK - AD vs. CN			TASK - MCI vs. CN		
	#AD Samples ADRC / DZNE	Overall Error rate	Difference in Error rates	#MCI Samples ADRC / DZNE	Overall Error rate	Difference in Error rates
Naive Pooling	30 / 60	13%	15%	56 / 92	7%	11%
Constrained Harmonization		9%	7% (**8% gain**)		9%	4% (**7% gain**)

Table 1. Site-Pooling study. We contrast the naive pooling method against the constraint harmonization method for two tasks, namely, AD vs CN classification and MCI vs. CN classification. A naive pooling method holds a high difference of errors across the groups. Such a difference is minimized by the proposed constrained harmonization method.

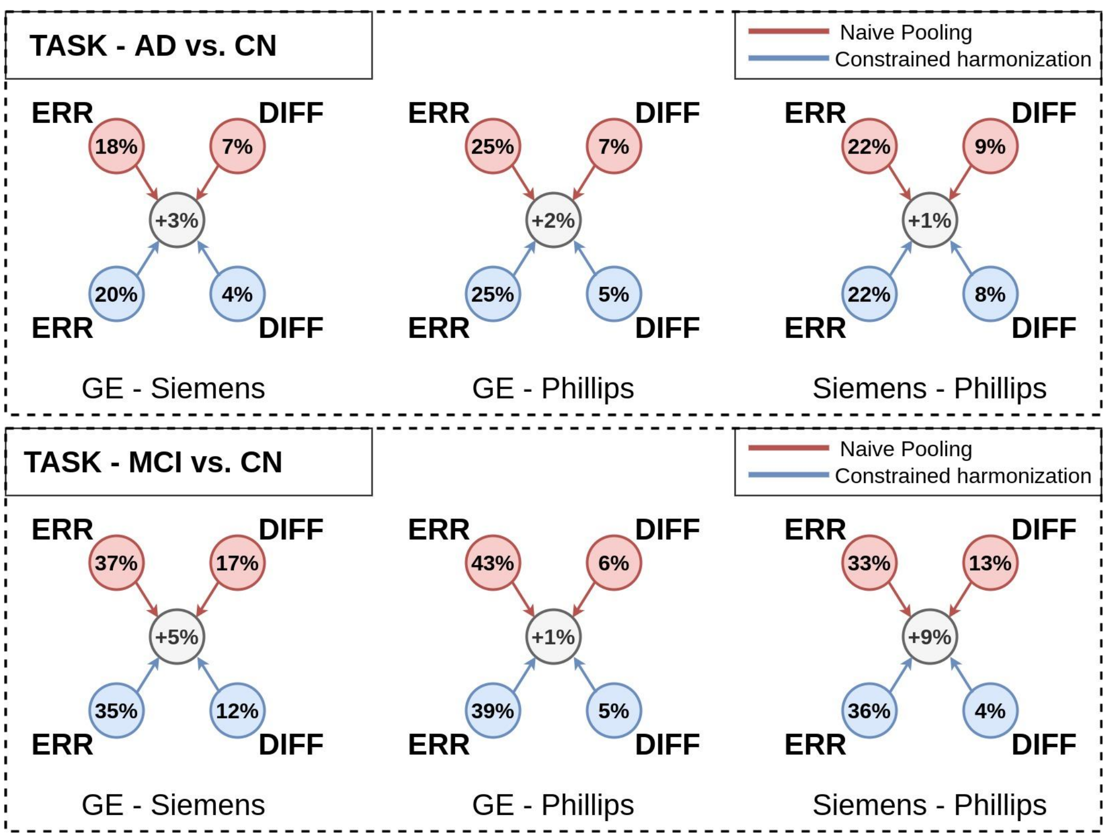

Figure 1. Scanner-Pooling Study. ERR denotes the overall error rate and DIFF denotes the difference of error rates across the groups. The performance of a naive pooling approach is indicated in red while blue shows the proposed constrained harmonization approach. The mid circle indicates the improvement in DIFF measure obtained from the proposed method.