HEINRICH HEINE UNIVERSITÄT DÜSSELDORF
SYSTEMS NEUROSCIENCE

Poster nr. 1657

Consistent sex classification accuracies across independent datasets

Lisa Wiersch[1,2], Kaustubh Patil[1,2] , Simon Eickhoff[1,2] & Susanne Weis[1,2]

[1]Institute of Systems Neuroscience, Heinrich Heine University Düsseldorf, Düsseldorf, Germany
[2]Institute of Neuroscience and Medicine (INM-7: Brain and Behaviour), Research Centre Jülich, Jülich, Germany
l.wiersch@fz-juelich.de

JÜLICH Forschungszentrum

Brain & Behaviour

Introduction

- Ongoing debate: Does a sexual dimorphism in the brain exist (1) or is the overlap of brain features is greater than the difference between the sexes (2)
- Inconsistent findings (3,4) for sex differences in brain organization as captured by resting-state functional connectivity (RSFC)
- The often used group comparison approach is insufficient to encompass the full complexity of sex diffferences in the brain (5)
- Machine learning (ML) approaches should be favored instead
- A ML-classifier that is able to accurately classify male from female brains can be taken as indicator that the expression of these brain features are more sex-specific than overlapping
- We aim to extend the work by Weis et al. (2020, 6) to five independent datasets

→ **How accurately can a subjects´ sex be classified according to the RSFC?**

→ **Are highly classifying regions consistent across datasets?**

Methods

Samples:

- Cambridge Centre for Ageing and Neuroscience sample (CamCAN): N = 622, age range: 20-87, mean age: 54.78)
- Enhanced Nathan Kline Institute-Rockland sample (eNKI): N = 458, age range: 20-85, mean age: 43.71
- 1000BRAINS study: N = 1042, age range: 20-85, mean age: 59.08
- Brain Genomics Superstruct Project (GSP): N = 870, age range: 21-35, mean age: 23.01)
- Human Connectome Project (HCP): N = 966, age range: 22-37, mean age: 28.29
- All samples were matched for age and sex within each sample

RS Connectome:

- Parcelwise approach with 436 parcels from the Schaefer Atlas (7) and the Brainnetome Atlas (8)
- Time course of activation in RS in each parcel summarized by the eigenvariate
- FC for each parcel computed as correlation of this parcel´s time course with each FC of the remaining 435 parcel

Sex classification:

- Support Vector Machine with radial basis function kernel (SVM-RBF, 9) model for classification of each subject´s sex from the RS connectome
- Nested optimization for cost and gamma hyperparameters
- 10 repetitions of 10-fold cross-validation (CV)
- Classification accuracy was averaged over all folds and repetitions of the outer CV-procedure

Results

CV classification accuracies:

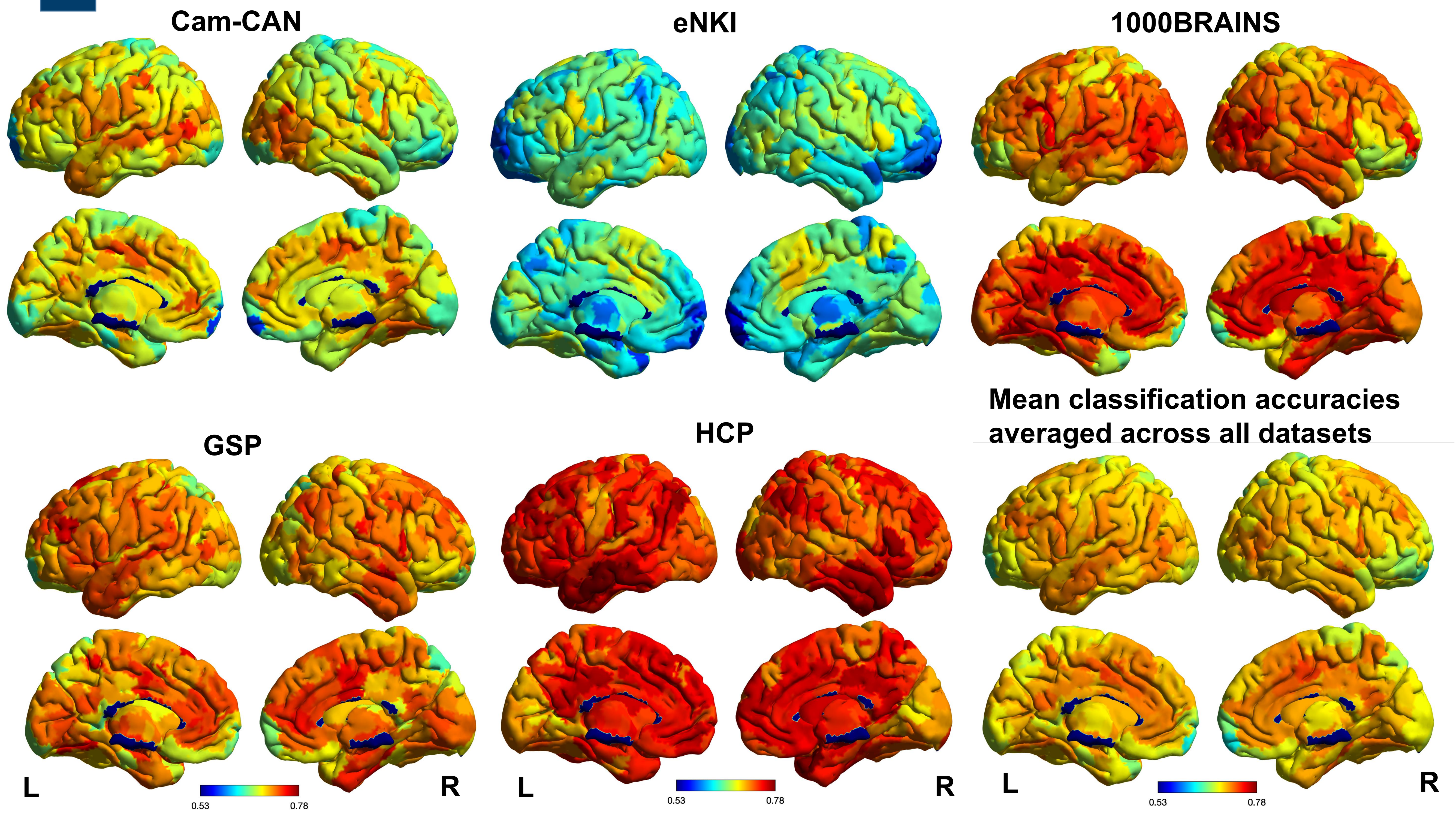
- HCP: M = 73.59%, SD = 1.86%
- 1000BRAINS: M = 72.51%, SD = 2.30%)
- GSP: M = 71.22%, SD = 2.25%
- Cam-CAN: M = 68.72%, SD = 2.50%
- eNKI: M = 64.34%, SD = 2.66%
- Highly predictive parcels were mainly located in the temporal lobe, cingulate cortex, inferior frontal gyrus and the insula

1 Spearman rank correlations displaying the order of classifying parcels across datasets

	CamCAN	eNKI	1000BRAINS	GSP	HCP
CamCAN	1	0.4001**	0.5710**	0.4028**	0.0909
eNKI	0.4001**	1	0.3138**	0.3639**	-0.0122
1000BRAINS	0.5710**	0.3138**	1	0.4575**	0.1092*
GSP	0.4028**	0.3639**	0.4575**	1	0.1860**
HCP	0.0909	-0.0122	0.1092*	0.1860**	1

* $p < 0.05$
** $p < 0.0001$

2 Lateral and medial view of the spatial distribution of parcel-based classification accuracies



Discussion

- Classification accuracies varied between datasets, which might be attributable to the variability in sample size, age range and imaging parameters
- Spearman rank correlations showed parcels are in a similar order regarding classification accuracies across datasets for the 1000BRAINS study and GSP dataset
- This pattern was not found for the correlations of HCP with CamCAN and eNKI
- This might be attributed to the sample size since the N for eNKI and CamCAN is not as high as for HCP, 1000BBRAINS and GSP
- Concerning the variance in classification accuracies, HCP has the lowest which might also lead to low correlations coefficients
- Lower variance in HCP may rely on the good quality of the HCP dataset, resulting in less overall variance in the results due to noise
- All datasets show a consistent pattern of brain regions displaying high classification accuracies
- Highly classifying regions are located in the temporal lobe, cingulate cortex, inferior frontal gyrus and the insula
- Weis et al. (2020, 6) found similar regions to be highly classifying
- Within-sample classification accuracies for all five datasets are also in a similar range as in the study by Weis et al. (2020,6)
- Highly classifying regions are related to the default mode network, high-level cognition and the subjective representation of the body (10)
- Classification accuracies were moderately high, indicating the features in RSFC are not fully sexual dimorphic
- The features can be rather seen as parts of the human brain mosaic which features may be common in males and females
- Still, there are similar brain regions for all datasets that distinguish between males and females on a high level

→ **Highly classifying brain regions are consistent across datasets, independent of sample size, age range or imaging parameters!**

References

1.Glezerman, M., *Yes, there is a female and a male brain: Morphology versus functionality.* Proc Natl Acad Sci U S A, 2016. **113**(14): p. E1971.
2.Joel, D., et al., *Sex beyond the genitalia: The human brain mosaic.* Proceedings of the National Academy of Sciences, 2015. **112**(50): p. 15468–15473.
3.Scheinost, D., et al., *Sex differences in normal age trajectories of functional brain networks.* Human brain mapping, 2015. **36**(4): p. 1524–1535.
4.Zhang, C., et al., *Sex and Age Effects of Functional Connectivity in Early Adulthood.* Brain Connect, 2016. **6**(9): p. 700-713.
5.Forde, N.J., et al., *Sex Differences in Variability of Brain Structure Across the Lifespan.* Cereb Cortex, 2020. **30**(10): p. 5420-5430.
6.Weis, S., et al., *Sex Classification by Resting State Brain Connectivity.* Cereb Cortex, 2020. **30**(2): p. 824-835.
7.Schaefer, A., et al., *Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI.* Cereb Cortex, 2018. **28**(9): p. 3095-3114.
8.Fan, L., et al., *The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture.* Cereb Cortex, 2016. **26**(8): p. 3508-26.
9.Chang, C.-C. and C.-J. Lin, *Libsvm.* ACM Transactions on Intelligent Systems and Technology, 2011. **2**(3): p. 1-27.
10.Lin, C.S., et al., *Neural network of body representation differs between transsexuals and cissexuals.* PLoS One, 2014. **9**(1): p. e85914.