

Control for brain volume in biological sex classification JÜLICI of cisgender and transgender individuals





Susanne Weis^{1,2}, Lisa Wiersch^{1,2}, Sami Hamdan^{1,2}, Felix Hoffstaedter^{1,2}, Mikhail Votinov^{3,4} Ute Habel 3,4, Benjamin Clemens 3,4, Birgit Derntl 5,6, Simon B. Eickhoff1,2, Kaustubh R. Patil1,2

WTh051

*Institute of Systems Neuroscience, Heinrich Heine University Düsseldorf, Germany; *Institute of Neuroscience and Medicine (INM-7), Research Centre Jülich, Germany; *Department of Psychiatry, Psychotherapy and Psychosomatics, Faculty of Medicine, RWTH Aachen University, Aachen, Germany; *Institute of Neuroscience and Medicine (INM-10), Research Centre Jülich, Jülich, Germany; *Department of Psychiatry and Psychotherapy, University of Tübingen, Tübingen, Germany; *LEAD Graduate School and Research Network, University of Tübingen, Tübingen, Germany; s.weis@fz-juelich.de

Introduction

- Total intracranial volume (TIV) differs prominently between males and females
- This difference can cause bias in machine learning (ML) based investigation of sex differences in brain structure
- High sex classification accuracies might be driven by TIV, rather than qualitative structural sex differences
- Such TIV bias might be especially relevant for populations for which modulations of brain size have been suggested ([1])
- Reduced sex classification accuracies in transgender individuals have been taken to indicate an interactive effect of biological sex and gender identity on brain organization ([1,2])
- Here, TIV bias in sex classification models applied to cisgender and transgender individuals was investigated by controlling for brain size either through feature-wise confound removal or by matching training samples for TIV, i.e. debiasing the model on the data level.

Methods

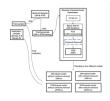
- mutually exclusive cisgender training samples (N = 1292, 646 f):
 - AM training set: Matched for age
 - ATM training set: Matched for age and TIV
- CAT 12.5 preprocessing, whole brain voxel-wise GMV extraction, PCA: n=1292 components [3]
- Support Vector Machine (SVM) classifiers, rbf kernel, Bayesian Hyper-Parameter Optimisation, nested 5-fold CV inner loop, stratified 10-fold CV outer loop) trained to classify sex [4]
- Models tested on:

Test sample: AM

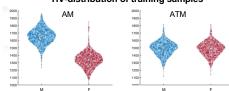
- held out AM and ATM test samples (n=322, 161f) and
- Application sample (n = 202) [5,6]:
- 43 cisgender men (CM)
- 47 cisgender women (CW)
- 62 transgender men (TM)
- 50 transgender women (TW)

Four sex classification models:

- 1) AM training sample without feature-wise TIV confound removal (AM);
- 2) AM training sample with feature-wise TIV confound removal (AM+cr);
- ATM training sample without feature-wise TIV confound removal (ATM)
- 4) ATM training sample with feature-wise TIV confound removal (ATM+cr),



TIV-distribution of training samples



- The AM model achieved highest sex classification accuracies for both test samples as well as cisgender and transgender individuals in the application sample
- TIV distributions revealed strong However. evidence for a TIV bias with the model tending to classify subjects with a higher TIV as male and those with a lower TIV as female
- Feature-wise control for TIV in the AM+cr model resulted in decreased classification accuracies for both test samples and the application sample, but the model output did not reveal any evidence for a TIV bias, in that classifications were not dependent on TIV
- Both the ATM and the ATM+cr models resulted in roughly similar classification accuracies, which were somewhat lower than those of the AM model. Here, additional feature-wise TIV control did not result in a pronounced decrease in model performance.
- Both the ATM and the ATM+cr model did not show any evidence of being biased by TIV
- The ATM model performed better on the ATM
- Across all models, there was no qualitative differences in classification performance of cisgender an transgender individuals for both the biased and the non-biased models

Results

BA = 0.9689

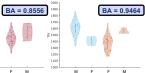
Test sample: ATM

Application: Cisgender

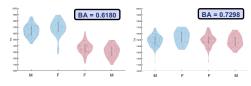
Application: Transgender

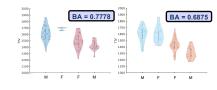
Age matched training sample: TIV biased model BA = 0.7919



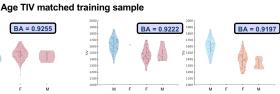


Age matched training sample with confound removal

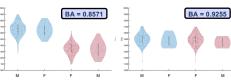


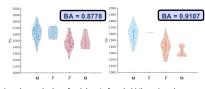


BA = 0.9255









TIV distributions by predicted label (sex). Shades of blue denote male birth assigned sex, shades of red denote female birth assigned sex. cisgender women, cisgender men, transgender women, transgender men

Discussion

- The TIV-biased AM model achieved very high sex classification accuracy, but inspection of the TIV-distributions shows for a strong TIV bias of the model
- The TIV-biased AM model performed worse on the ATM as compared to the AM test sample, since in the ATM sample the model could rely less on sex differences in TIV
- Feature-wise TIV confound removal in the AM+cr model reduced the bias of the model at the cost of reduced classification accuracies
- While a lack of bias in a model is desirable, so is
- high accuracy, suggesting that feature-wise confound removal might not be the ideal approach to reduce bias in structural sex classification
- Both the ATM and the ATM+cr model achieved high classification accuracies for both test samples and the application sample, while not showing any indication of a TIV bias
- We propose that matching men and women for TIV in the training sample provides an appropriate approach for creating an unbiased model that achieves high classification accuracies for both TIV matched and nonmatched test samples
- All models performed at similar levels of accuracy for cisgender and transgender individuals in the application sample
- Present results suggest that non-TIV-biased models are able to classify the sex of both cisgender and transgender individuals with high accuracy, highlighting the relevance of appropriate TIV control in sex classification studies
- Furthermore, present results contradict previous studies reporting reduced sex classification accuracies and higher uncertainty for transgender individuals

1. Flint, C., Forster, K., Koser, S. A., Konrad, C., Zwitserlood, P., Berger, K., . . . Grotegerd, D. (2020). Biologicalsex classification with structural MRI data shows increased misclassification in transgender women. Neuropsychopharmacology. doi:10.1038/s41386-020-0666.3. 2. Baldinger-Meilör, P., Urquijo Castro, M. F., Seiger, R., Ruef, A., Dwyer, D. B., Kranz, G. S., Koutsouleris, N. (2020). Sex Matters: A. Multivariate Pattern Analysis of Sex- and Gender-Related Neuroanatomical Differences in Cis- and Transgender Individuals Using Structura Magnetic Resonance Imaging. Cereb Cortex,30(3), 1345-1356. doi:10.1093/cercor/bhz170. 3. http://www.neuro.uni-jena.de/cat12/

https://juaml.github.io/julearn/main/index.html
 Clemens B, Derntl B, Smith E, Junger J, Neulen J, Mingoia G, Schneider F, Abel T, Bzdok D, Habel U.Predictive Pattern C Can Distinguish Gender Identity Subtypes from Behavior and Brain Imaging Cereb Cortex. 2020 May 14;30(5):2755-2765. do 10.1093/cercor/bhz272. PMID: 31999324.

10.1093/cercorinn2z/cz_MMID: 31999342.
6. Uribe, C., Junque, C., Gomez-Gil, E., Abos, A., Mueller, S. C., & Guillamon, A. (2020). Data for functionalMRI connectivity in transgender people with gender incongruence and cisgender individuals. Data Brief, 31,105691. doi:10.1016/j.dib.2020.105691