Anatomically-Informed Data Augmentation for Functional MRI with Applications to Deep Learning

Kevin P. Nguyen, Cherise Chin Fatt, Alex Treacher, Cooper Mellema, Madhukar H. Trivedi, and Albert Montillo

Department of Bioinformatics

UT Southwestern Medical Center, Dallas, Texas

SPIE Medical Imaging February 2020

Deep learning in data-limited situations

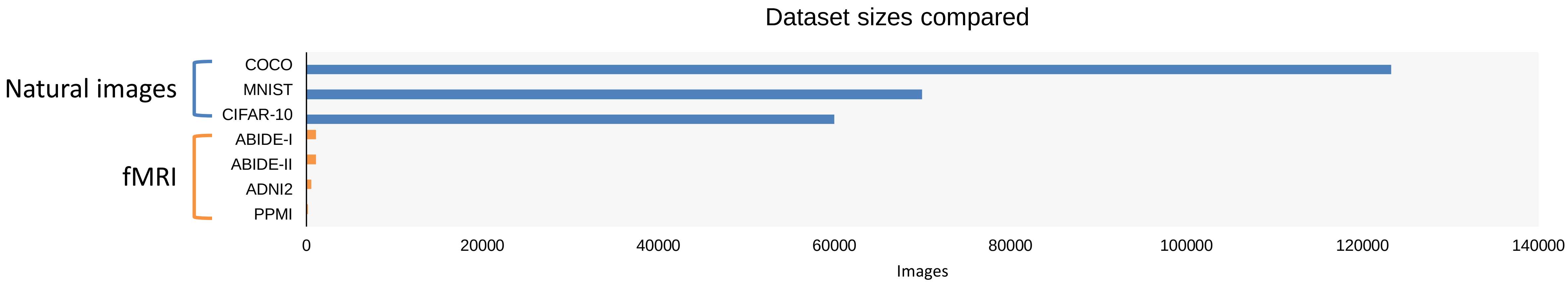
Neuroimaging (esp. fMRI) dataset sizes are frequently limited compared to commonly-used DL datasets.

Natural image datasets

Dataset	Images
ImageNet	14 million
SVHN	630,420
COCO	123,287
CIFAR-10	60,000

fMRI datasets

Study	Unique subjects with fMRI
ABIDE-I	1112
ABIDE-II	1081
ADNI2	551
PPMI	185



Data augmentation in deep learning

Data augmentation in natural image problems involves:

Geometric transformations

- Scaling
  - $\boldsymbol{I}' = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{I}$
- Rotation
  - $\boldsymbol{I}' = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{I}$
- Translation
  - $\boldsymbol{I}' = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{I}$
- Shearing
  - $\boldsymbol{I}' = \begin{bmatrix} 1 & c_x & 0 \\ c_y & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{I}$
- etc.

Image (pixel intensity) transformations

- Saturation
- Contrast
- Gamma
- Blurring
- Cutout

Original

Scaling

Rotation

Shearing

Desaturation

Cutout

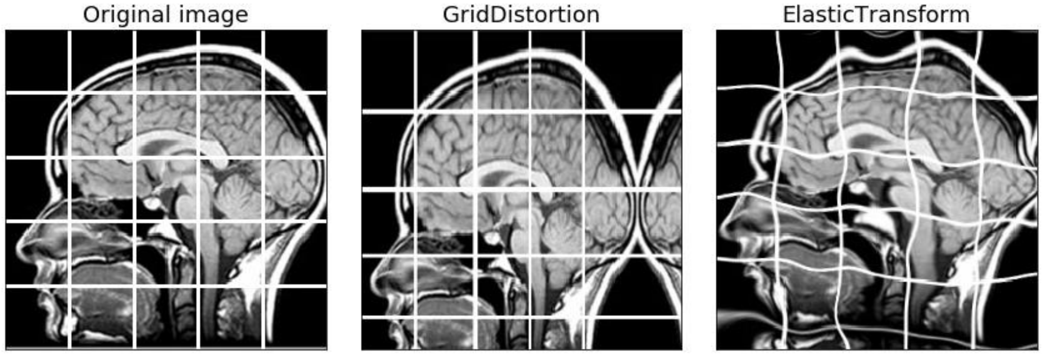
Recent methods for automated natural image augmentation achieve large performance benefits:

- AutoAugment (Cubuk et al. CVPR 2019): up to **29%** relative improvement on CIFAR-10
- Population-based augmentation (Ho et al. ICML 2019): up to **37%** relative improvement on CIFAR-10

Cubuk et al. AutoAugment: Learning Augmentation Policies from Data. CVPR 2019.
Ho et al. Population Based Augmentation: Efficient Learning of Augmentation Policy Schedules. ICML 2019.

Data augmentation for fMRI

Augmentation techniques for natural images do not create anatomically realistic images.



Github.com/albumentations-team/albumentations

Previous methods for structural MRI augmentation:

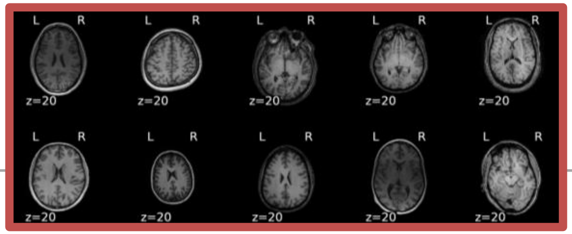
- Ulloa et al., 2015; Castro et al., 2015: ICA and random loading matrices, up to **8%** relative improvement in Schizophrenia diagnosis
- Relatively less work on fMRI augmentation

Goals:

- Should be constrained to **neuroanatomically realistic** brain morphology and appearance
- Requires minimal user parameterization
- Should readily scale to large augmentation targets
- Tangible benefit for deep learning models

Ulloa et al. Synthetic structural magnetic resonance image generator improves deep learning prediction of schizophrenia, MLSP 2015.
Castro et al. Generation of synthetic structural magnetic resonance images for deep learning pre-training, ISBI 2015.

Proposed method



Select target

Raw **source** fMRI ($\boldsymbol{I}_f$)

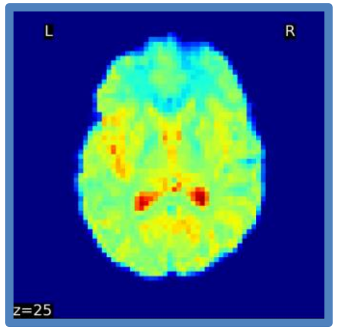
Brain extraction (FSL BET & AFNI 3dAutomask)

Raw **source** sMRI ($\boldsymbol{I}_S$)

Brain extraction (ROBEX)

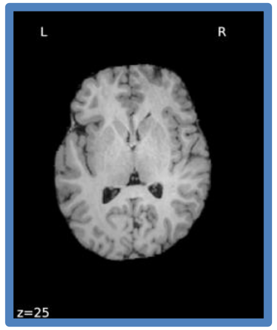
Raw **target** sMRI ($\boldsymbol{I}_S$)

Brain extraction (ROBEX)



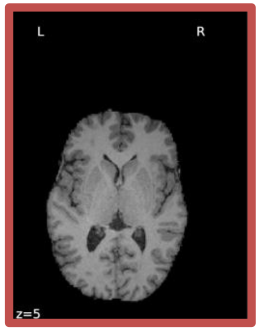
Nonlinear registration*

$T_{fs}(\boldsymbol{I}_f)$



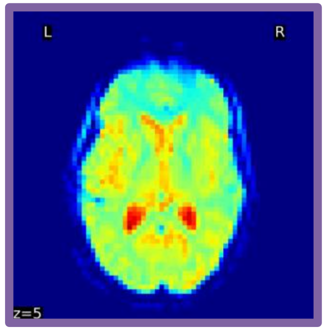
Nonlinear registration**

$T_{ss}(\boldsymbol{I}_s)$



Final transformation

$T_{ss}\left(T_{fs}(\boldsymbol{I}_f)\right)$

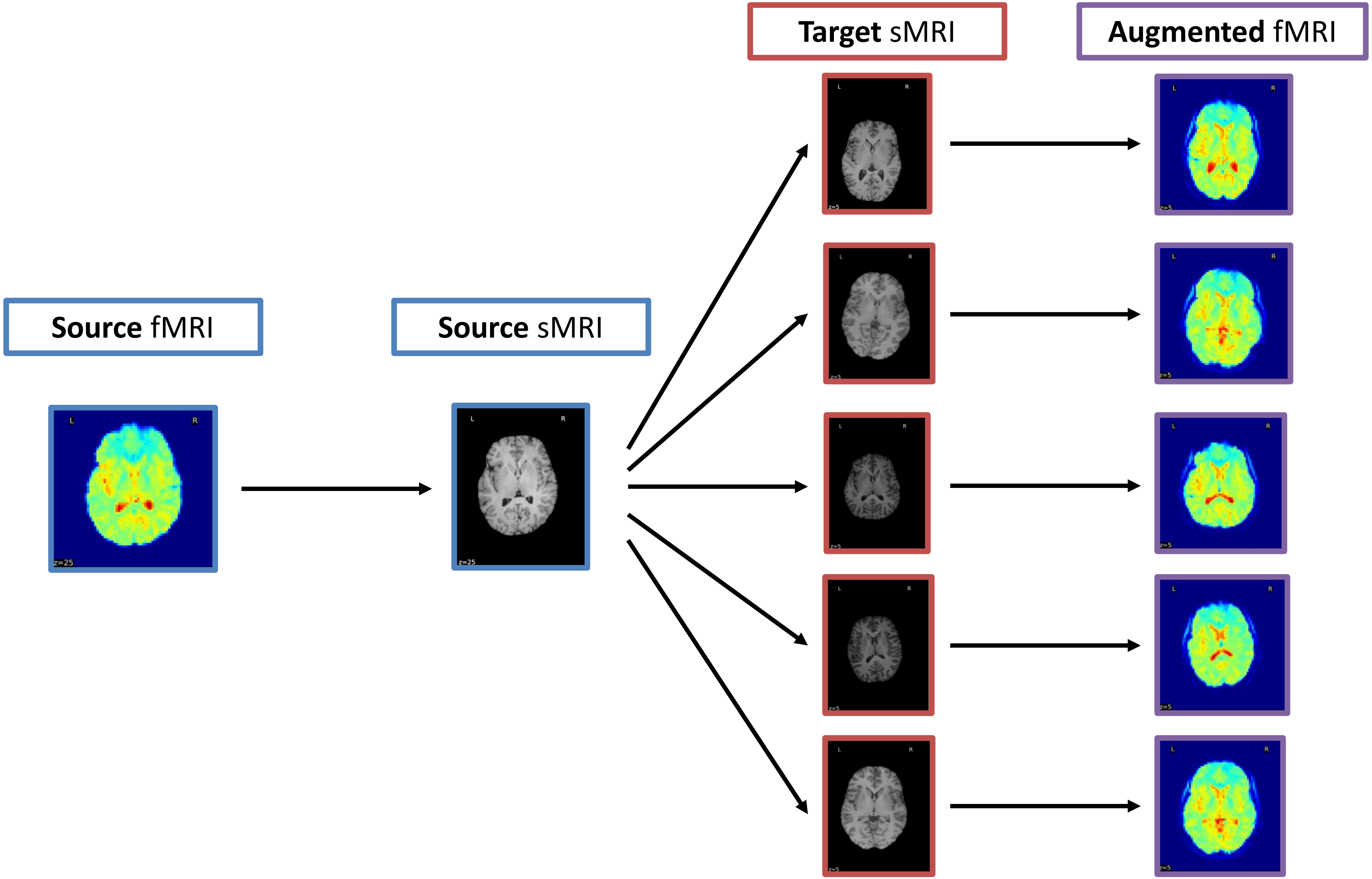


$$\{\boldsymbol{v}_1^*, \boldsymbol{v}_2^*\} = \underset{v_{1,2}}{\operatorname{argmin}} \left\{ \int_0^{0.5} \|L\boldsymbol{v}_1(x,t)\|^2 \, dt + \int_0^{0.5} \|L\boldsymbol{v}_2(x,t)\|^2 \, dt + \lambda \int_\Omega \Pi(\mathcal{J} \circ \phi_1(\boldsymbol{x}, 0.5), \mathcal{J} \circ \phi_2(\boldsymbol{x}, 0.5)) \, d\Omega \right\}$$

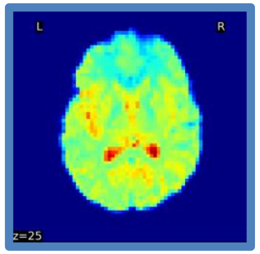
*ANTS RegistrationSynQuick

**ANTS RegistrationSyn

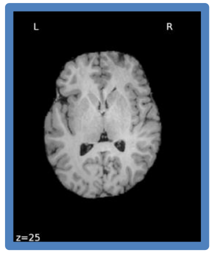
Avants et al. A Reproducible Evaluation of ANTs Similarity Metric Performance in Brain Image Registration NeuroImage 2010.


Source fMRI
Source sMRI
Target sMRI
Augmented fMRI

Source fMRI



Source sMRI



ABIDE
Autism Brain Imaging
Data Exchange

PARKINSON'S
PROGRESSION
MARKERS
INITIATIVE
Play a Part in Parkinson's Research

ADNI
ALZHEIMER'S DISEASE
NEUROIMAGING INITIATIVE

Human Connectome Project

OASIS
OPEN ACCESS SERIES
OF IMAGING STUDIES

Selection of target images

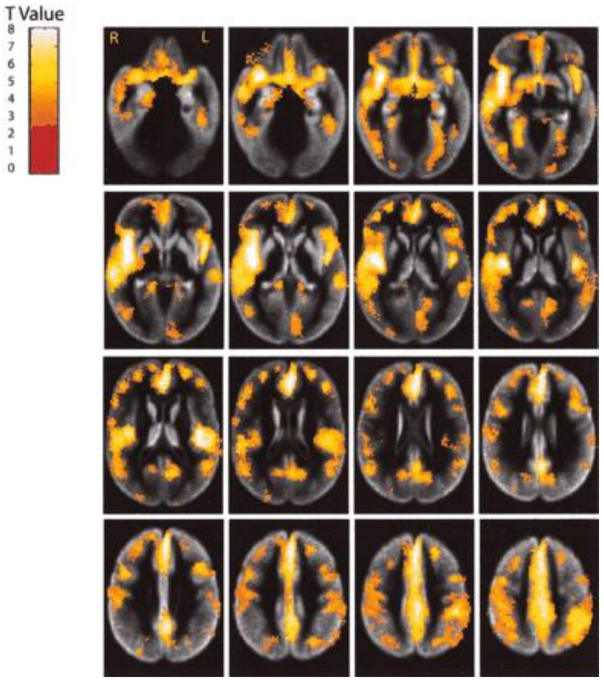
Considerations:

- Age
  - Cortex atrophies with age[1,2]
- Sex
  - Gray matter volume distributes differently in male vs. female brains[3,4]
- Disease state

Longitudinal Magnetic Resonance Imaging Studies of Older Adults: A Shrinking Brain

Susan M. Resnick, Dzung L. Pham, Michael A. Kraut, Alan B. Zonderman, and Christos Davatzikos

Journal of Neuroscience 15 April 2003, 23 (8) 3295-3301; DOI: https://doi.org/10.1523/JNEUROSCI.23-08-03295.2003



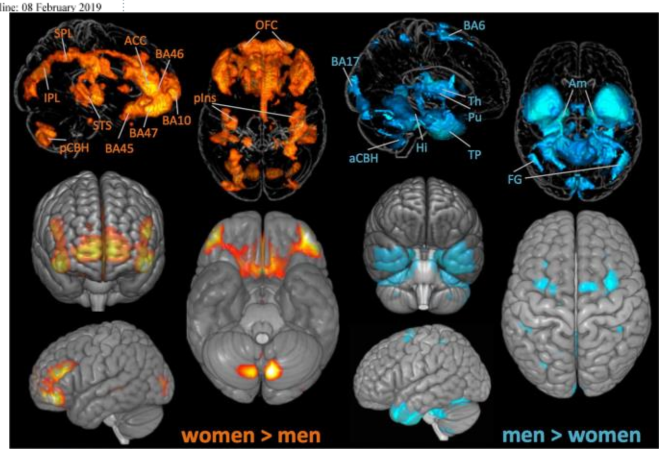
SCIENTIFIC REPORTS

OPEN

Novel findings from 2,838 Adult Brains on Sex Differences in Gray Matter Brain Volume

Received: 26 June 2018
Accepted: 18 December 2018
Published online: 08 February 2019

Martin Lotze[1], Martin Domin[1], Florian H. Gerlach[1], Christian Gaser[2], Eileen Lueders[3,4], Carsten O. Schmidt[5] & Nicola Neumann[1]

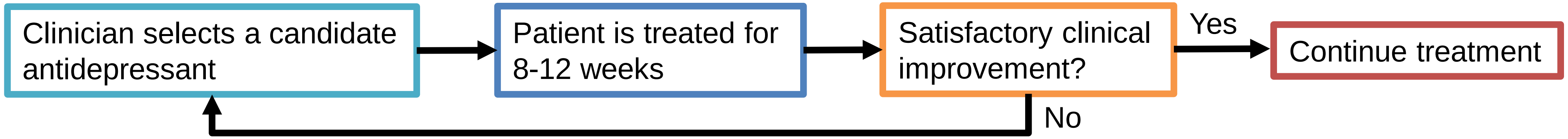


1. Resnick et al. Longitudinal Magnetic Resonance Imaging Studies of Older Adults: A Shrinking Brain. J Neurosci, 2003.
2. Good et al. A voxel-based morphometric study of ageing in 465 normal adult human brains. Neuroimage, 2001.
3. Lotze et al. Novel findings from 2,838 Adult Brains on Sex Differences in Gray Matter Brain Volume. Sci Reports, 2019.
4. Ritchie et al. Sex Differences in the Adult Human Brain: Evidence from 5216 UK Biobank Participants. Cerebral Cortex 2018.

Application: depression treatment outcome prediction

Major Depressive Disorder (MDD) is a leading cause of disability with 16% lifetime prevalence[1].

- Individual antidepressant response is unpredictable, each drug has a ~40% response rate
- Treatment selection is largely based on trial-and-error



EMBARC dataset[2]

- 163 subjects treated with sertraline for 8 weeks
- Structural and task-based fMRI acquired before treatment
- fMRI uses a number-guessing task that probes reward processing circuitry

<u>Deep learning task: use pre-treatment fMRI to predict individual outcomes to sertraline treatment (change in Hamilton Rating Scale for Depression, HAMD)</u>

1. Kessler et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA 2003
2. Trivedi et al., Establishing moderators and biosignatures of antidepressant response in clinical care (EMBARC): Rationale and design. J. Psych. Res. 2016

Data preparation

fMRI data was augmented 5-fold:

- For each **source** subject, 5 age- and gender-matched **target** subjects were selected from a separate treatment group (placebo) from the dataset
- 163 subjects → 978 subjects after augmentation
  - Model takes ~600 input features
  - Subjects:features ratio is improved 163:600 → 978:600
- Proposed method (**nonlinear** registration) was compared to a basic **affine** registration approach
- Augmented data used in training set only

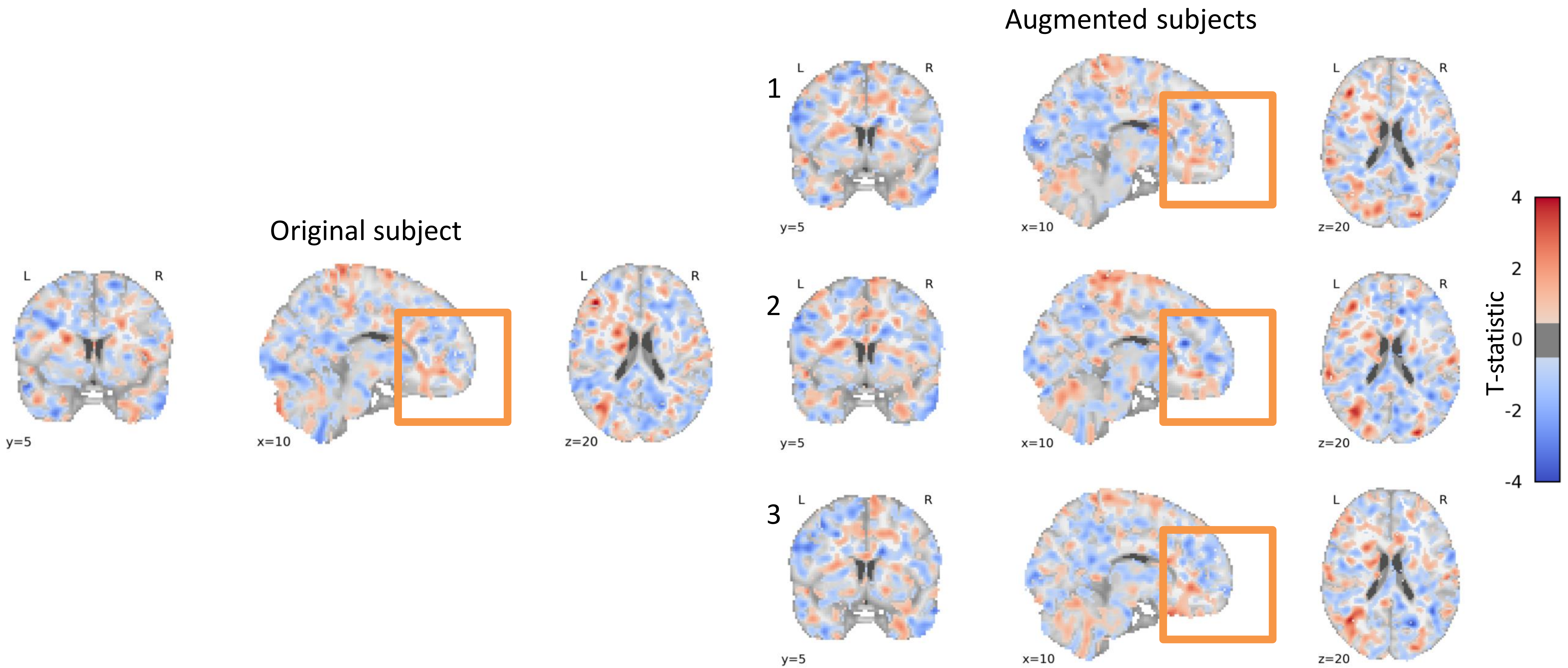
fMRI preprocessing:

Head motion correction → Brain extraction → Spatial normalization to MNI152 EPI template → Smoothing with 6 mm FWHM Gaussian kernel

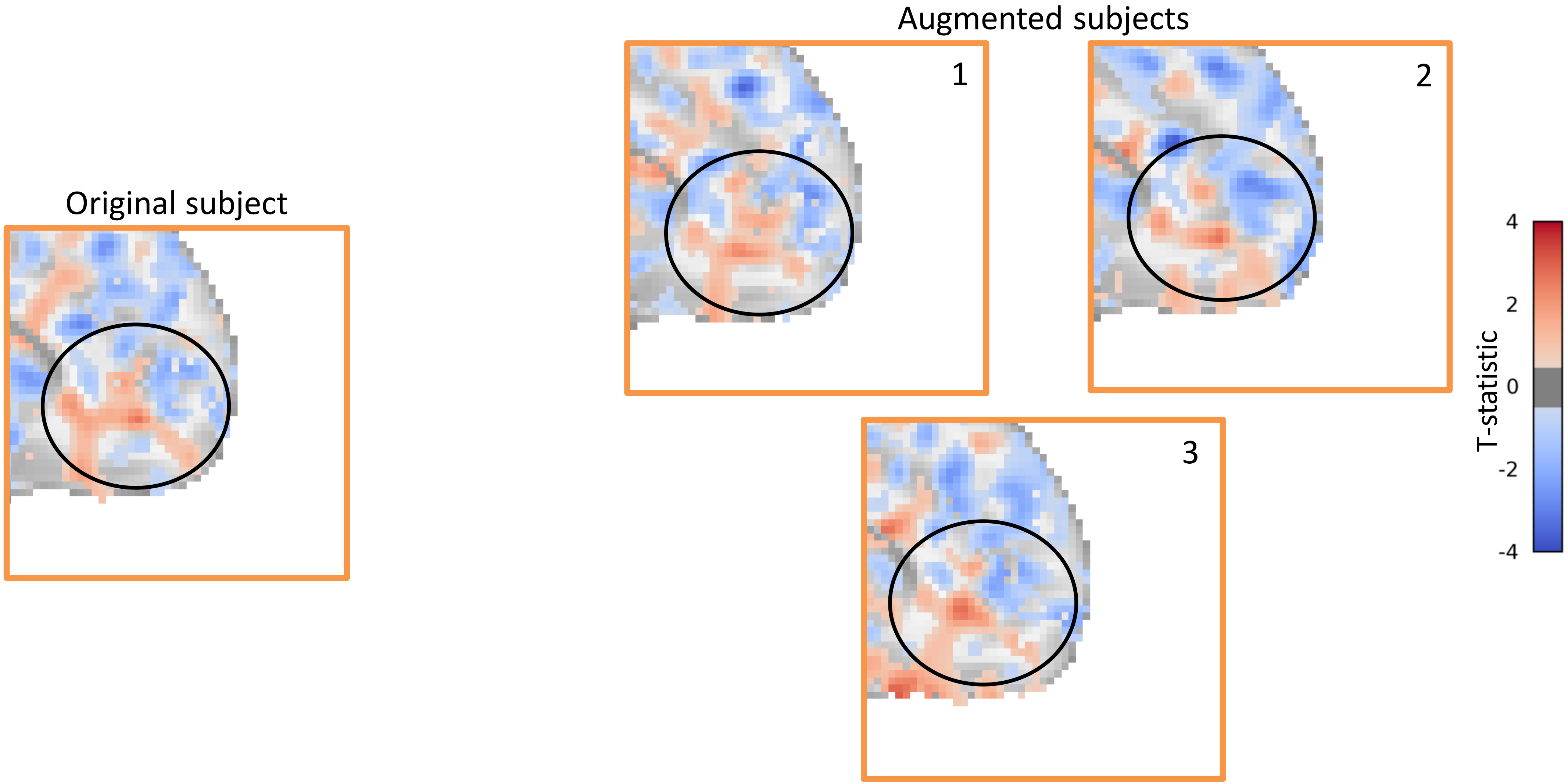
Feature extraction:

- 1st-level GLM fitted to task conditions → voxel-level contrast maps
- Parcellation with study-specific functional atlas → ~600 mean regional contrast values

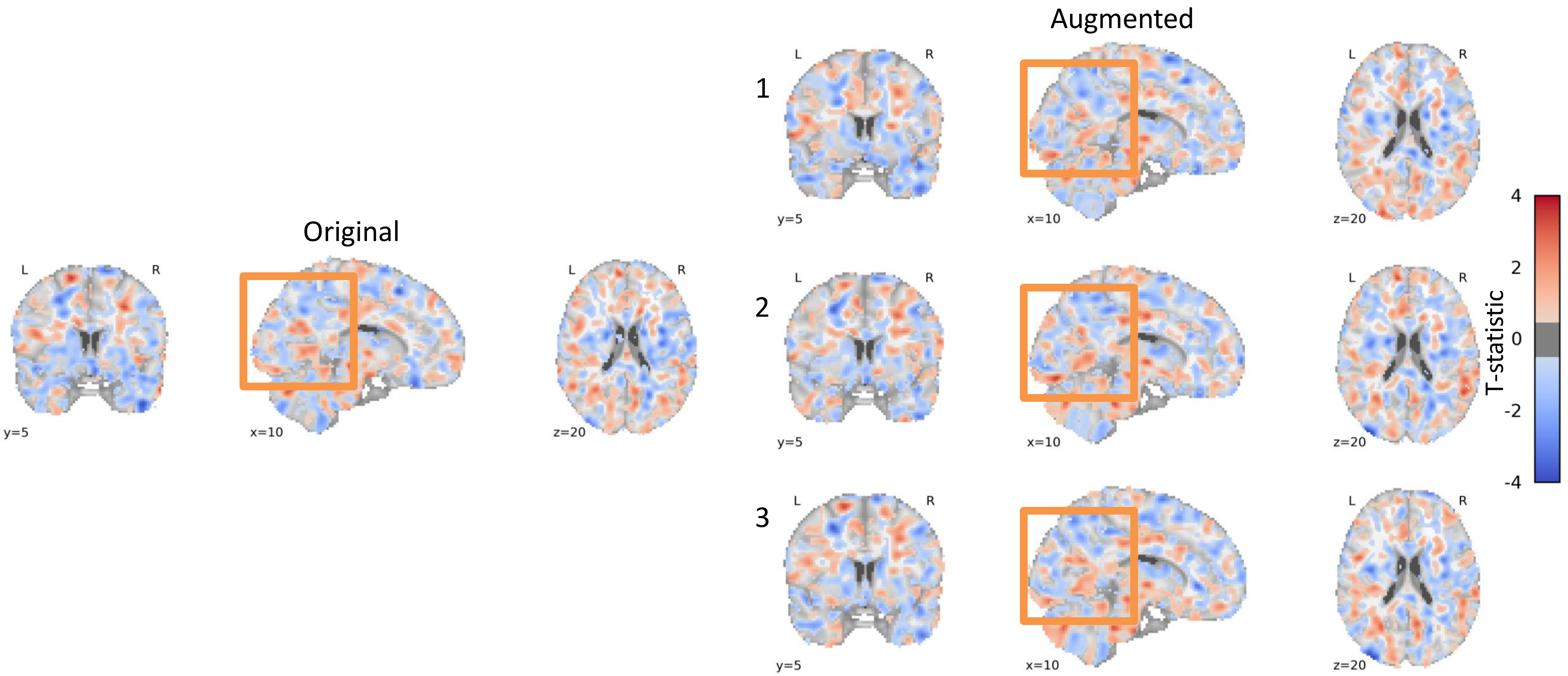
Contrast maps derived from augmented fMRI – example 1



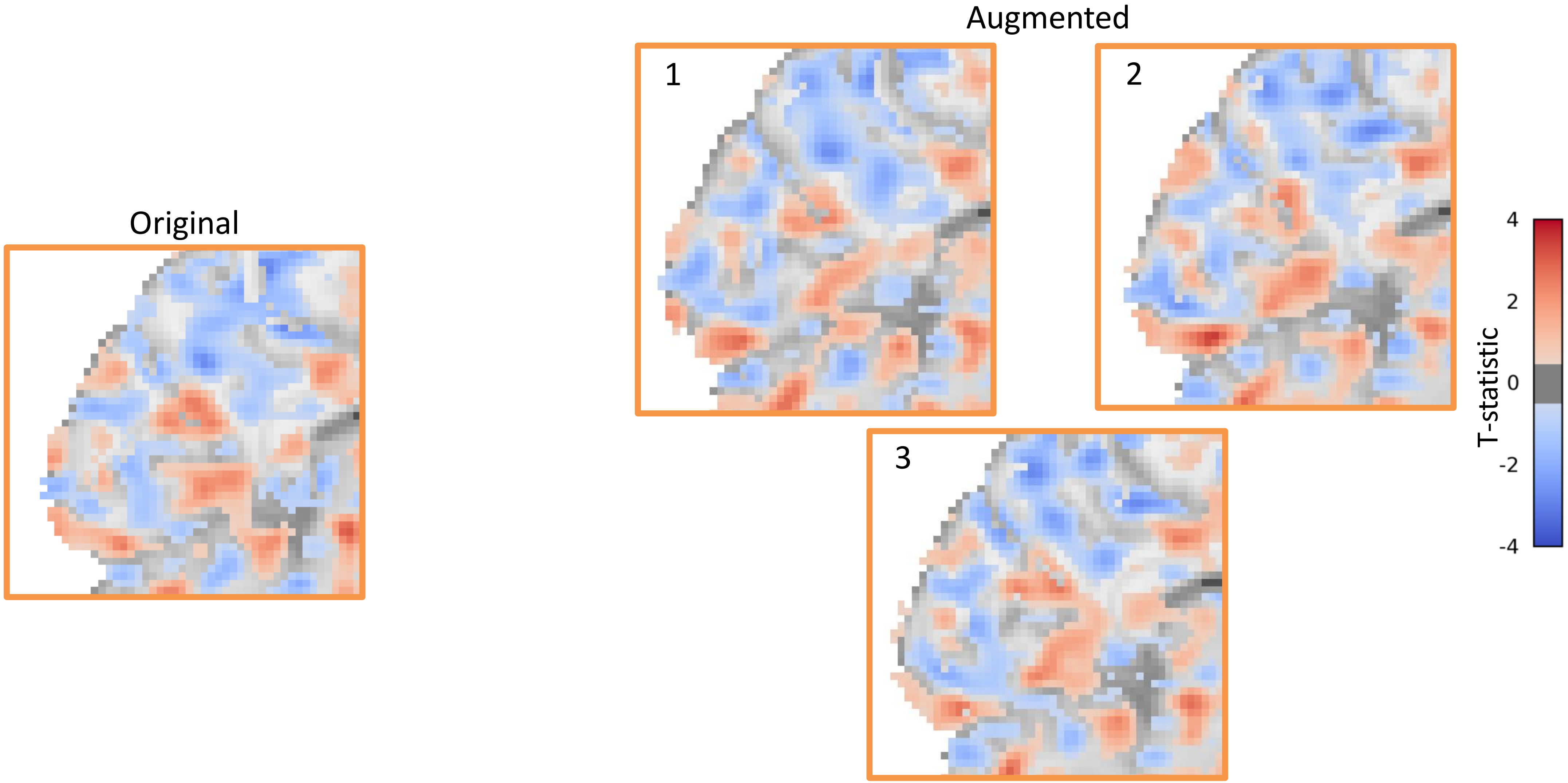
Contrast maps derived from augmented fMRI – example 1



Contrast maps derived from augmented fMRI

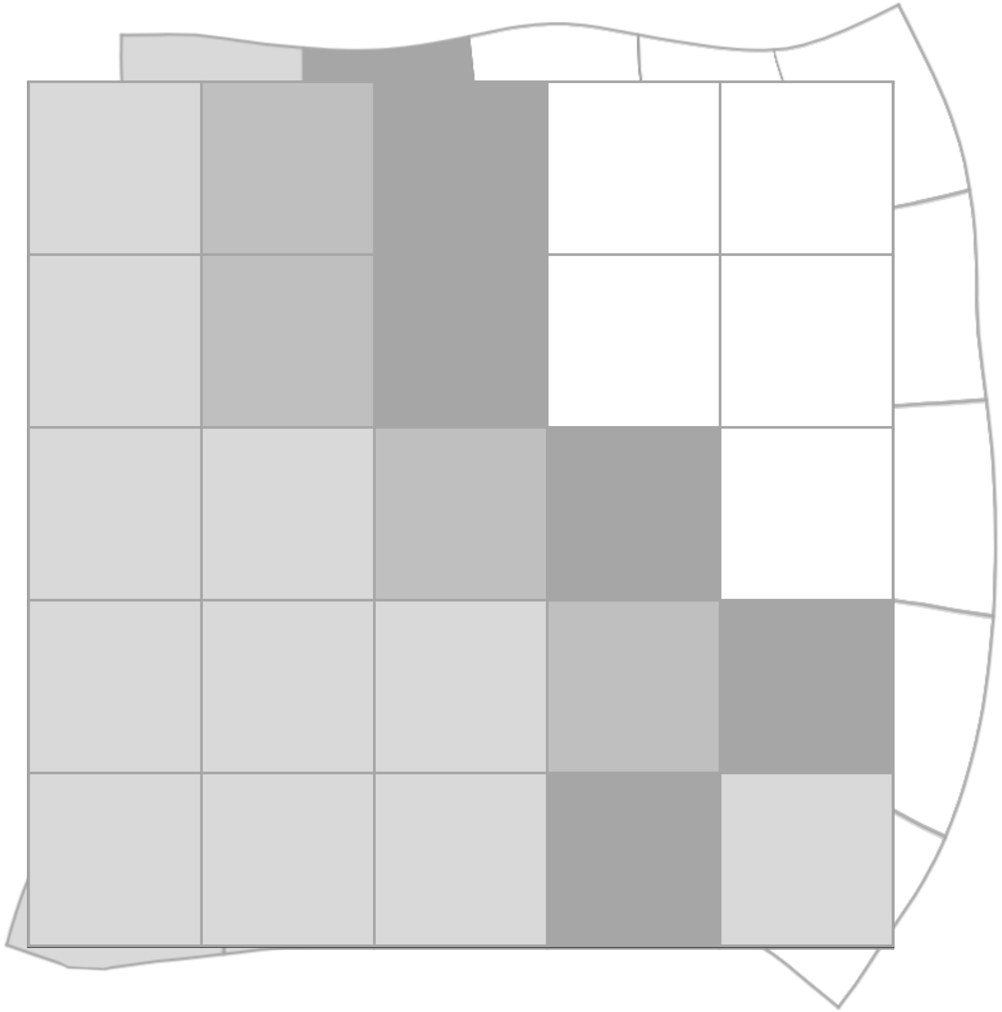


Contrast maps derived from augmented fMRI



Effects of nonlinear transformations

Nonlinear warp field



Model training, optimization, and validation

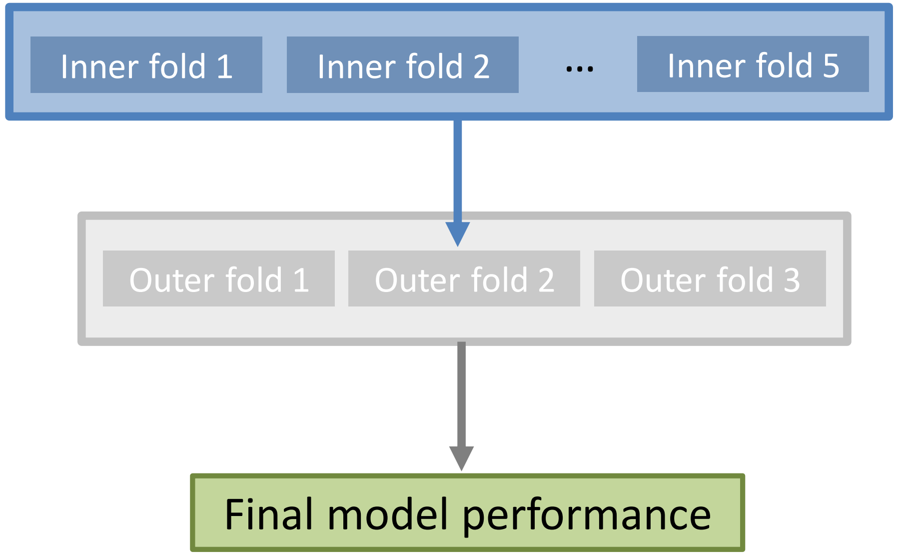
Feed-forward fully-connected neural networks were trained to predict treatment outcome (ΔHAMD) from the regional contrast values

Model hyperparameters were optimized using random searches[1] over 300 hyperparameter configurations, with 3 x 5 nested K-fold cross-validation

- Number of layers
- Layer size
- Activation function
- Weight regularization, batch normalization, dropout rate
- Learning rate
- Atlas granularity (number of regions)

3 model searches conducted

- No augmentation (**baseline**)
- **Proposed** augmentation method
- Basic **affine** augmentation



1. Train and evaluate model in each inner fold

2. Select best model based on R^2, retrain on all data from inner folds

3. Evaluate best model in each outer fold, report mean performance over 3 models

1. Bergstra and Bengio, Random Search for Hyper-parameter Optimization, JMLR 2012.

Model search results

The proposed augmentation increased model performance and outperformed the basic affine augmentation.

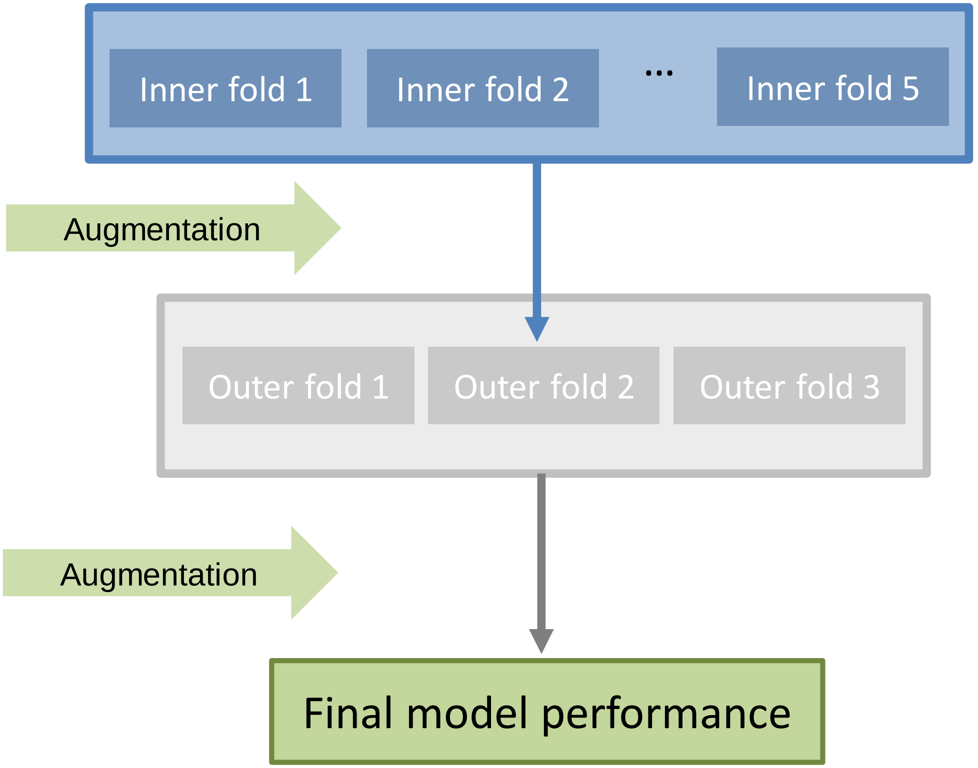
Augmentation method	RMSE	R^2
Baseline (no augmentation)	6.57	0.112
Proposed (nonlinear)	**6.46**	**0.141**
Affine	6.53	0.114

Differences in performance were significant at $p < 0.001$ after retraining each model 100 times with random reinitializations.

Where is augmentation making a difference?

Model search
(optimize model *hyperparameters*)

Final model training
(optimize model *parameters*)



1. Train and evaluate model in each inner fold

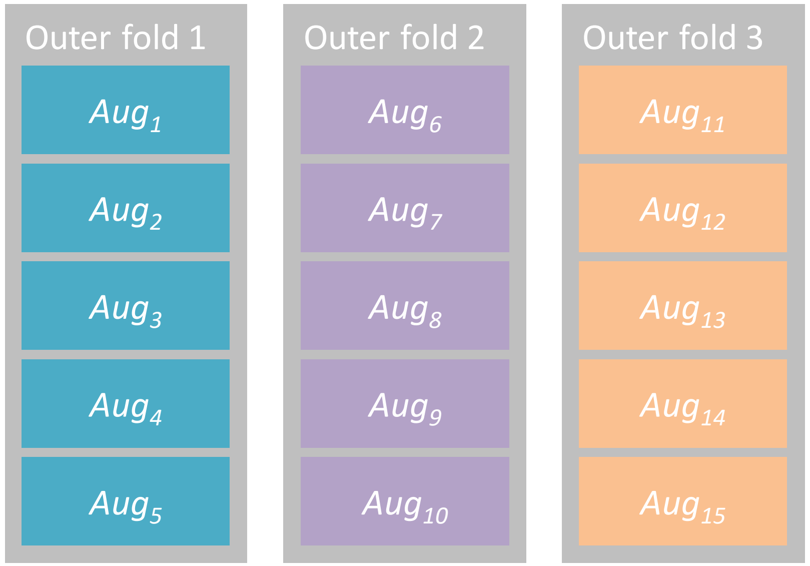
2. Select best model based on R^2, retrain on all data from inner folds

3. Evaluate best model in each outer fold, report mean performance over 3 models

Impact of augmentation on model training (parameter optimization)

Ablative experiment:

The top 5 models from each outer fold in *Aug* search (Aug_1, Aug_2,…, Aug_{15}) were retrained **without** augmented data.

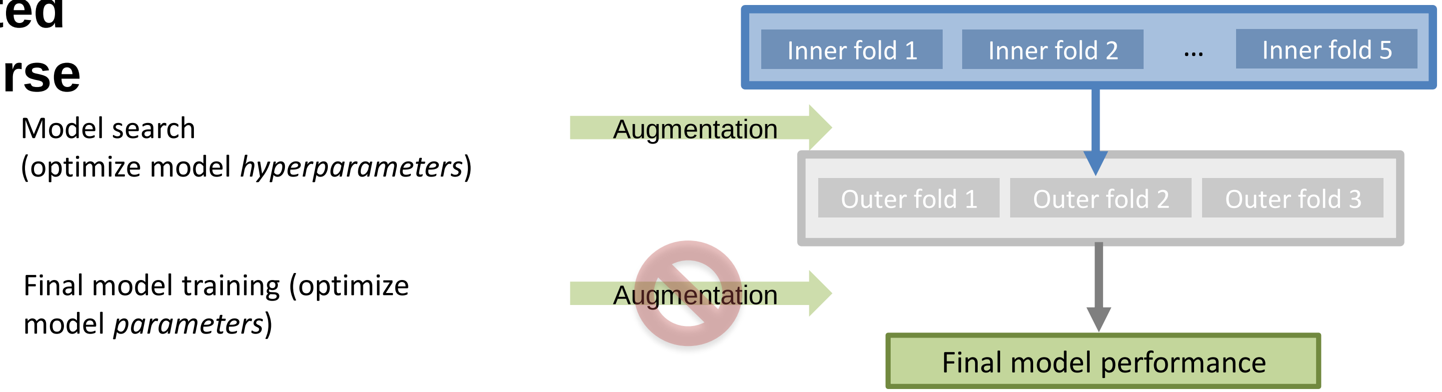


<u>Performance decreased significantly:</u>

R^2 decreased by 0.058 ± 0.051 ($p = 0.0006$)

RMSE increased by 0.21 ± 0.18 ($p = 0.0006$)

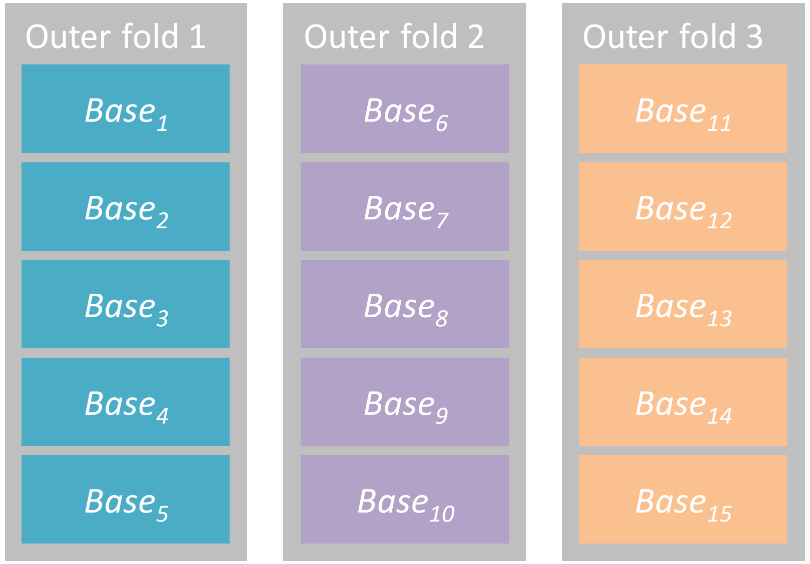
Hyperparameter configurations selected with augmented searches perform worse without augmented training.



Impact of augmentation on model search (hyperparameter optimization)

Additive experiment:

The top 5 models from each outer fold in *Base* search (*$Base_1$*, *$Base_2$*, …, *$Base_{15}$*) were retrained **with** augmented data.

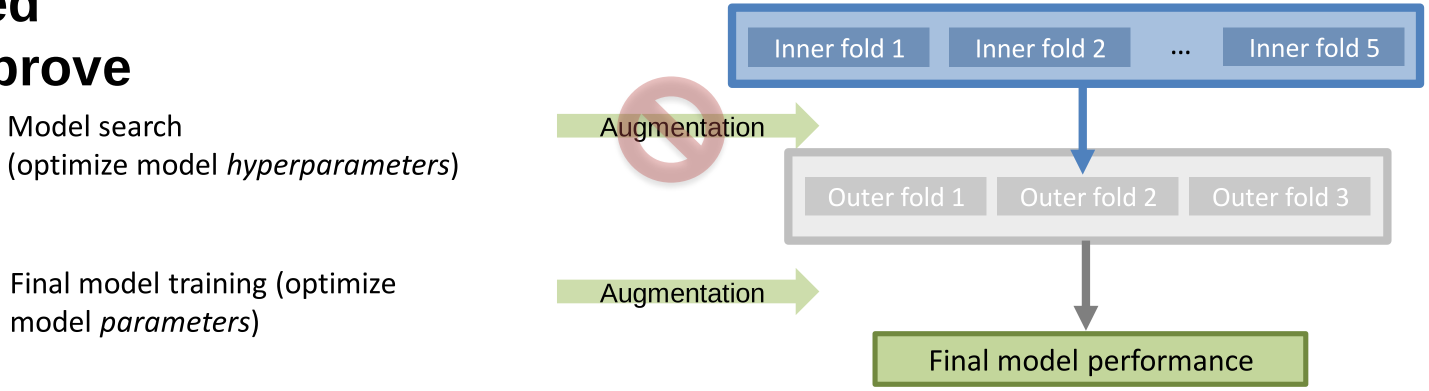


Performance did not change significantly:

R^2 increased by 0.015 ± 0.044 ($p = 0.209$)

RMSE decreased by 0.05 ± 0.16 ($p = 0.214$)

Hyperparameter configurations selected without augmented searches fail to improve with augmented training.



Conclusions

We propose a parameter-free fMRI data augmentation method that demonstrates high performance benefit in a deep learning prognostic problem.

Augmenting the data 5x with the proposed method increased performance in antidepressant outcome prediction by a 26% in R^2 (relative).

- This is consistent with natural image augmentation methods, e.g. AutoAugment and PBA (22-37% performance boost).

Basic affine augmentation had no significant performance benefit, in comparison.

The most benefit comes from using augmented data throughout **both** model search and final model training.
Model search (hyperparameter optimization) on limited data can result in less statistically powerful models that fail to increase performance on additional data in the future

Limitations:

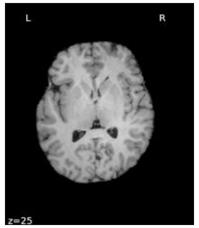
- Nonlinear registrations are computationally intensive (~30 minutes per augmentation, but highly parallelizable)
- Experiments were limited to 5x augmentation, but we already see a benefit
- One application was shown (MDD and task-fMRI), but we anticipate extensibility to other MRI contrasts, datasets

Cubuk et al. AutoAugment: Learning Augmentation Policies from Data. CVPR 2019.
Ho et al. Population Based Augmentation: Efficient Learning of Augmentation Policy Schedules. ICML 2019.

Acknowledgements

Deep Learning for Precision Health Lab
Department of Bioinformatics

Albert Montillo

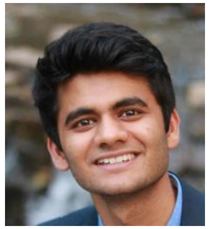


Krishna Kanth Chitta

Cooper Mellema

Son Nguyen

Alex Treacher



Vyom Raval

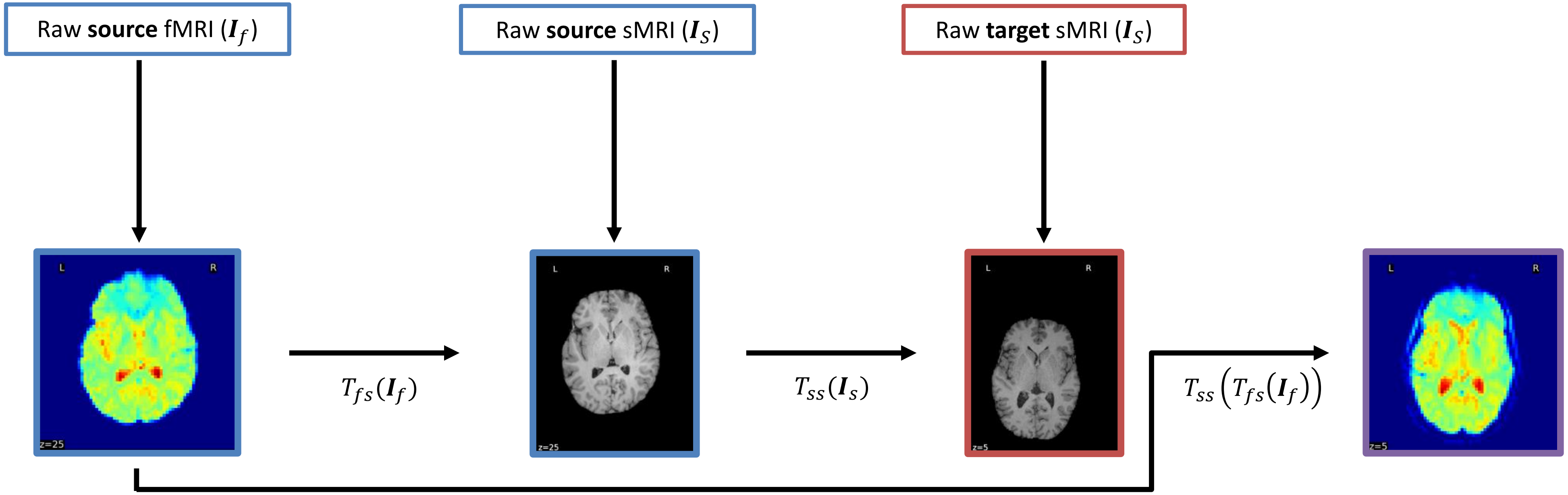
Jana Windsor

Center for Depression Research and Clinical Care
Department of Psychiatry
Madhukar Trivedi
Cherise Chin Fatt

Author Contact:
kevin3.nguyen@utsouthwestern.edu

Post-doc position in neuroimaging & machine learning available
UT Southwestern, Dallas TX
Contact PI Albert Montillo: albert.montillo@utsouthwestern.edu

Summary



Augmentation method	RMSE	R^2
Baseline (no augmentation)	6.57	0.112
Proposed (nonlinear)	**6.46**	**0.141**
Affine	6.53	0.114

Contact:
Kevin3.Nguyen@utsouthwestern.edu