Age prediction and site based bias correction

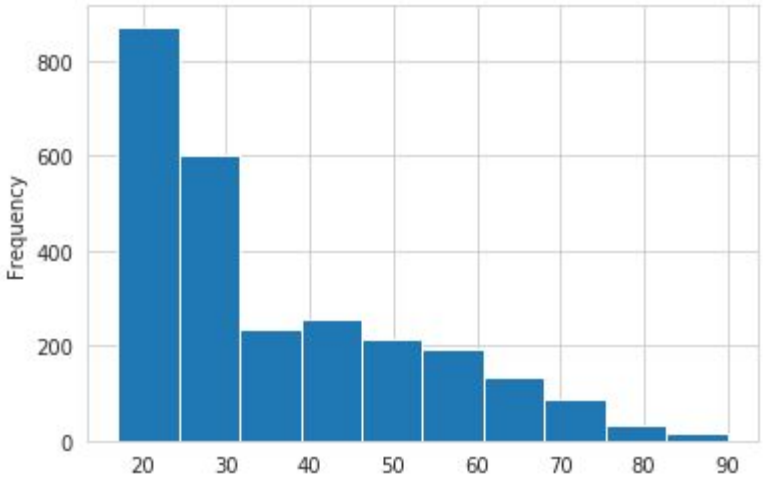
Overview

What we are going to talk about:

- age prediction via a convolutional neural network
- impact of the site on the performance of the predictor and a possible correction
- Future improvements

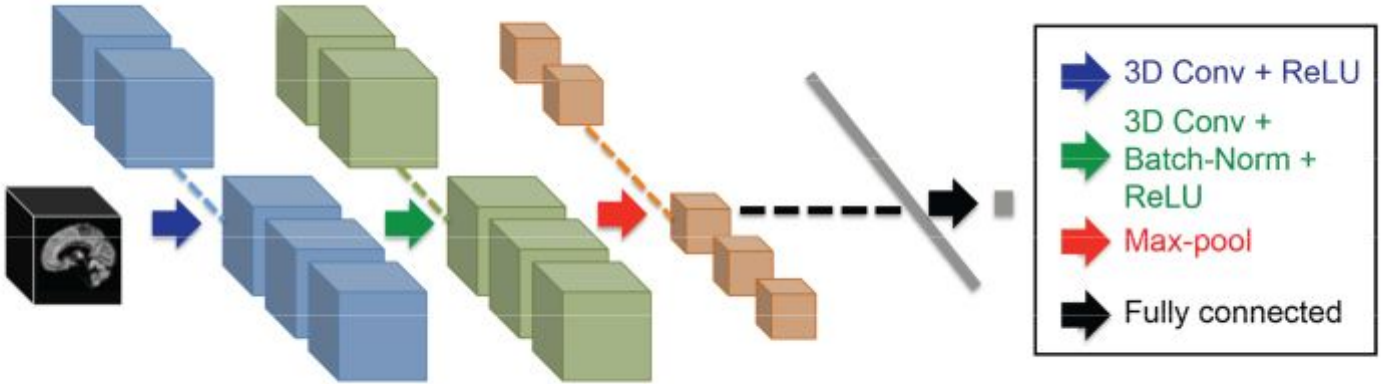
Dataset & preprocessing procedure

- Dataset composition:
  - 2647 T1-weighted MR images
  - 17 different acquisition sites
  - 53 % males, 47 % female
- Preprocessing:
  - registration to the MNI space
  - resampling to the same dimension: 181,217,181
  - Min-max normalization of the intensity value to the (-1,1) range



Age distribution of the dataset

CNN architecture



5 convolutional blocks.
The number of filters is 8, at start, and doubled after each block. The kernel size is 3x3x3.
Pooling is carried out on 2x2x2 volumes.
The fully connected part is composed by 19200 neurons connecting to the prediction one.

J. H. Cole *et al.*, “Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker,” *NeuroImage*, vol. 163, pp. 115–124, Dec. 2017, doi: 10.1016/j.neuroimage.2017.07.059.

Achieved performances

The architecture of the model has been tested using different hyper-parameters and optimizers (adam, sgd, rmsprop).

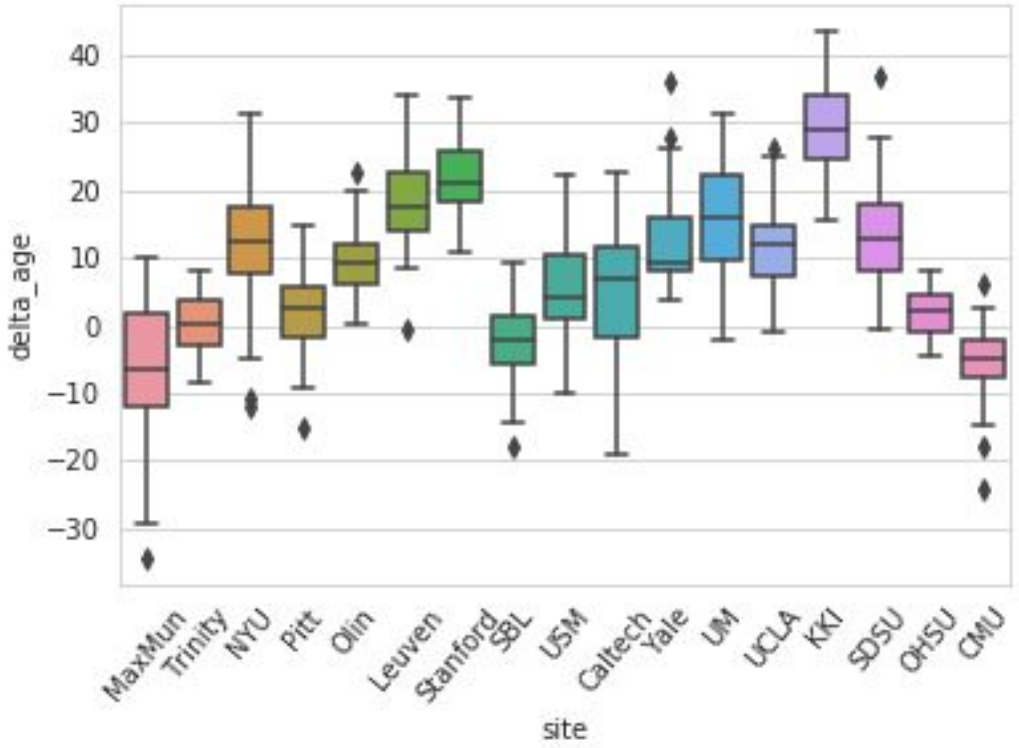
All the different setups ended up in a network capable of predicting age with a MAE (mean absolute error) of 5.5 year.

How is the prediction error distributed across different sites?

Let’s see how it behaves on the ABIDE dataset

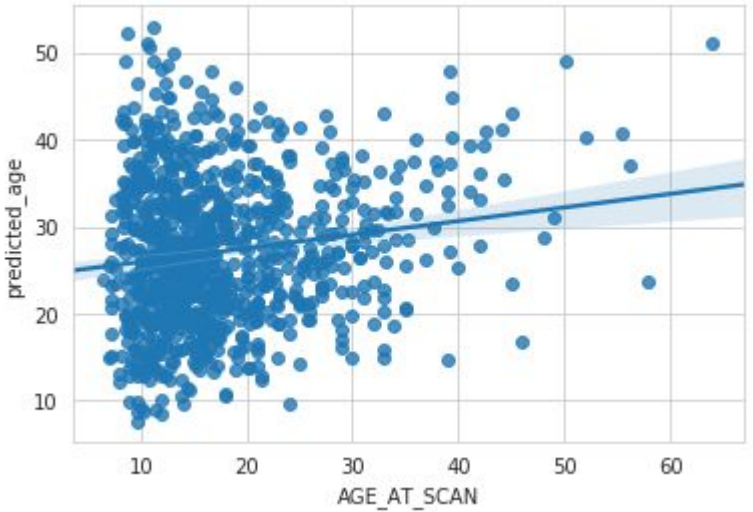
Considerations on the ABIDE dataset

site	delta_age(p-r)	abs_delta_age
CMU	-5.8	6.8
Caltech	4.8	9.1
KKI	29.0	29.0
Leuven	18.2	18.2
MaxMun	-5.7	8.8
NYU	12.3	13.0
OHSU	1.7	3.2
Olin	9.5	9.5
Pitt	2.4	5.0
SBL	-2.6	5.3
SDSU	13.7	13.8
Stanford	21.9	21.9
Trinity	0.5	3.5
UCLA	11.5	11.6
UM	15.9	16.0
USM	5.3	6.4
Yale	12.2	12.2

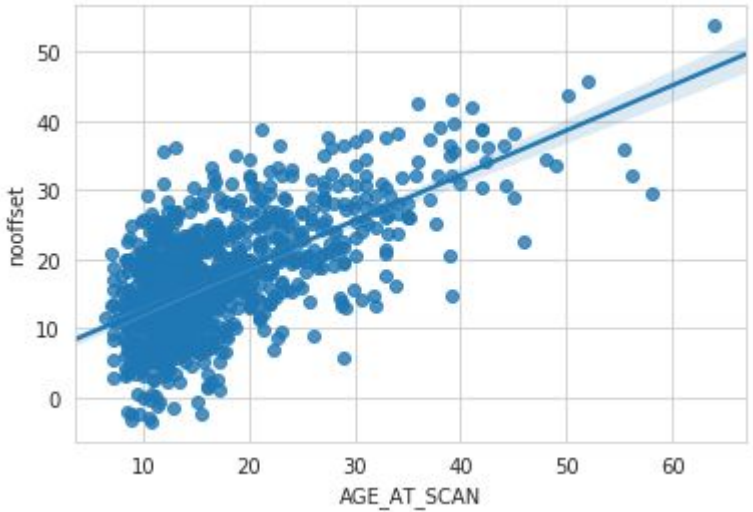


Site related offset

It seems likely that the site-effect can be modeled by means of a site-specific offset.

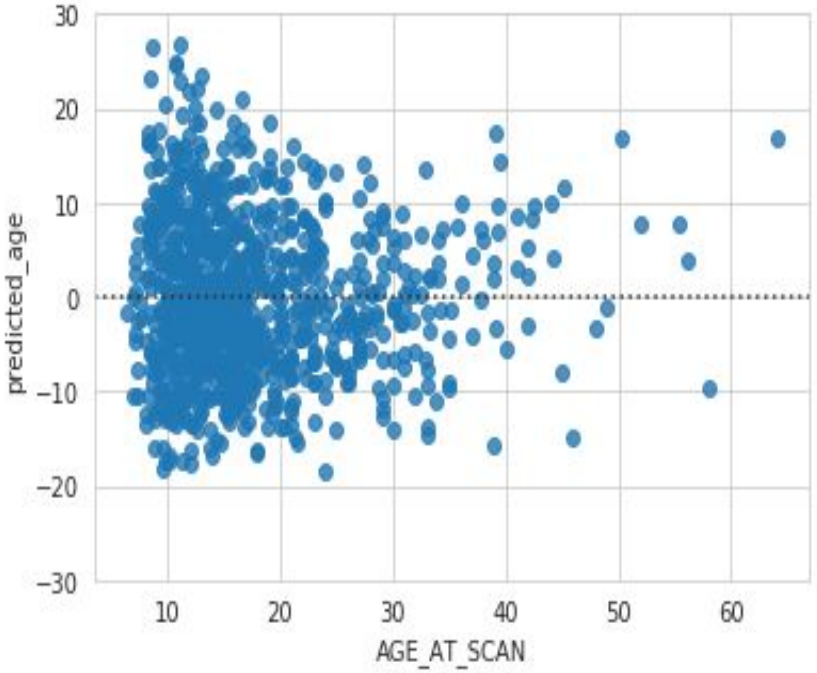


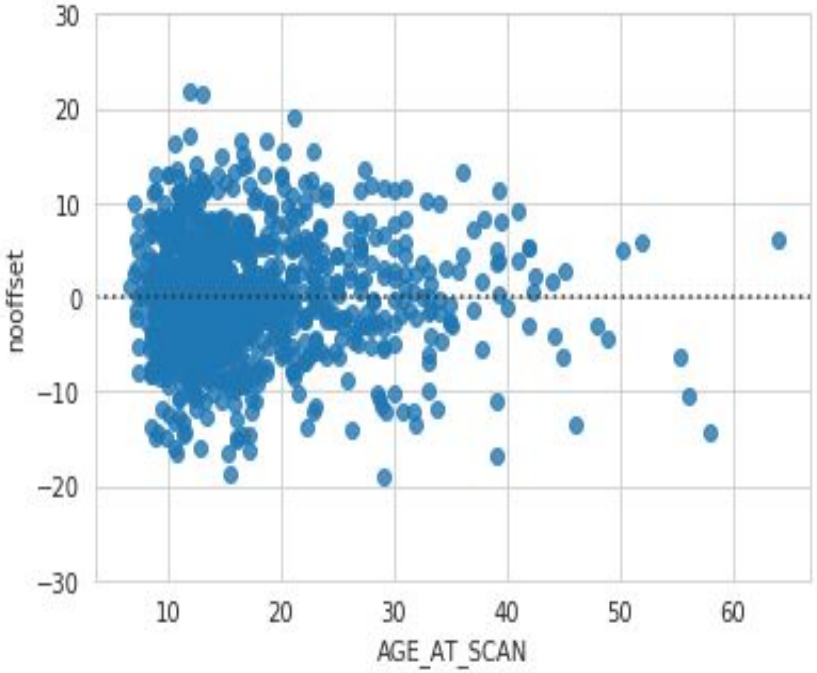
y = 0.16*x+34; R = 0.15



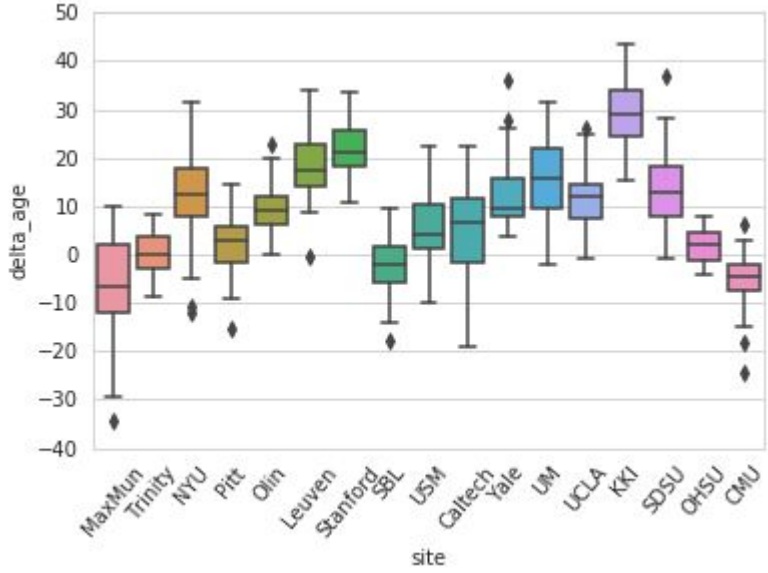
y = 0.65*x+6; R = 0.63

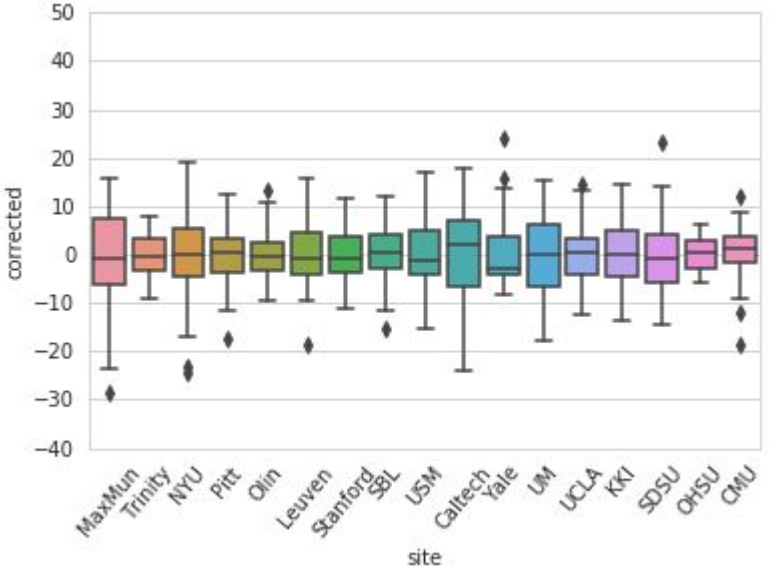
Residual plot





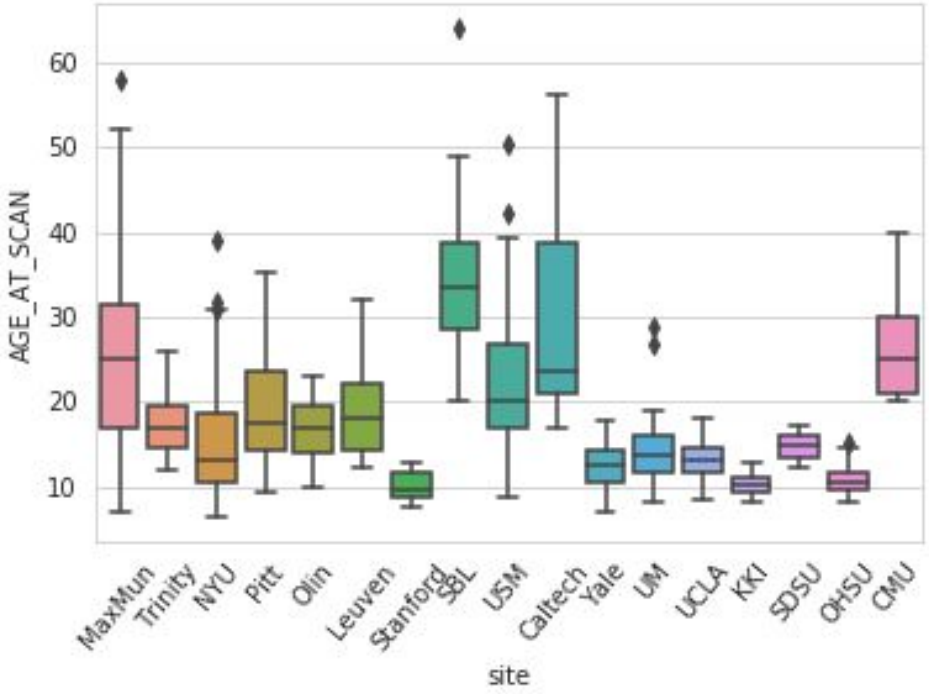
Original vs corrected boxplot





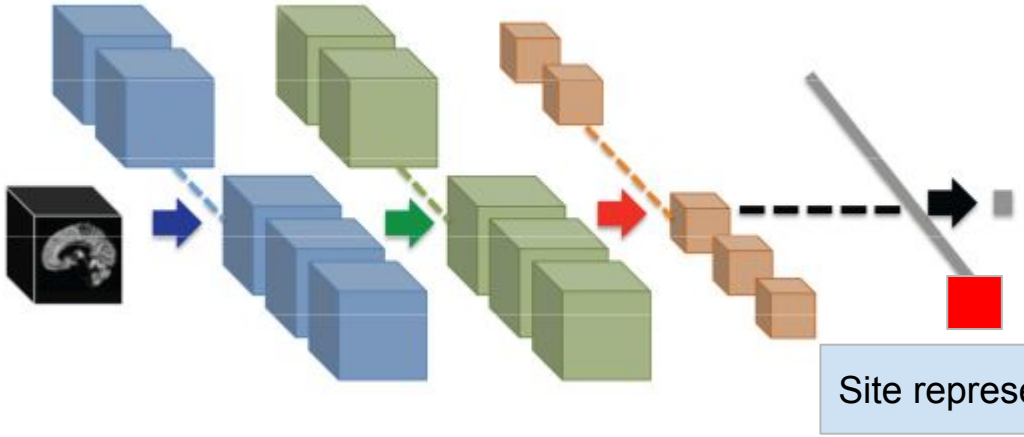
Future improvements

- fine tune the network on a narrower and more represented range of age such as 0-20y or 0-40y
- Implement a site-aware network architecture maybe using the pretrained network as a starting point and then introducing the sit into the fully-connected part (transfer learning)



Transfer learning example

Site could be represented as a single value or as a low-dimensional vector using well-established encoding strategies.



Pretrained part

Modified
Fully-connected part