Script

March 13, 2019

1 This demo uses dataset of raw T1 structual MRI images collected by Steve.

541 subjects in total.
from 6 to 85 years old.
10% for test, 90% for training.
Because of the incompatible in the pixdim and dim, we get rid of three images from the dataset:

- A00058503
- A00058952
- A00059344

Two images are missing according to the phenotypic information:

- A00040181
- A00039084

```
In [1]: %load_ext autoreload
        %autoreload 2
        %matplotlib inline
```

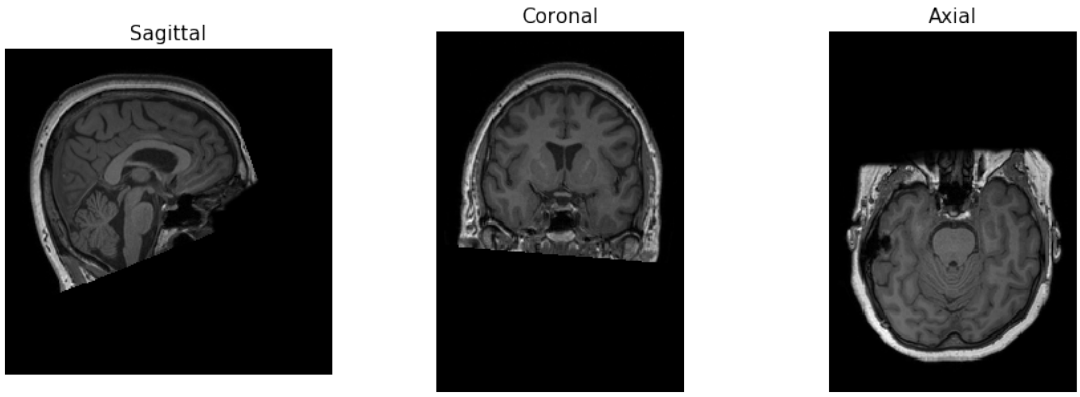
```
In [2]: from preprocess import *
```

1.1 1. The raw image

1.1.1 Some images looks like this:

```
In [12]: nii_img = nib.load('/media/woody/Elements/Steve_age_data/ANAT/A00008326.nii')
         npy_img = nii_img.get_data()
         print2d(npy_img)
```

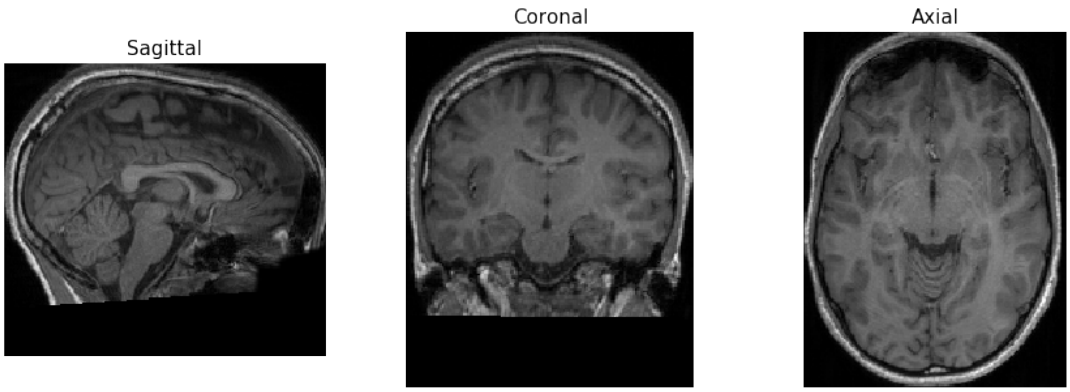
```
Dimension:  (176, 256, 256)
```



1.1.2 Some has different dimension:

```
In [13]: nii_img = nib.load('/media/woody/Elements/Steve_age_data/ANAT/A00037483.nii')
         npy_img = nii_img.get_data()
         print2d(npy_img)
```

```
Dimension:  (152, 205, 187)
```



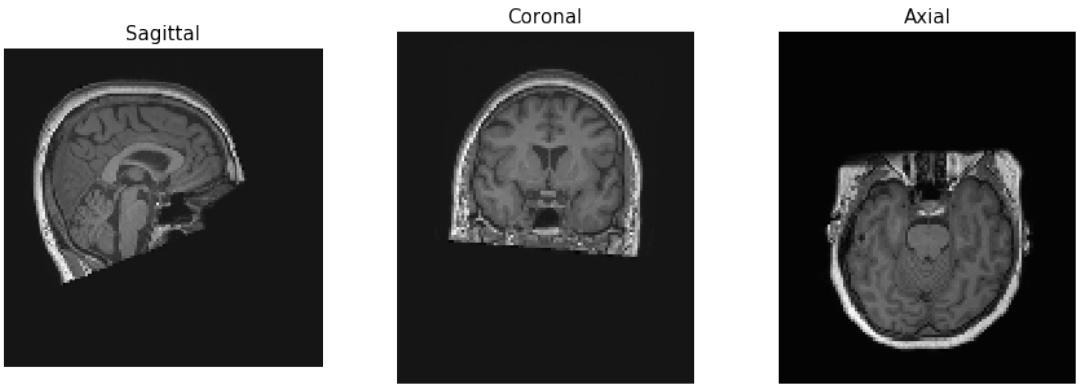
1.1.3 That's why we padded all of them into the same size (176,256,256)

1.1.4 Plus, since the limit in our GPU's memory, it is difficult to input bunches of such large matrix directly to our prediction model. So we resampled them from (176,256,256) to (110,130,130), which also complies with the dimension used in James Cole's article. However, for further study with multi-center dataset, pay attention that the images may need to be rotated.

1.1.5 After the padding and resampling processes, the image looks like this:

```
In [11]: crop_padded_img = np.load('./data_npy/origin/A00008326.npy')
         print2d(crop_padded_img)
```

```
Dimension:  (110, 130, 130)
```



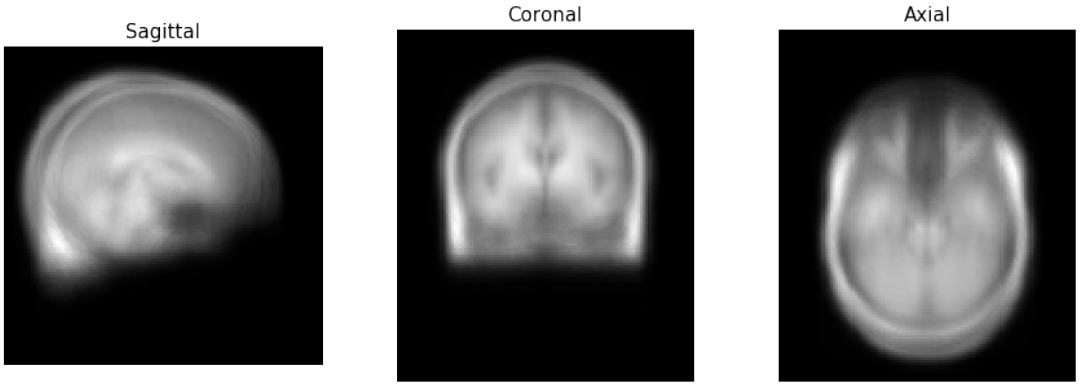
1.1.6 The mean image of all the training data is computed and is subtracted from all training and test data.

1.1.7 It is worth noting that the test data does not contribute to the mean image. This is because the training data, and only training data, needs to have zero mean for better training performance.

1.1.8 How does the mean values look like:

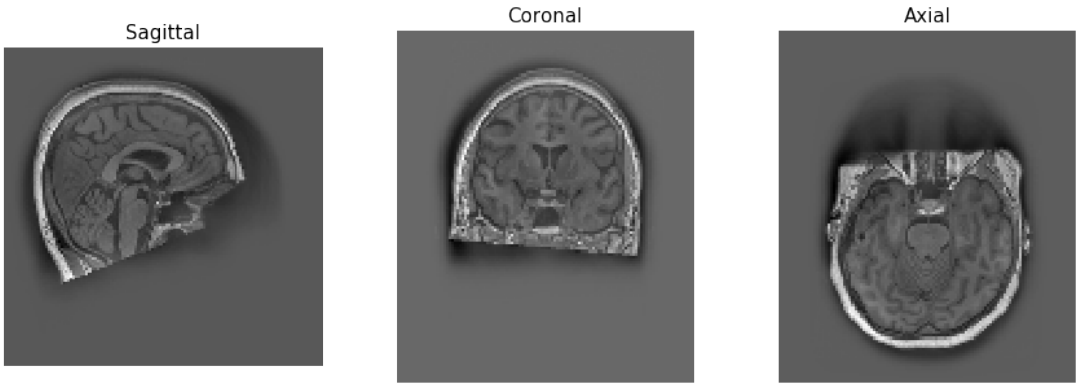
```
In [14]: mean_npy = np.load('./data_npy/mean_npy.npy')
         print2d(mean_npy)
```

```
Dimension:  (110, 130, 130)
```



1.1.9 Finally, the images fed into the model look like this:

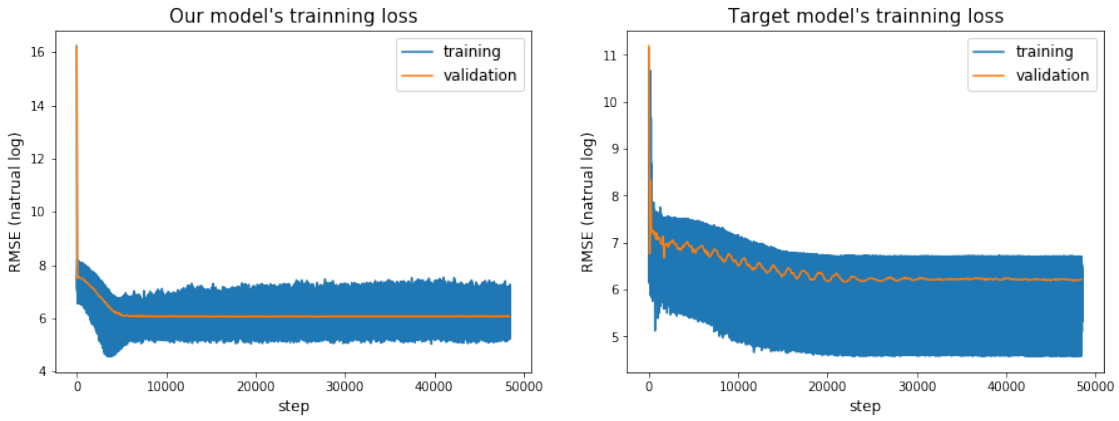
```
In [15]: final_img = np.load('./data_npy/mean_subtracted/A00008326.npy')
         print2d(final_img)

Dimension:  (110, 130, 130)
```



1.2 2. Training and test results

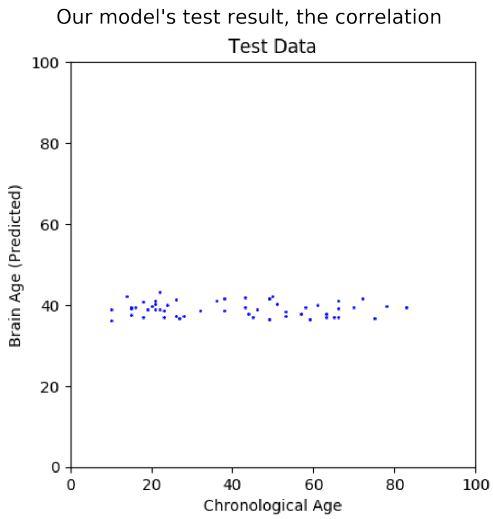
- We collect the best shot who has the smallest loss value (RMSE) in the training process.
- The validation uses the test dataset.
- Our model uses 3D-CNN with dropout, batch normalization and L2 regularization.
- Target model uses 3D-CNN with dropout.

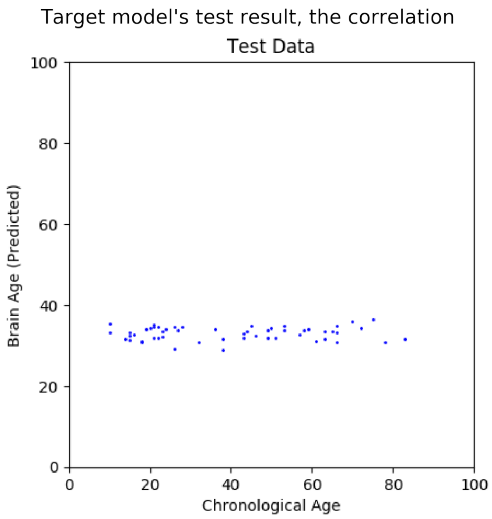
```
In [23]: def subdraw(ax,filename):
             arr = np.load(filename)
             ax.plot(arr[0], np.log(arr[1]),label='training')
             ax.plot(arr[2], np.log(arr[3]),label='validation')
             ax.legend(fontsize=12)
             ax.set_xlabel('step',fontsize=12)
             ax.set_ylabel('RMSE (natrual log)',fontsize=12)

         f, (ax1, ax2) = plt.subplots(1, 2, figsize=(15,5))
         subdraw(ax=ax1,filename='./img/demo_1_1_pltdata_2019.03.13.09:21:34.npy')
         ax1.set_title('Our model\'s trainning loss',fontsize=15)
         subdraw(ax=ax2,filename='./img/demo_1_2_pltdata_2019.03.13.15:39:09.npy')
         ax2.set_title('Target model\'s trainning loss',fontsize=15)

Out[23]: Text(0.5, 1.0, "Target model's trainning loss")
```



1.2.1 Let's check the correlation between the predicted age and the chronological age.

```
In [25]: f, (ax1, ax2) = plt.subplots(1, 2, figsize=(40,20))
         ax1.imshow(plt.imread('./img/demo_1_1_test.png'))
         ax1.axis('off')
         ax1.set_title('Our model\'s test result, the correlation',fontsize=40)
         ax2.imshow(plt.imread('./img/demo_1_2_test.png'))
         ax2.axis('off')
         ax2.set_title('Target model\'s test result, the correlation',fontsize=40)

Out[25]: Text(0.5, 1.0, "Target model's test result, the correlation")
```





1.3 Conclusion

1.3.1 It seems either this dataset endures the similar features as the IXI dataset that the prediction model can hardly estimate the brain age from the raw images, or there is something wrong with the model code.

1.3.2 We are in great need of figuring this problem out by using the ABIDE and ABIDEII dataset for training and test.

In [ ]: