Investigating brain ageing in HIV-positive individuals using neural networks

Rachel Catzel [1]

Jonathan Shock [1,2] Deshen Moodley [1]

[1]University of Cape Town, [2]INRS Montreal

Introduction

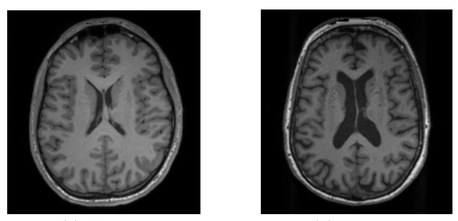
Machine learning techniques, and more recently, neural networks are being utilised in the field of health care and, for example, put to the task of predicting brain age. Studies have shown that this predicted brain age can be biomarker of brain health [1] and that for certain diseases this predicted brain age may be higher than the individual's chronological age. Due to the neurological effect HIV may have, there is a question of whether HIV accelerates brain ageing and what factors may affect this accelerated ageing. Being able to determine and monitor the progression of HIV in the brain is important to assess factors such as treatment efficacy. The use of neural networks to investigate brain ageing in HIV-positive individuals is relatively unexplored.

Research Questions

- Does HIV lead to an acceleration in brain ageing?
- Is this accelerated brain ageing correlated to an individual's CD4 count, nadir CD4 count, ART status and AIDS status?

Brain ageing

As one ages, both the body and brain change. Brain components, such as grey matter tend to decrease with age, while others, such as ventricular volumes increase with age. It has been seen that HIV may have a neurological impact on the brain and can cause, for example, atrophy and lesions. This leads us to question whether the brain ageing of individuals with HIV is accelerated.



(a) 18 years (b) 87 years

Figure 1. Illustration of brain changes that occur with age [2]

Methodology

- **Data pre-processing**: Re-orientate, skull strip and register T1-weighted brain MRI scans for input
- **Training**: Train, validate and test the network on dataset of healthy controls
- **Testing**: Apply trained network to testing dataset from ENGIMA-HIV Working Group
- This is a regression problem as age falls along continuous spectrum
- The network takes axial brain slices as input and outputs a predicted brain age

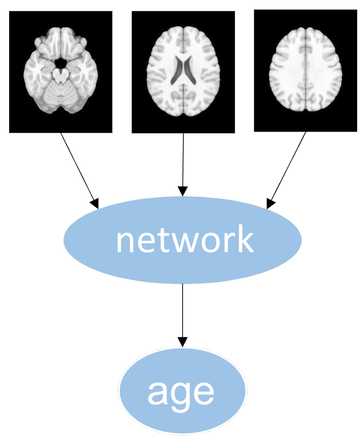


Figure 2. Network pipeline

Experimental Design

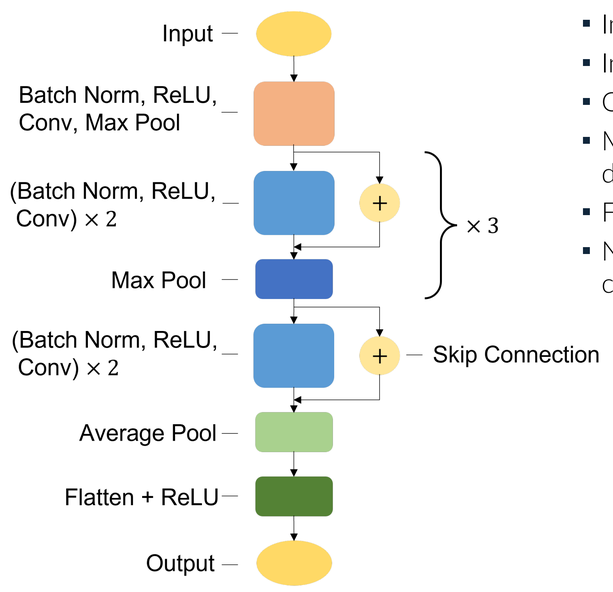


Figure 3. Network architecture

- Implement a CNN with four residual blocks
- Input of spatial dimensions 225×225×225
- Output scalar age value
- Max pooling for downsizing of spatial dimensions
- Feature maps increase in order 32, 64, 128, 256
- NVIDIA Tesla V100 SXM2 GPU used for computations

Datasets

We acquire datasets of healthy controls, for training a network to map brain scans to age values, and a dataset of individuals with HIV for investigating the research questions.

Dataset	Size	Age Range (y)	Status
UK Biobank	21366	45-82	Controls
Cam-CAN	652	18-88	Controls
ENIGMA-HIV	1194	17-82	HIV+ and controls

Table 1. Details of datasets used

References

[1] James H. Cole and Katja Franke. *Predicting age using neuroimaging: Innovative brain ageing biomarkers*. Trends in Neurosciences, 40:681-690, 2017

[2] D Taylor, *Saliency Mapping in Convolutional Neural Networks to Determine Brain Age Trajectories*, MSc thesis, University of Cape Town, Cape Town, 2022.

Analysis

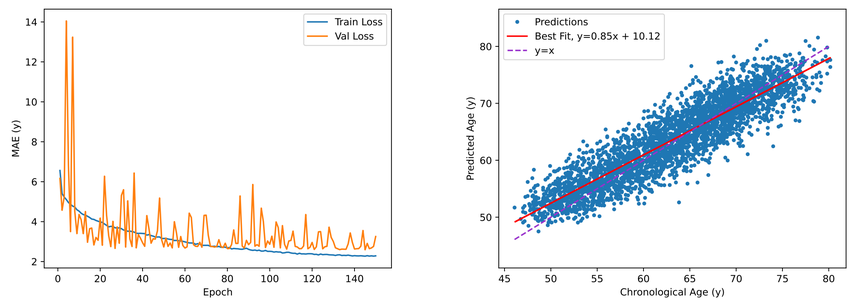
- Network predicted age used to calculate brain age delta:

$$\delta = age_{pred} - age_{chron}$$

- Correlation between δ and CD4 count, nadir CD4 count, ART status, and AIDS status to be analyzed
- Confounding factors include lack of information on duration of disease, treatment timeline, and date of nadir CD4 count measurement
- Not all individuals in the dataset have details on ART and AIDS status

Preliminary Results

We first pre-train the network on the large UK Biobank dataset.



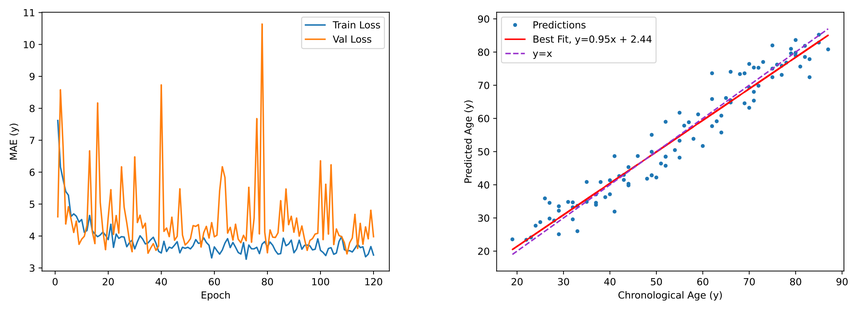
(a) Training and validation loss (b) Correlation plot

Figure 4. Pre-training on UK Biobank dataset

Dataset	Min Train MAE (y)	Min Val MAE (y)	Test MAE (y)	r	δ (y)
UK Biobank	2.27	2.59	2.57 ± 1.94	0.90	0.46

Table 2. Results of pretraining network on UK Biobank dataset

Next we transfer learn with the Cam-CAN dataset to expose the network to a wider age range.



(a) Training and validation loss (b) Correlation plot

Figure 5. Transfer learning on Cam-CAN dataset

From Tab. 3 we see that pre-training improves the accuracy obtained on the Cam-CAN dataset.

Pretraining	Min Train MAE (y)	Min Val MAE (y)	r	MAE (y)	δ (y)
Without	5.44	5.95	0.92	5.86 ± 4.36	-1.48
With UKB	3.53	3.44	0.97	3.74 ± 2.63	-0.32

Table 3. Comparison of results with and without pretraining with UK Biobank dataset

Following transfer learning we input the ENIGMA-HIV dataset to the trained network and analyse results. We first analyse results for cohorts containing HIV+ and HIV- groups for accurate comparison.

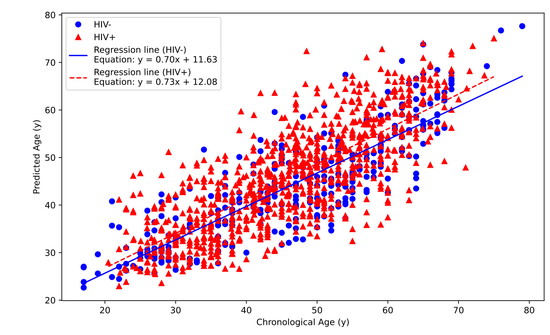


Figure 6. Comparison of predicted age verse chronological age, with line of best fit, for HIV+ and HIV- individuals in testing dataset

	$\delta_+ \pm \sigma_+$ (y)	n_+	$\delta_- \pm \sigma_-$ (y)	n_-	t-statistic	p-value
Arch	3.25 ± 5.87	79	2.83 ± 6.06	55	0.400	0.345
Brown	-3.65 ± 5.96	76	-3.56 ± 5.85	47	-0.082	0.533
Hinkin	-6.47 ± 4.85	13	-7.32 ± 4.58	52	0.567	0.286
Valcour	-3.91 ± 5.71	53	-4.41 ± 4.76	30	0.422	0.337
Boban	-0.04 ± 6.26	81	2.78 ± 5.62	76	-2.972	0.998
Cysique	-9.15 ± 6.37	82	-10.14 ± 4.71	40	0.966	0.168

Table 4. Summary of mean and associated standard deviation, sample size, t-statistic, and p-value for the δ values within cohorts featuring HIV+ and HIV- subsets. We denote the HIV+ subset with a " + " subscript and HIV- with " − " subscript.

While preliminary results do indicate that on average a more positive δ is obtained for the HIV+ subgroup compared to the HIV- subgroup, current results are not statistically significant and further investigation is required.

Future Work

- Harmonisation of scans from different sites in testing dataset
- Compute statistics for age range subsets
- Analysis of correlations

For queries contact rachelcatzel@gmail.com