Does pre-training on brain-related tasks results in better deep-learning-based brain age biomarkers?

Bruno M. Pacheco, Victor H. R. de Oliveira, Augusto B. F. Antunes,
Saulo D. S. Pedro, and Danilo Silva

UFSC

3778 GAMA

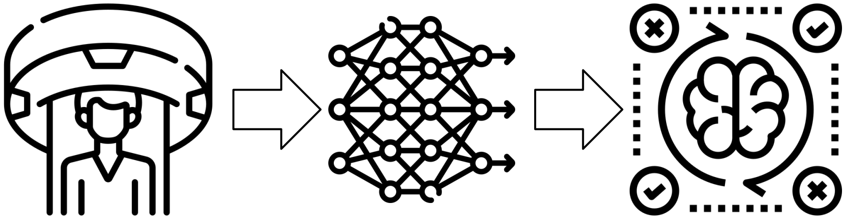
Deep Learning & Brain Age

Aging patterns of the brain are associated with several pathologies.

Train a deep learning model to quantify these aging patterns (brain age) from MRIs.

Hypothesis: brain age (BA) of healthy subjects is equal to their chronological age (CA)

1. Regression task with X=MRI and y=CA (healthy subjects only)
2. CNN quantifies aging patterns through BA
3. Δ_{BA} = BA - CA
   - Healthy $\Rightarrow \Delta_{BA} \approx 0$
   - Unhealthy $\Rightarrow \Delta_{BA} \gg 0$



Franke, K. et al. (2010). Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: Exploring the influence of various parameters. NeuroImage, 50(3), 883–892.
Cole, J. et al. (2017). Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. NeuroImage, 163, 115–124.
Poloni, K., & Ferrari, R. (2022). A deep ensemble hippocampal CNN model for brain age estimation applied to Alzheimer's diagnosis. Expert Systems with Applications, 195, 116622.

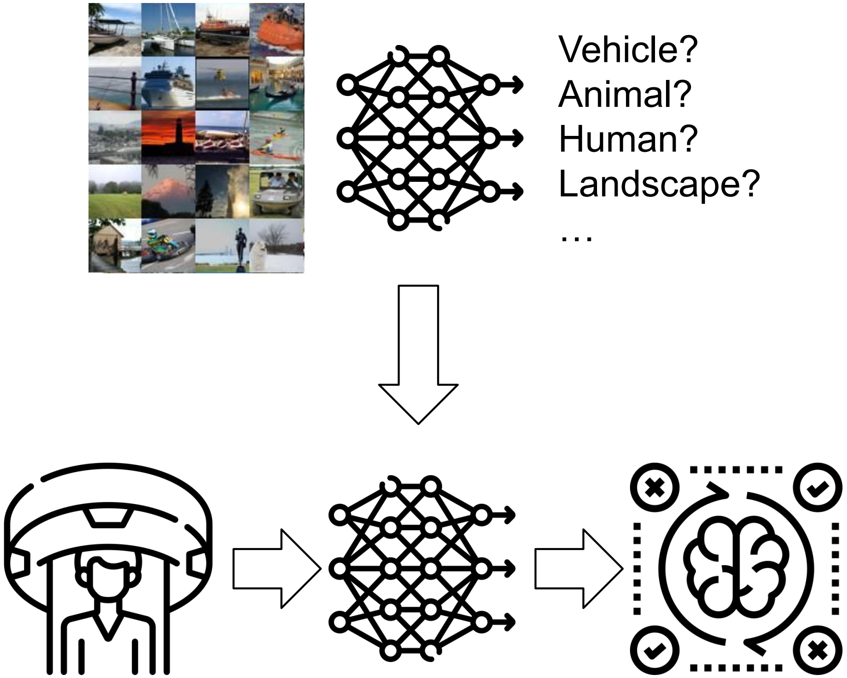
Icons: Flaticon.com

Pre-training

Train the model on another task prior to the task of interest (transfer learning).

- Improves performance by exploiting the knowledge from the other task
- Common on image tasks (ImageNet's pre-trained models)
- Published results of pre-training for brain age (Bashyam et al., 2020)

If pre-training on natural imaging helps brain age prediction, what about pre-training on a brain-related tasks?



Icons: Flaticon.com

In this paper…

- Evaluation of brain-related pretraining
  - BraTS x ImageNet
  - Our approach overcomes all results reported in the literature
- Fully reproducible experiments ⇒ enable direct comparisons
  - Publicly available data
  - Standard split
  - Code available
- Evaluation of the resulting biomarker
  - Δ_{BA} as an indicator of cognitive impairment
  - Better brian age model $\nRightarrow$ Better biomarker

Materials and Methods

Data

ADNI

Alzheimer's Disease Neuroimaging Initiative dataset contains MRIs of healthy patients and patients with mild cognitive impairment (MCI) and Alzheimer's disease (AD).

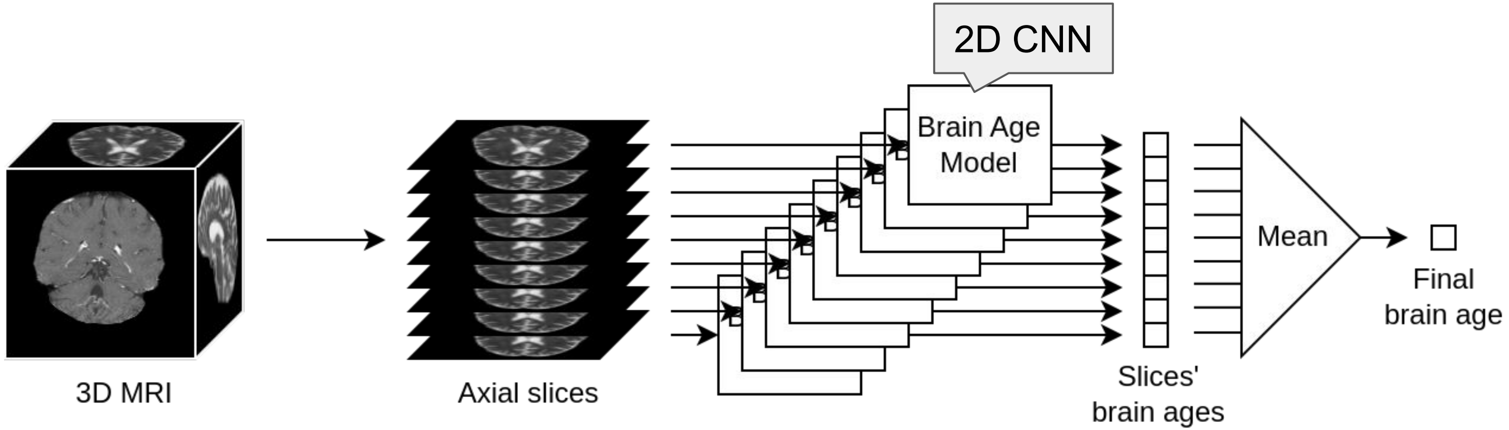
Training: ADNI-2 + ADNI-3 + ADNI-GO

Validation: ADNI-1 Standard Training set

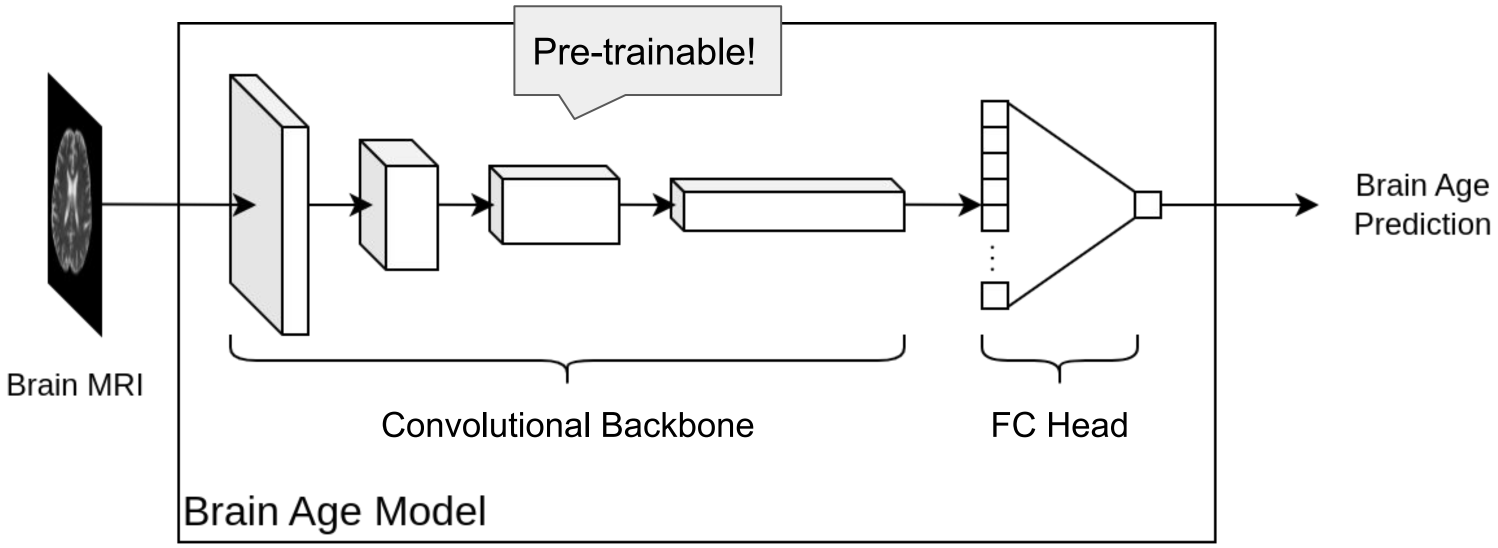
Test: ADNI-1 Standard Test set

	CN	MCI	AD	Age Range (mean)	Sex
ADNI-1	229	401	188	55-91 (75.2)	342 F / 476 M
ADNI-GO	0	142	0	55-88 (71.25)	67 F / 75 M
ADNI-2	150	373	109	55-91 (72.61)	324 F / 307 M
ADNI-3	456	354	64	51-97 (72.78)	475 F / 399 M

Deep learning model



Deep learning model



Pre-training tasks

ImageNet

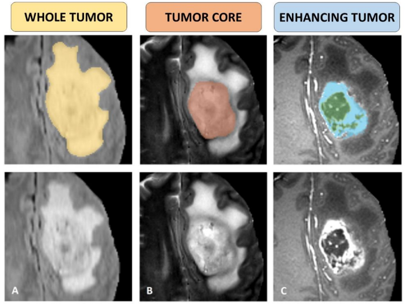
Natural image classification.

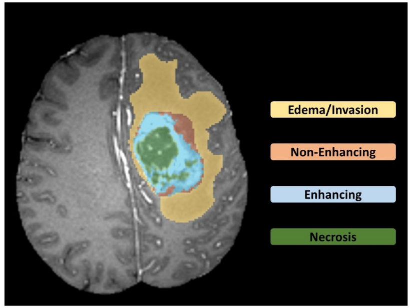
X = Photograph
y = Image label

BraTS'2020

Brain MRIs tumor segmentation.

X = Brain MRI
Y = Tumor segmentation



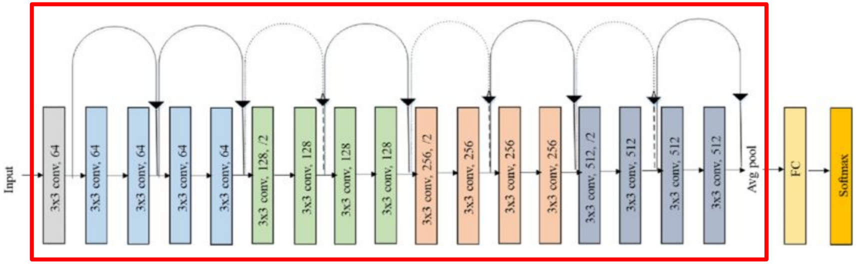


Pre-trained backbones

ImageNet

Deep Learning Model:

- ResNet (publicly available) ⇒ ResNet backbone

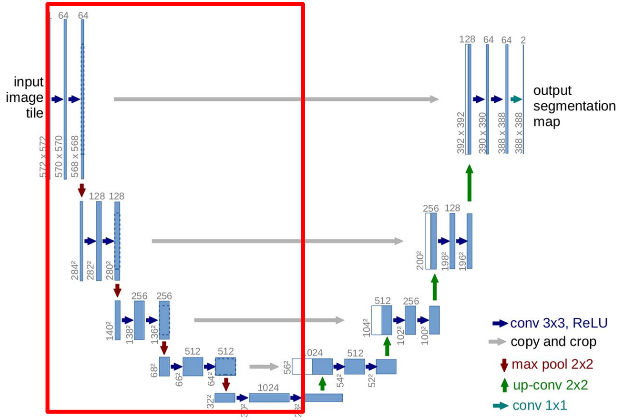


Backbone

BraTS’2020

Deep Learning Model:

- U-Net ⇒ U-Net backbone
- ResU-Net ⇒ ResNet backbone



Backbone

Pre-training

Model	Backbone	ImageNet Pre-training	BraTS Pre-training
1	U-Net		
2	U-Net		X
3	ResNet		
4	ResNet		X
5	ResNet	X	
6	ResNet	X	X

Experiments and Results

Model Training

1. Train on healthy subjects;
2. Hyperparameter tuning using the validation set;
3. Final models trained on training+validation sets;
4. Test set performance through the Mean Absolute Error (MAE) on healthy patients.

Brain Age Results - Healthy Patients

Backbone	BraTS Pre-training	Validation MAE	Test MAE	Test MAE (train+validation models)
U-Net		3.203 ($\sigma = 0.042$)	3.358 ($\sigma = 0.099$)	3.138 ($\sigma = 0.069$)
	X	**3.186** ($\sigma = 0.084$)	**3.284** ($\sigma = 0.071$)	**3.079** ($\sigma = 0.077$)
ResNet		3.474 ($\sigma = 0.117$)	3.556 ($\sigma = 0.258$)	3.223 ($\sigma = 0.071$)
	X	3.361 ($\sigma = 0.079$)	3.413 ($\sigma = 0.123$)	3.141 ($\sigma = 0.030$)
ResNet (ImageNet)		3.509 ($\sigma = 0.086$)	3.638 ($\sigma = 0.160$)	3.210 ($\sigma = 0.047$)
	X	3.452 ($\sigma = 0.086$)	3.411 ($\sigma = 0.066$)	3.238 ($\sigma = 0.022$)

3778 GAMA

Brain Age Results - Healthy Patients

Backbone	BraTS Pre-training	Validation MAE	Test MAE	Test MAE (train+validation models)
U-Net		3.203 ($\sigma = 0.042$)	3.358 ($\sigma = 0.099$)	3.138 ($\sigma = 0.069$)
	X	**3.186** ($\sigma = 0.084$)	**3.284** ($\sigma = 0.071$)	**3.079** ($\sigma = 0.077$)
ResNet		3.474 ($\sigma = 0.117$)	3.556 ($\sigma = 0.258$)	3.223 ($\sigma = 0.071$)
	X	3.361 ($\sigma = 0.079$)	3.413 ($\sigma = 0.123$)	3.141 ($\sigma = 0.030$)
ResNet (ImageNet)		3.509 ($\sigma = 0.086$)	3.638 ($\sigma = 0.160$)	3.210 ($\sigma = 0.047$)
	X	3.452 ($\sigma = 0.086$)	3.411 ($\sigma = 0.066$)	3.238 ($\sigma = 0.022$)

Reduced distribution shift

Brain Age Results - Healthy Patients

Backbone	BraTS Pre-training	Validation MAE	Test MAE	Test MAE (train+validation models)
U-Net		3.203 (σ = 0.042)	3.358 (σ = 0.099)	3.138 (σ = 0.069)
	X	**3.186** (σ	**3.284** (σ = 0.071)	**3.079** (σ = 0.077)
ResNet		3.474 (σ	3.556 (σ = 0.258)	3.223 (σ = 0.071)
	X	3.361 (σ = 0.079)	3.413 (σ = 0.123)	3.141 (σ = 0.030)
ResNet (ImageNet)		3.509 (σ = 0.086)	3.638 (σ = 0.160)	3.210 (σ = 0.047)
	X	3.452 (σ = 0.086)	3.411 (σ = 0.066)	3.238 (σ = 0.022)

BraTS
x
ImageNet

3778 GAMA

Brain Age Results - Healthy Patients

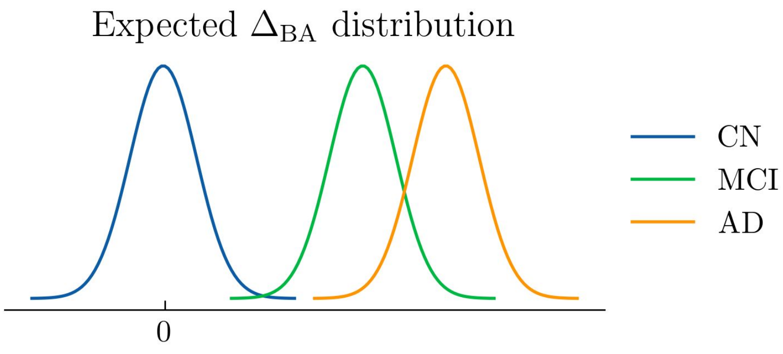
Backbone	BraTS Pre-training	Validation	Test	Test MAE (train+validation models)
U-Net				3.[illegible]9)
	X			**3.079** (σ = 0.077)
ResNet				3.223 (σ = 0.071)
	X			3.141 (σ = 0.030)
ResNet (ImageNet)				3.210 (σ = 0.047)
	X			3.238 (σ = 0.022)

Best results!

Brain Age Model	MAE
Lam et al. (2020)	3.96
Ly et al. (2020)	3.70
More et al. (2023)	6.56
Lee et al. (2022)	3.10
Poloni and Ferrari (2022)	3.66

Evaluation of Δ_{BA}

- Δ_{BA} = BA - CA
- Aging patterns correlate to cognitive impairment levels
  - $\Delta_{BA}(AD) > \Delta_{BA}(MCI) > \Delta_{BA}(CN) \approx 0$
- Pairwise Mann-Whitney U-test between CN, MCI and AD
- Δ_{BA} significantly differentiates between CN and MCI, and between CN and AD



$\Delta_{BA}(MCI) > \Delta_{BA}(CN) \Rightarrow$ Easy

$\Delta_{BA}(AD) > \Delta_{BA}(CN) \Rightarrow$ Easy

$\Delta_{BA}(AD) > \Delta_{BA}(MCI) \Rightarrow$ Hard

Δ_{BA} between MCI and AD

Backbone	BraTS Pre-training	Validation p-value	Test p-value	Test p-value (train+validation models)
U-Net		**0.011** (**0.017**)	**0.025** (0.053)	0.091 (0.107)
	X	**0.011** (**0.035**)	**0.036** (**0.049**)	0.190 (0.213)
ResNet		0.051 (0.066)	**0.046** (0.071)	0.192 (0.245)
	X	0.051 (0.064)	0.051 (0.089)	0.261 (0.298)
ResNet (ImageNet)		**0.044** (0.055)	0.063 (0.112)	0.177 (0.230)
	X	0.095 (0.110)	0.171 (0.225)	0.345 (0.390)

Δ_{BA} between MCI and AD

Backbone	BraTS Pre-training	Validation p-value	Test p-value	Test p-value (train+validation models)
U-Net		**0.011** (**0.017**)	**0.025** (0.053)	0.091 (0.107)
U-Net	X	**0.011** (**0.035**)	**0.036** (**0.049**)	0.190 (0.213)
ResNet		0.051 (0.066)	**0.046** (0.071)	0.192 (0.245)
ResNet	X	0.051 (0.064)	0.051 (0.089)	0.261 (0.298)
ResNet (ImageNet)		**0.044** (0.055)	0.063 (0.112)	0.177 (0.230)
ResNet (ImageNet)	X	0.095 (0.110)	0.171 (0.225)	0.345 (0.390)

Reduced distribution shift

Conclusions

Conclusions

- Brain-related pre-training was consistently a better option;
  - Brain-related pre-training led to overcoming the state-of-the-art
  - Fully-reproducible results, open to future comparisons
- Better brain age deep learning models do not imply in more reliable brain age biomarkers
  - Results for differentiating cognitive impairment levels
  - Training brain age models only on healthy subjects needs further research

Thank you!

gama-ufsc/brain-age

3778 GAMA

Bruno M. Pacheco
UFSC

Victor H. R. de Oliveira
UFSC

Augusto B. F. Antunes
Alliar-NEPIA

Saulo D. S. Pedro
3778 Healthcare

Danilo Silva
UFSC