



OPEN ACCESS

EDITED BY

Jinping Xu,
Chinese Academy of Sciences (CAS), China

REVIEWED BY

Zonglei Zhen,
Beijing Normal University, China
Li Lin,
Jinan University, China

Gang Li,
University of North Carolina at Chapel Hill,
United States

*CORRESPONDENCE

Bang-Bon Koo

✉ bbkoo@bu.edu

Yangming Ou

✉ yangming.ou@childrens.harvard.edu

[†]These authors have contributed equally to this work

RECEIVED 28 June 2023

ACCEPTED 10 October 2023

PUBLISHED 01 November 2023

CITATION

He S, Guan Y, Cheng CH, Moore TL, Luebke JL, Killiany RJ, Rosene DL, Koo B-B and Ou Y (2023) Human-to-monkey transfer learning identifies the frontal white matter as a key determinant for predicting monkey brain age. *Front. Aging Neurosci.* 15:1249415. doi: 10.3389/fnagi.2023.1249415

COPYRIGHT

© 2023 He, Guan, Cheng, Moore, Luebke, Killiany, Rosene, Koo and Ou. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Human-to-monkey transfer learning identifies the frontal white matter as a key determinant for predicting monkey brain age

Sheng He^{1†}, Yi Guan^{2†}, Chia Hsin Cheng², Tara L. Moore², Jennifer I. Luebke², Ronald J. Killiany², Douglas L. Rosene², Bang-Bon Koo^{2*} and Yangming Ou^{2*}

¹Harvard Medical School, Boston Children's Hospital, Boston, MA, United States, ²Department of Anatomy & Neurobiology, Boston University Chobanian and Avedisian School of Medicine, Boston, MA, United States

The application of artificial intelligence (AI) to summarize a whole-brain magnetic resonance image (MRI) into an effective "brain age" metric can provide a holistic, individualized, and objective view of how the brain interacts with various factors (e.g., genetics and lifestyle) during aging. Brain age predictions using deep learning (DL) have been widely used to quantify the developmental status of human brains, but their wider application to serve biomedical purposes is under criticism for requiring large samples and complicated interpretability. Animal models, i.e., rhesus monkeys, have offered a unique lens to understand the human brain - being a species in which aging patterns are similar, for which environmental and lifestyle factors are more readily controlled. However, applying DL methods in animal models suffers from data insufficiency as the availability of animal brain MRIs is limited compared to many thousands of human MRIs. We showed that transfer learning can mitigate the sample size problem, where transferring the pre-trained AI models from 8,859 human brain MRIs improved monkey brain age estimation accuracy and stability. The highest accuracy and stability occurred when transferring the 3D ResNet [mean absolute error (MAE) = 1.83 years] and the 2D global-local transformer (MAE = 1.92 years) models. Our models identified the frontal white matter as the most important feature for monkey brain age predictions, which is consistent with previous histological findings. This first DL-based, anatomically interpretable, and adaptive brain age estimator could broaden the application of AI techniques to various animal or disease samples and widen opportunities for research in non-human primate brains across the lifespan.

KEYWORDS

deep learning models, human brain age estimation, monkey brain age estimation, transfer learning, brain MRIs

Introduction

The advances of artificial intelligence (AI)-based predictor of phenotypic brain aging have generated new tools for mapping normative aging trajectory from large-scale neuroimaging data sets, by summarizing the whole-brain magnetic resonance imaging (MRI) into a holistic, simple, yet effective metric (Cole and Franke, 2017). Deep learning (DL) methods which are more advanced techniques of AI models were shown useful in providing reference information on healthy brain aging trajectories, disease-specific profiles,

and for predicting disease progression. However, several challenges remain before the broader application of DL models to clinical settings. The models often require large sample sizes [at least 2,000 samples (Schulz et al., 2020; Abrol et al., 2021) for human brain MRIs] to minimize biased and confounded predictions, and the biological interpretability of the "brain age" measure is often complicated and not transparent (Moore et al., 2019).

The use of laboratory animals offers a way to run tests and experiments before human trials or when human samples (e.g., healthy brain tissues) are impossible to obtain (Gray and Barnes, 2019) and may help with improving AI models. For example, studies with rhesus monkeys (*Macaca mulatta*) have demonstrated how environmental and lifestyle factors such as diet or exercise impact the brain and cognition across the lifespan, and for understanding the effect of different disease pathophysiology including neurodegenerative disorders, stroke, hypertension, infectious diseases, and affect the brain as it is easier to systematically quantify and manipulate these factors in monkeys than in humans (Mellus, 1905, 1907; Roth et al., 2004; Moss et al., 2007; Lacreuse and Herndon, 2009; Mattison and Vaughan, 2017; Koo et al., 2018; Kuchan et al., 2020). However, applying AI predictors to animal models of aging (i.e., non-human primates) is rare and challenging because of the scarcity of brain MRIs (Franke et al., 2010). Contrarily, the number of publicly available human brain MRIs (>8,000) is larger than the minimum request numbers (2,000) (Schulz et al., 2020; Abrol et al., 2021) to train a DL model, yielding a good accuracy of brain age estimation on human MRIs (Feng et al., 2020; He et al., 2021a; Peng et al., 2021; Lee et al., 2022). To mitigate this issue, we implemented the transfer learning strategy, which uses knowledge learned from large-sample benchmark tasks (i.e., brain age estimation in humans) to improve prediction accuracy in small-sample target tasks (i.e., brain age estimation in animal models) (Kermany et al., 2018) in our recently developed pre-trained 3D AI models with improved transferability in 3D images (Feng et al., 2020; He et al., 2021b, 2022).

Here, we applied a DL approach to MRIs of non-human primates (NHP), which is a well-established animal model for studying cognitive aging with high genetic homology (>92%), similar cognitive and sensory abilities to those of humans, and longer lifespan (approximately 1/3 to that of human) than more distant laboratory animals such as rodents (Roth et al., 2004; Gray and Barnes, 2019). We hypothesized that five state-of-the-art 2D and 3D AI models pre-trained on 8,859 human brain MRIs could reduce the errors of AI brain age estimation in 29–290 monkey brain MRIs compared to models without transfer learning and that the key features selected by the AI models are consistent with previous biological investigations. Saliency maps generated from our models are used to examine brain image features that are most important for brain age estimation and are compared with those from our previous human brain age models. We demonstrated that 3D DL models with transfer learning are highly accurate, stable, and biologically meaningful in analyzing relatively small samples of NHPs, broadening the opportunities for applying cutting-edge AI paradigms to various laboratory animals.

Methods

Monkey subjects and MRI acquisition

In total, 290 T1-weighted brain MRIs from 274 rhesus monkeys were used in this study. They include 172 male and 102 female monkeys, 1–30 years of age at the time of brain MRI acquisition, 126 from the Boston University (BU) NHP database, and 148 from the PRIMatE Data Exchange (PRIME-DE) database (Milham et al., 2018).

The monkeys at BU were obtained from the Emory National Primate Research Center, the New England National Primate Research Center, and World Wide Primates for other funded projects over 3 decades. Prior to entering the study, all monkeys received medical examinations that included serum chemistry and health screenings to ensure that monkeys did not have a history of malnutrition, diabetes, chronic illness, or any neurological diseases. All monkeys were housed in the Laboratory Animal Science Center of Boston University Medical Campus, an AAALAC-accredited facility. Experiments were conducted in accordance with the Guide for the Care and Use of Laboratory Animals from the National Institute of Health's Office of Laboratory Animal Welfare and were approved by the Boston University Institutional Animal Care and Use Committee (IACUC).

MRIs on BU monkeys were obtained as previously described (Koo et al., 2012, 2013). In brief, the monkeys were anesthetized with an initial dose of ketamine (10 mg/kg, IM) followed by administration of propofol IV (25 mg/kg, IV) using a syringe pump to maintain anesthesia throughout the session or with repeated doses of ketamine (10 mg/kg, IV) and valium (1 mg/kg, IV). In the MRI scanner, monkeys were held in a stereotactic MRI-compatible head holder designed to fit within the 8-channel phase array head coil in a fixed position preventing any inadvertent movement from respiration. Respiration and oxygenation levels were monitored, and body temperature was maintained throughout the imaging session. MRI scans were performed on an Intera (for 16 scans) or Achieva 3T whole-body MRI scanner (Philips Healthcare, Best, The Netherlands). T1-weighted images were acquired using a T1-weighted 3D-turbo field echo (TFE) sequence: repetition time (TR)/echo time (TE) = 7 ms/3 ms, flip angle = 8°, NEX = 6, inter shot delay = 2,800 ms, TFE = factor 200, voxel size = 0.6 mm isotropic, and sagittal plane acquisition.

The PRIME-DE database combines NHP neuroimaging data sets across 22 sites. Details have been described in a 2018 article (Milham et al., 2018). Of them, this study included data from 15 sites that contain *in vivo* structural MRI data for rhesus monkeys, details are described in [Supplementary Table S2](#).

Among the 290 T1-weighted MRI scans, 260 monkeys had only one scan session, 12 had two MRI sessions, and 2 had three MRI sessions (at least 6 months apart between each scan). These longitudinal scans are treated as independent data for brain age estimation. Multiple scans from human subjects were also used in human brain age prediction studies with dozens or lower hundreds of subjects, but our study included a much lower percentage (6%) of human subjects with two scans.

Human subjects and MRI acquisition

We collected 8,859 T1-weighted human brain MRIs (0–97 years of age, 45.2% male) from 13 public data sets, including ABIDE-I (Di Martino et al., 2014) (567 scans, age 6–57), BeijingEN¹ (180 scans, age 17–28), CamCAN (Taylor et al., 2017) (442 scans, age 18–68), DLBS (Park et al., 2012) (315 scans, age 20–89), DutchDonders² (171 scans, age 18–32), IXI³ (556 scans, age 20–86), Narratives (Nastase et al., 2021) (335 scans, age 18–53), NIH-PD (Evans et al., 2006) (1,210 scans, age 0–23), KI-RS_Enhanced (Nooner et al., 2012) (2,032 scans, age 6–85), OASIS-3 (LaMontagne et al., 2018) (1,838 scans, age 42–97), SALD (Wei et al., 2018) (488 scans, age 19–80), Wayne (Daugherty and Raz, 2017) (606 scans, age 18–91), and YaleHires (Finn et al., 2015) (119 scans, age 18–59). These T1-weighted brain MRIs were all acquired at 1 isotropic mm voxel size or slightly higher resolution.

Pre-processing of human and monkey MRI

Following the human brain age estimation (He et al., 2021a,b, 2022), we performed identical minimal pre-processing to both monkey and human brain MRIs.

For human brain T1-weighted brain MRI, the pre-processing included N4 bias correction (Tustison et al., 2010), field of view normalization (Ou et al., 2018), multi-atlas skull stripping (MASS) (Doshi et al., 2013; Ou et al., 2015), and a deformable registration of the skull-stripped image into the SRI24 atlas space (Rohlfing et al., 2010) by the deformable registration via attribute matching and mutual-saliency weighting (DRAMMS) algorithm (Ou et al., 2011, 2014). The deformable registration split the T1-weighted human brain MRI into two channels, both in the atlas space: a registered intensity image channel for the **contrast** information and a tissue density map known as RAVENS map (regional analysis of volumes examined in normalized space) containing **morphometry** information (Davatzikos et al., 2001). See Figure 1B for these two channels for exemplar subjects. Our previous study demonstrated that splitting a single T1-weighted MRI into these two channels led to a higher brain age estimation accuracy (He et al., 2021b, 2022). Therefore, this study used both channels for age estimation.

Monkey T1-weighted brain MRIs underwent the AFNI pipeline (animalwarper tool) to generate skull-stripped images (Saad et al., 2009; Jung et al., 2021), by non-rigidly transforming NIMH Macaque Template (NMT v2) into each monkey's brain MRI space. The skull-stripped monkey brain MRI in the native space then went through DRAMMS deformable registration to be spatially normalized into the NMT atlas space. This split the monkey brain MRI into a registered intensity image (i.e., the **contrast** channel) and the RAVENS map (i.e., the **morphometry** channel) similar to those in human brain MRIs. The right panel of Figure 1B shows exemplar monkey subjects for their images and the two channels we used for age estimation.

After cropping unused background in the pre-processed images, the final dimensions were $100 \times 128 \times 156$ for monkey images in the NMT atlas space and $120 \times 130 \times 170$ for brain images in the SRI24 atlas space. For the purpose of transfer learning, further resizing was done to linearly scale (not transforming) the human brain images to have the same size as monkey brains.

To train the deep learning model, the intensity map and RAVENS map are concatenated as one input image with two channels, with the size of $2 \times 100 \times 128 \times 156$.

Five AI models

Five different DL AI models were evaluated for monkey brain age estimation. They included the general 2D/3D ResNet models (He et al., 2016), which were not specifically designed for but have been previously validated for brain age estimation (He et al., 2021a, 2022). We also included three models that were specifically designed for brain age estimation: 3D simple fully convolutional network (SFCN) (Peng et al., 2021), 2D global-local transformer (GL-Transformer) (He et al., 2021a), and 3D deep relation transformer (Relation-Transformer) (He et al., 2022). For the 2D models, we consider the 3D MRIs as a stack of 2D slices with different channels.

We chose these five AI models for their representativeness. On one note, these five models represent 2D and 3D typical deep learning architectures. The 2D ResNet and GL-Transformer are 2D deep learning architectures. The 3D ResNet, SFCN, and Relation-Transformer are 3D deep learning architectures. Moreover, these five models represent both the convolutional neural network (CNN) and transformer-based algorithms. The 2D/3D ResNet and SFCN are CNN-based, whereas the GL-Transformer and Relation-Transformer are based on transformers, which are recently popular deep learning architectures based on attention mechanisms (Vaswani et al., 2017; He et al., 2022). Loosely speaking, CNN uses voxel/pixel arrays, whereas transformer models split the images into visual tokens. They are two main screams of deep learning models for image computing.

The algorithms behind these models are briefly described below. The ResNet model contains residual block with skip connection to reuse the feature feeding into the residual block. The SFCN is a short version of the VGGnet (Simonyan and Zisserman, 2014), with only five convolutional layers to extract deep features. Global-Local transformer contains a transformer (Vaswani et al., 2017) block to optimal fuse global context and local information from patches to explore more detailed brain age information on brain MRIs. Relation-Transformer (He et al., 2022) uses the attention mechanism to learn four relations, such as the cumulative, relative, maximal, and minimal relations from a pair of input images.

Training from scratch in monkey brain MRIs

The models trained from scratch on monkey MRIs used the same training configurations as in human MRIs: batch size was 8, the initial learning rate of the Adam optimizer (Kingma and Ba,

1 https://fcon_1000.projects.nitrc.org/indi/PRIMEcollections.html

2 <https://data.donders.ru.nl/?1>

3 <https://brain-development.org/ixi-dataset>

2014) was 0.0001 and was reduced to half at every 50 epochs with a total of 200 training epochs.

Transferring and refining pre-trained models from human to monkey MRIs

Pre-training was done in human brain MRIs. All five models were trained with the mean absolute error (MAE) as the loss function, the same as in other human brain age estimation studies (Bashyam et al., 2020; Feng et al., 2020; He et al., 2021a,b, 2022; Peng et al., 2021). Parameters on models were optimized using the Adam optimizer (Kingma and Ba, 2014), with an initial learning rate of 0.001, which was reduced to half at every 20 epochs with a total of 80 training epochs. The batch size was set to 16. All other parameters were set to default provided by the PyTorch platform, similar to existing studies (He et al., 2021a,b, 2022).

Transfer learning was done using the pre-trained models from human brain MRIs as an initialization. The refinement was done in two stages. In Stage 1, we refined the last fully connected layer from scratch for monkey age estimation. We fixed the parameters on convolutional layers for feature extraction. This was done with 100 epochs on monkey brain MRIs with a learning rate of 0.001. After that, Stage 2 fine-tuned the parameters in all layers of the deep learning AI models with a learning rate of 0.0005, which was quartered at the 100th epoch. All models were trained with 200 epochs. Although the age ranges of human and monkey are not the same, age matching is not necessary for transfer learning since the last fully connected layer is randomly initialized for monkey and fine-tuned in transfer learning.

Ranking AI strategies by accuracy and stability

Ten AI strategies (five AI models, each with or without transfer learning) were compared.

Accuracy. We used a 10-fold cross-validation to measure the accuracy. Given N monkey brain MRIs, we randomly split them into 10 equal numbers of non-overlapping samples. Each time, one fold was selected for testing, and the rest were used for training. One metric for accuracy was the mean absolute error (MAE), $MAE = \frac{1}{N} \sum_i^N |\hat{y}_i - y_i|$, where \hat{y} is the AI-estimated age, while y is the chronological age. Smaller MAE corresponds to higher accuracy. The second accuracy metric we used was Spearman's correlation coefficient between the chronological and AI-estimated monkey brain ages. Spearman's correlation coefficient ranges from 0 (lowest accuracy) to 1 (highest accuracy). In addition to reporting the overall accuracy across the lifespan in the main text, we also computed the accuracy metrics in different age groups in [Supplementary material](#). A monkey's lifespan capacity is approximately 0–35 years, with a rough ratio of 1:3 to human lifespan ages (Tigges et al., 1988). Therefore, we defined and compared three age groups within the current cohort, as previously defined (Simmons, 2016): young monkeys (age ≤ 5 , $n = 88$), middle-aged (5 < age < 20, $n = 157$), aged (age ≥ 20 , $n = 45$).

Stability. Each AI strategy was applied to varying sample sizes in our monkey brain MRI dataset. A strategy is more stable if the accuracy dropped the least going from using all 290 monkey brain MRIs to using only 10, 20, 30, ..., and 90% of them. Training samples were randomly selected on each fold (10-fold cross-validation). For example, when sampling 20% for training, we randomly selected 20% samples from the 9-fold for training and the rest fold for testing and repeat 10 times.

Ranking by accuracy and stability. We ranked the 10 AI strategies from 1 (least preferable) to 10 (most preferable) in each of the four accuracy metrics and three stability metrics. The total rank score is a summation of those 7 scores ranging from 7 (least preferable) to 70 (most preferable). The four accuracy metrics were MAE ranked from highest (rank score 1) to the lowest (rank score 10) when using 100% of the monkey brain MRIs; Spearman's ρ ranked from the lowest (rank score 1) to the highest (rank score 10) when using 100% of the monkey brain MRIs; MAE ranked from highest (rank score 1) to the lowest (rank score 10) when using 30% of the monkey brain MRIs; and Spearman's ρ ranked from the lowest (rank score 1) to the highest (rank score 10) when using 30% of the monkey brain MRIs. The three stability metrics were the increase of MAE ranked from the highest (rank score 1) to the lowest (rank score 10) from using 100% to 30% of the monkey brain MRIs, the increase of MAE ranked from the highest (rank score 1) to the lowest (rank score 10) from using 100% to 20% of the monkey brain MRIs, and the increase of MAE ranked from the highest (rank score 1) to the lowest (rank score 10) from using 100% to 10% of the monkey brain MRIs. The AI strategies with the highest ranking scores (ideally close to 70) were recommended among the 10 strategies.

Interpretation

For further interpretation of the brain age models, saliency maps were generated to show the most informative voxels from MRI scans selected by the AI model for brain age estimation. In brief, the integrated gradients were used (Sundararajan et al., 2017) to visualize the relationship between input images and predictions. The baseline was the image with zero values and a sequence of images was generated by a linear interpolation between the baseline and input image, yielding a sequence of output prediction and a sequence of gradient maps on input images.

The saliency maps shown in [Figure 4](#) were computed as the average saliency maps across all subjects in corresponding age groups (<5 and ≥ 5 years for monkeys and <15 and ≥ 15 for humans). To quantify the tissue-specific and regional contribution to brain age estimations, we applied lobular segmentations for gray and white matter regions, as well as the inclusion of the subcortical structure as a separate region. Mean saliency values normalized by whole sample minimum and maximum saliency values for each age group were quantified for tissue-based regional segmentation and plotted to visualize the trajectories of mean saliency.

Topographical saliency maps are prepared for two age groups for each channel used in AI models (intensity/contrast & RAVENS/morphometry) and captured from three different axes.

Similar procedures were carried out to generate human saliency maps and quantified saliency trajectories in matching age groups (under 15 years old and 15 or older) (Figure 4).

Results

Setup of transfer learning

The benchmark task was human brain age estimation (Figure 1A, left panel). It was based on using 8,859 T1-weighted human brain MRIs 0-97 years of age. The target task was monkey brain age estimation (Figure 1A, right panel), using T1-weighted monkey brain MRI from 290 monkeys with a lifespan of 1–30 years of age. We applied minimum pre-processing to split each bias-corrected and skull-stripped T1-weighted MRI into a 3D intensity image (for contrast information) and a 3D RAVENS (regional analysis of volumes examined in normalized space) map (for morphometry information) in the human or monkey atlas spaces (Seidlitz et al., 2018; He et al., 2021a,b, 2022; Jung et al., 2021), as shown in Figure 1B. Age estimation in this study, with or without transfer learning, was based on concatenating and using these two channels of images.

The five AI models we used in the study included two 2D AI models [the classic 2D ResNet He et al., 2016, our 2D global-local transformer (GL-Transformer) He et al., 2021a] and three 3D AI models (the classic 3D ResNet He et al., 2016, a recently-developed 3D SFCN Peng et al., 2021, and our 3D Relation Transformer He et al., 2022). These five models also include three convolutional neural network models (2D/3D ResNet, 3D SFCN) and two transformer models, that are the two main branches of deep learning AI architectures.

Transfer learning improved accuracy, especially in very small sample sizes

We performed transfer learning by first fixing the parameters from the pre-trained benchmark (human) model on convolutional layers, to refine the last fully-connected layer from scratch, and then performed refinement of all the parameters during target (monkey) model training. Figure 2 shows that transferring AI models from human to monkey brain MRIs (dashed curves) led to reduced age estimation errors (y-axis) for all five AI models compared to training directly in monkey data from scratch (solid curves). The effect varied, though, by sample size in the target task (x-axis) and across models.

The effect of transfer learning on AI accuracy was more pronounced when we used only <150, especially <50 ($p < 0.05$ for t-Test), randomly-sampled monkey MRIs. The benefit of transfer learning became smaller or even vanished for some AI models when using more than 200 randomly-sampled monkey MRIs. This general trend is consistent with all five models we tested although the specific cutoff sample size varied for different underlying AI models, as described below.

Effects of transfer varied by AI models

We evaluated the performances of the AI models using mean absolute error (MAE) and Spearman's correlation coefficient (ρ) between the actual and the AI-estimated ages. A lower MAE and a higher ρ usually correspond to higher accuracy. The lowest MAE in the analysis of the 290 monkey brain MRIs occurred when using 3D ResNet with transfer learning (MAE=1.83 \pm 0.39 years). The highest Pearson correlation ρ occurred when using the 3D SFCN with transfer learning ($\rho = 0.921$).

Two 3D AI models (3D SFCN and 3D relation transformer) benefited more from transfer learning than the two 2D AI models (2D ResNet and 2D GL-Transformer). In general, 3D AI models often required larger sample sizes than 2D models (Schulz et al., 2020).

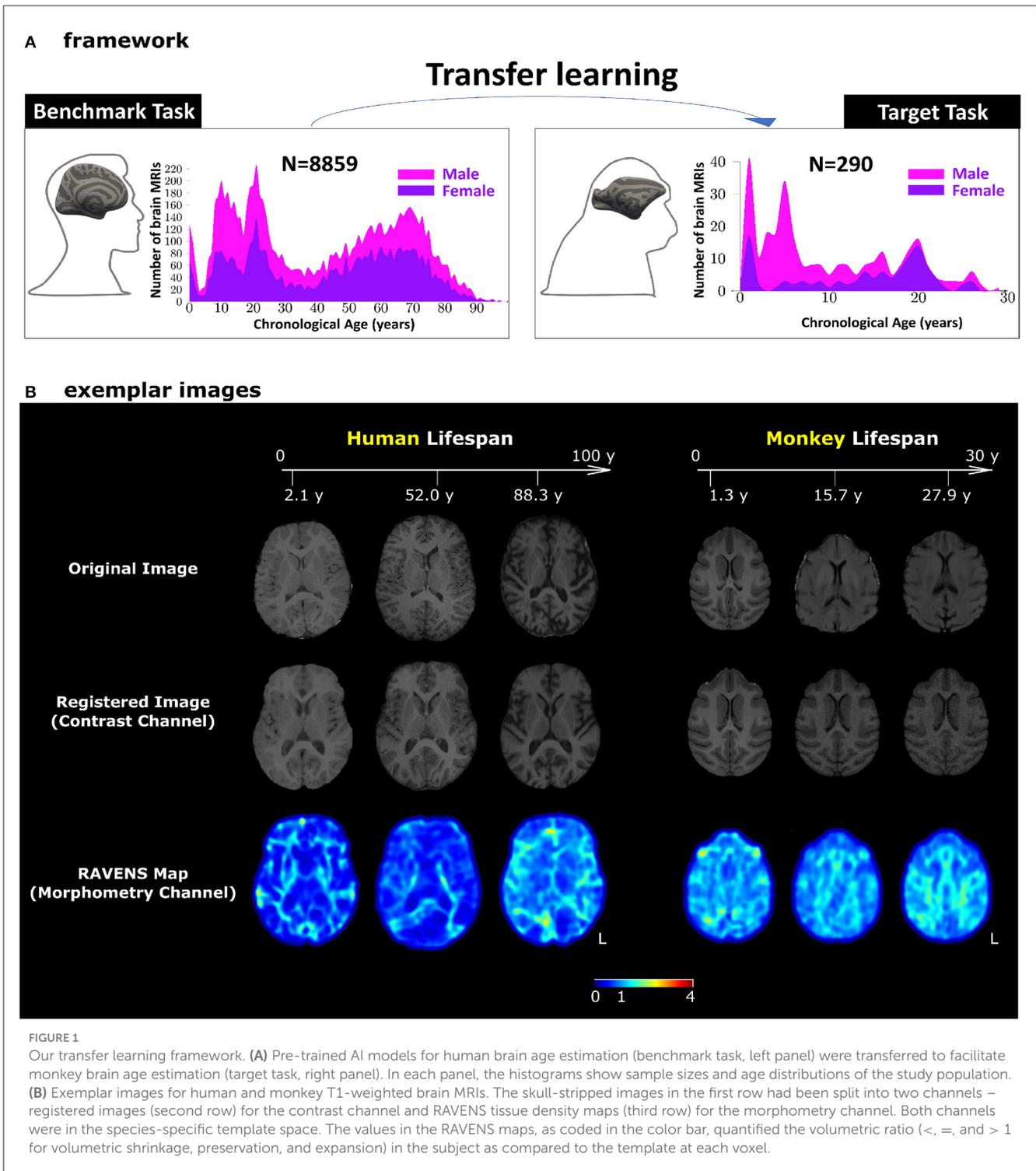
The AI model that benefited the most was the model with the largest errors (3D SFCN). AI models that had a relatively smaller error (<4 years), even with just 30 randomly-sampled monkey brain MRIs, benefited less from transfer learning.

Of special note is the 3D relation transformer, which, given N input images, used N^2 pairs of images as training samples. The 3D relation transformer estimated the relationship (sum, difference, bigger/smaller than) on every pair of input images (He et al., 2022). Therefore, transfer learning reduced errors compared to without transfer learning for the 3D relation transformer, with 220 or fewer input images (<48,400 input image pairs). The MAE of transfer learning became larger than the MAE without transfer learning with more than 220 input monkey brain MRIs (more than 48,400 input MRI pairs).

In general, the comparison across the underlying AI models showed that those AI models that needed more input MRIs (e.g., 3D SFCN and 3D ResNet) would benefit more from transfer learning.

Transfer learning also improved stability

We next ranked the 10 strategies (five AI models; with or without transfer learning) based on their accuracy and stability (Figure 3). The rank score ranged from 1 (least recommended) to 10 (most recommended) in four accuracy metrics and three stability metrics (with each row being a metric). Accuracy was measured as the mean absolute errors (MAE, smaller is better) and Spearman's correlation (higher is better) at 30% and 100% of the samples between AI-estimated and actual chronological monkey ages. Stability was measured by the increase of MAE (smaller is better) from when using 100% of the samples to when using 10%, 20%, and 30% of the samples. Overall, the 2D GL-Transformer with transfer learning and the 3D ResNet with transfer learning received the highest total rank scores, both scored 56 out of 70. They were followed by 2D ResNet with transfer learning, with a total rank score of 51. These recommended choices were all with transfer learning.



Neuroanatomical interpretations

After examining the accuracy and stability of different AI models, we visualized the key MRI regions and features (from the contrast and morphometry channels) that contributed the most to age estimation in the best-performing model, 3D ResNet with transfer learning, which had the highest rank scores

in [Figure 3](#) and the lowest MAE when all 290 MRIs were used in cross-validation. We generated saliency maps, which are topographical representations of feature importance, showing the most informative voxels selected by the AI model for brain age estimation ([Figure 4](#)). Three main observations are noted when comparing saliency maps generated from the pre-trained human model:

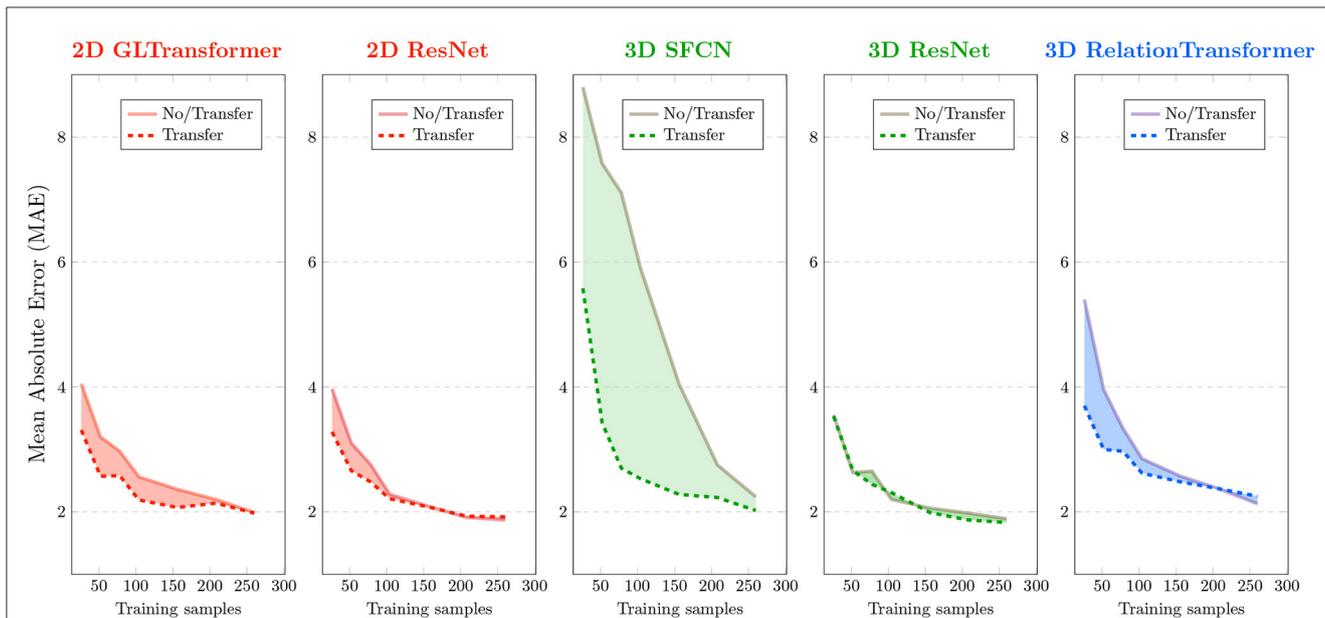


FIGURE 2
Effects of transfer learning in monkey brain age estimation. Each panel corresponds to an AI model that was used in this study. Brain age errors (mean absolute errors or MAEs) in 10-fold cross-validation of monkey brains are plotted on the y-axis, and various sample sizes used for AI are plotted on the x-axis. Solid curves are errors without transfer (i.e., training from monkey brain MRIs from scratch). Dashed curves are errors with transfer (from human to monkey). Red colors are used for 2D AI models, green for 3D AI models, and blue for a special 3D AI model that uses N^2 input image pairs given N input images.

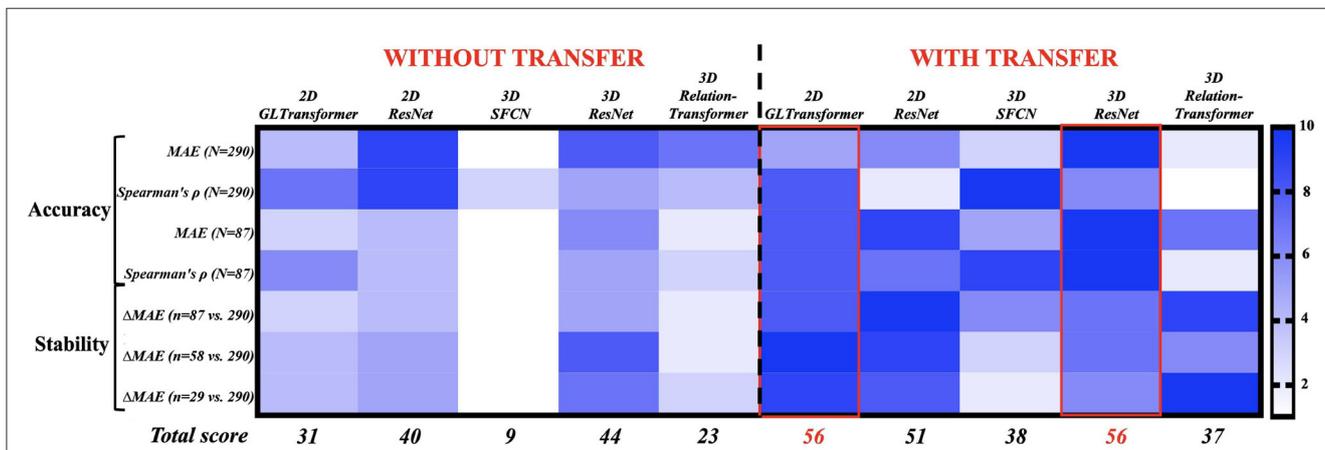


FIGURE 3
Ranking 10 AI strategies by accuracy (Rows 1–4) and stability (Rows 5–7). Each row assigns 10 AI strategies a rank score of 1 (least preferable, light blue) to 10 (most preferable, dark blue). The AI strategies with the highest total rank scores (**bottom row**) are the most accurate and stable choices for monkey brain age estimation. The exact values in those 10 accuracy and stability metrics can be found in Figure 2 and Supplementary Table S1.

- For rhesus monkeys 5 years of age or younger, the salient regions selected by the AI model were more global throughout the brain. This can be seen in the right panels in Figure 4(a). Quantitative results in 1-5 years in Figure 4(b) also confirmed this observation.
- For rhesus monkeys 5 years of age and older, the aging effect became more localized in specific regions or tissue types. Compared to monkeys younger than 5 years of age, AI estimated ages for monkeys 5 years and older was higher/more reliant in the frontal white matter [pointed out by red arrows in Figure 4 (a, left)], temporal white matter (orange arrows

and subcortical structures (yellow arrows, including basal ganglia, diencephalon, brain stem, and cerebellum), followed distantly by the parietal white matter. The age information came from both the contrast (intensity) and the morphometry (RAVENS) channels. This can be visualized under monkey panels in Figure 4(a), as well as quantified in Figure 4(b) for 5 years and up.

- A similar trend is observed in humans under the age of 15 where salient regions are globally distributed and mean saliency values are comparable across regions [Figure 4 (left)], whereas in humans 15 years old and older (matching to 5 years

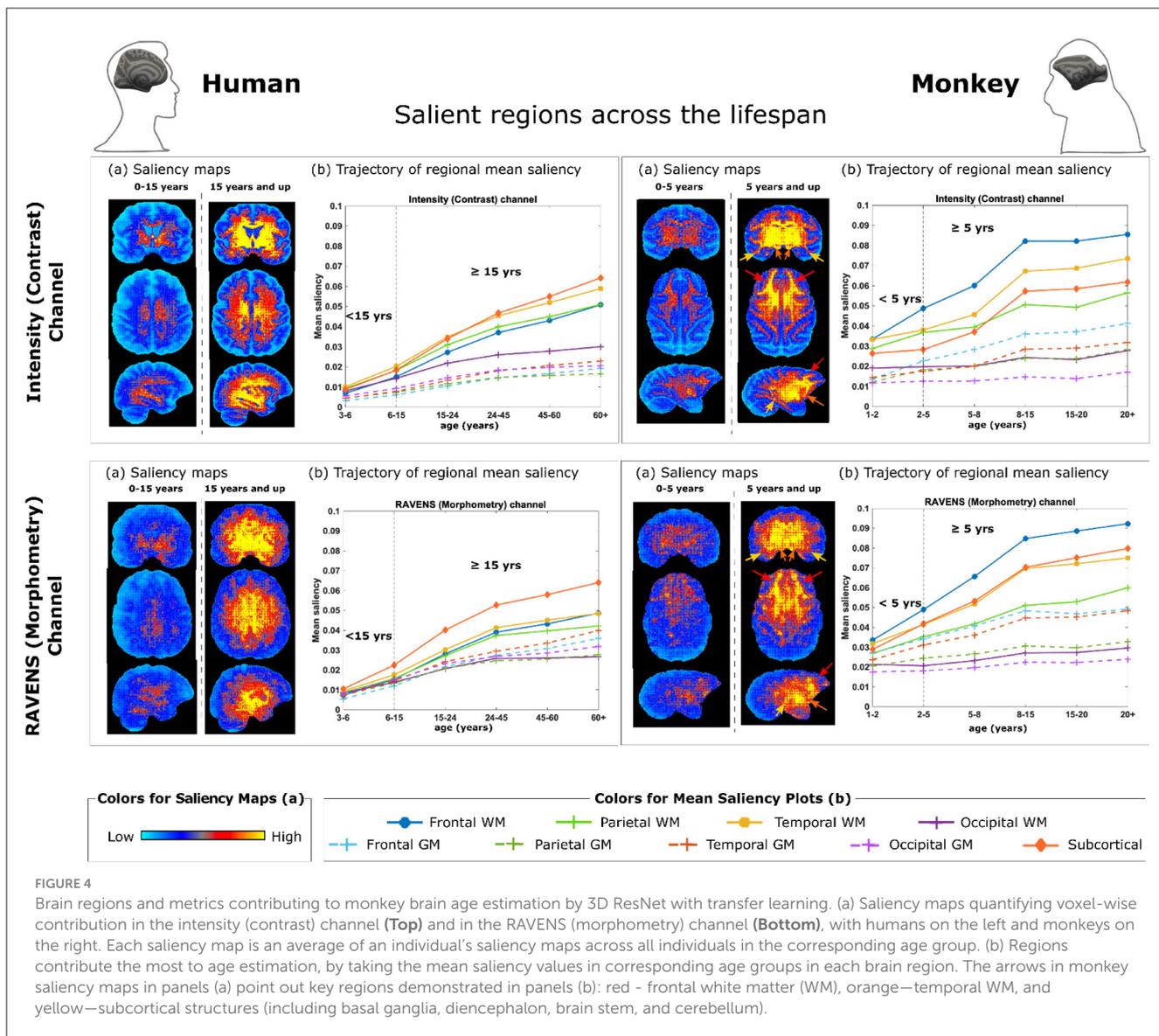


FIGURE 4
Brain regions and metrics contributing to monkey brain age estimation by 3D ResNet with transfer learning. (a) Saliency maps quantifying voxel-wise contribution in the intensity (contrast) channel (**Top**) and in the RAVENS (morphometry) channel (**Bottom**), with humans on the left and monkeys on the right. Each saliency map is an average of an individual's saliency maps across all individuals in the corresponding age group. (b) Regions contribute the most to age estimation, by taking the mean saliency values in corresponding age groups in each brain region. The arrows in monkey saliency maps in panels (a) point out key regions demonstrated in panels (b): red - frontal white matter (WM), orange—temporal WM, and yellow—subcortical structures (including basal ganglia, diencephalon, brain stem, and cerebellum).

old and older in monkeys), subcortical structures were the most prominent, followed by temporal, frontal, and parietal white matters [Figure 4 (left)].

Discussion

Following the advances in medicine and nutrition, the longer life expectancy worldwide led to a growing elderly population and an increased healthcare burden on neurodegenerative diseases. Rhesus monkeys are particularly important in mapping longitudinal normal and pathological aging trajectories because of their high genetic and physiological similarity to humans, which effectively helped to bridge the critical gap between rodent experiments and human trials for developing potential therapeutics (Gray and Barnes, 2019). However, the limited sample sizes of monkey studies present a major obstacle to the reproducibility and wider application of this model in biomedical

research. As a potential solution, we built and improved AI models with transfer learning from a large-sample (human) benchmark task to small-sample (monkey) target tasks (Kermay et al., 2018) to study brain aging using human and monkey MRI scans. While existing studies primarily focused on transferring pre-trained 2D AI models to 2D images or 2D slices within 3D images (Kather et al., 2019; Bashyam et al., 2020; Hollon et al., 2020; Ström et al., 2020), recent advances in 3D AI models (He et al., 2016; Feng et al., 2020), including our models (He et al., 2021a,b, 2022), which have been pre-trained on thousands to tens of thousands of 3D images (Figure 1), have created the opportunity to transfer pre-trained 3D AI models to 3D medical images. In the current study, we demonstrated that (1) 3D ResNet with transfer learning achieved the highest ranking in accuracy and stability out of 10 strategies (five AI models, with or without transfer), as well as the lowest errors when all 290 monkey MRIs were used (Figure 3), (2) transfer learning improved the accuracy and stability for monkey brain age estimation for all models (see

the total ranking scores before versus after transfer in Figure 3), and the effect was more pronounced in very small sample sizes (especially for sample size ranges between 150 and 50, Figure 2), (3) the saliency maps generated by our best-performing model identified important brain age features in the frontal white matter, which is consistent with previous monkey studies (Wisico et al., 2008; Peters and Kemper, 2012). To the best of our knowledge, this is the first study that applied the DL framework combined with the transfer learning technique for brain age prediction using brain MRI from rhesus monkeys. Of note, there was one study that trained a non-deep-learning (DL) algorithm (relevance vector regression) on the tissue segmentations from 29 T1-weighted baboon brain MRIs for age estimation (Franke et al., 2017).

The current study sought to mitigate the small sample issue in building DL brain age models by applying transfer learning from pre-trained models from human to monkey. Indeed, we observed improved age prediction accuracy (indicated by the reduction of MAE in Figure 2) in all five models at various degrees. Interestingly, while we initially expected that more complex AI models [complexity is determined by the number of parameters and the floating point operations per second (flops)] may need more samples and thus would benefit more from transfer learning, our results did not conform with such an expectation. For instance, 3D SFCN [2.81 M(ega) parameters and 26.66 G(iga) flops] and 3D relation transformer (2.91 M parameters and 6.46 G flops) had fewer parameters than other AI models we used in this study but benefited the most from transfer learning. In contrast, 2D GL-Transformer (19.68 M parameters and 16.75 G flops), 2D ResNet (11.25 M parameters and 3.6 G flops), and 3 D ResNet (31.65 M parameters and 34.71 G flops)—three AI models with larger numbers of parameters (suggesting higher complexity)—actually benefited less from transfer learning. Instead of complexity as the major factor for determining the success of transfer learning, our results showed that AI models with larger errors at very small sample sizes (50 or under) actually benefited more from transfer learning. This suggests that, in addition to model complexity, sample sizes might be a key factor to consider when selecting the appropriate AI model for different tasks.

After evaluating the performances (based on accuracy and stability) of our DL models, we showed that the two top-ranked models (see Figure 3) are 3D ResNet with transfer learning (MAE = 1.83 years) and 2D GL-Transformer with transfer learning (MAE = 1.99 years). For brain age models, improved prediction accuracy may suggest a potentially higher sensitivity to identify subtle deviations from normal brain aging curves. In human, AI-estimated ages have identified accelerated aging for cognitive impairment (Liem et al., 2017; Poddar et al., 2019), traumatic brain injuries (Cole et al., 2015), schizophrenia (Cole et al., 2018), Alzheimer's disease (Franke et al., 2010), diabetes (Franke et al., 2013; Guan et al., 2022), smoking or alcohol use (Guggenmos et al., 2017; Ning et al., 2020), and early signs of future psychosis (Chung et al., 2018). Similarly, AI quantified the delayed human aging in long-term meditation practice (Luders et al., 2016), music-making (Rogenmoser et al., 2018), and higher education (Steffener et al., 2016). In these studies, the AI age estimation errors on

human brain MRIs are around 3–6 years (Goyal et al., 2019; Bashyam et al., 2020; Smith et al., 2020; He et al., 2021b). With improvements in prediction accuracy and stability from transfer learning for monkey brain age estimation, it becomes possible to better quantify how diseases, genetics, environment, lifestyle, and other epigenetic factors have shifted monkey brains from normal aging curves (Jeon et al., 2018; Kuchan et al., 2020; Souder et al., 2021). This is especially useful because many of these factors can be well controlled in the laboratories for monkeys and other animal models, providing neuroscience testbeds that are otherwise difficult to obtain in human brains (Moss et al., 2007).

The current study also sought to improve the interpretability of the DL models and pinpoint the most important MRI features used by the best-performing models for age predictions. To achieve this, we generated saliency maps and showed that the highest-ranked AI strategy was able to learn monkey brain ages from both morphometry (RAVENS map) and the contrast (intensity image) information (Figure 4). The aging information was more widespread throughout the whole brain in monkeys under the age of 5 years but became progressively more localized for monkeys 5 years and more, which was also observed in the age-matched groups of humans (Figure 4). Interestingly, AI predicted monkey brain ages mostly based on MRI voxels located in the frontal white matter (WM) regions, followed by parietal WM and subcortical (mostly GM) regions (Figure 4). This agrees with previously established histological findings on monkey brain development across the lifespan, where WM degeneration was prevalent in aged animals (Peters, 2009; Peters and Kemper, 2012; Chen et al., 2013; Kubicki et al., 2019). These have mostly focused on the forebrain regions, showing mild dendritic and synaptic loss, and prominent WM degeneration in the dorsal prefrontal cortex, area 46, anterior cingulate cortex, and hippocampus during aging (Bowley et al., 2010; Luebke et al., 2010; Hara et al., 2012; Koo et al., 2012, 2013). We also observed that the rhesus monkey brain develops throughout young ages (under 5) (Malkova et al., 2006; Scott et al., 2016; Kim et al., 2020), which is similar to human brain development (Ou et al., 2017; Sotardi et al., 2021); WM undergoes more maturation and neurodegeneration similar to that in human; whereas the GM in monkey brain, unlike in human, experiences less to no significant changes in volume beyond very early maturation phases (Rakic et al., 1986; Ge et al., 2002; Peters and Kemper, 2012; Shi et al., 2013). The different feature salience from our model between humans and monkeys may be explained by neuroanatomical differences across species. For example, monkeys do not show increased neurofibrillary tangles and amyloid plaques, which are pathological hallmarks primarily affecting human hippocampal and parahippocampal regions during early stages of aging and dementia (Luebke et al., 2010). Moreover, evidence suggests functional specialization of the human brain for language related to the expansion of temporal WM connections, which may give rise to variation in WM connectome between humans and monkeys (Sierpowska et al., 2022). The differences in monkey salience features compared to those of humans may reflect that our pre-trained AI models were well-adapted to the animal dataset after transfer learning.

There are several limitations of the study that we plan to address in our future analyses. First, we observed a regression-to-the-mean

(RTM) issue, similar to previous studies on age estimation in human brain MRIs (Bashyam et al., 2020; Feng et al., 2020; He et al., 2021b). Specifically, younger monkey brains were predicted older, and older brains were predicted younger (Supplementary Table S1 and Supplementary Figure S1). This is a well-known methodological issue for AI-based age estimation models. For our follow-up analyses, we aimed to address the RTM issue by re-balancing sample distribution across ages (Feng et al., 2020) or post-correction of AI-estimated brain ages (Liang et al., 2019; Peng et al., 2021). These procedures are ongoing work in the field (de Lange and Cole, 2020) but outside the main focus of the current study (transfer learning). Second, this study only focused on T1-weighted brain MRIs, similar to most AI brain age estimation articles in the human brain, as T1-weighted MRIs are most common in a large cohort (Bashyam et al., 2020; Feng et al., 2020; He et al., 2021b; Peng et al., 2021). Future studies may include diffusion and functional MRIs to further improve model performances. Third, among the 290 T1-weighted brain MRIs from 274 monkeys, there were 16 longitudinal scans or 6% of the monkeys had longitudinal scans, which is a much lower percentage of longitudinal scans compared to human brain age prediction studies with only dozen to lower hundreds of subjects (Zhang et al., 2018; Hu et al., 2019). It is possible that the addition of more longitudinal scans or using longitudinal information as a prior may improve model predictions. Fourth, we are yet to test the effect of scanner or site on AI accuracy in this study, similar to other human brain age estimation studies that also merged data from 10+ sites, and also due to heavily unbalanced data across sites (eight sites had less than five subjects, see Supplementary Table S2) and across scanners (95.5% of the scans were acquired on 3T scanners). Fifth, the saturated and highest possible accuracy in monkey age estimation remains unanswered due to the lack of large-scale monkey brain MRIs. It took at least 8,000 brain MRIs to saturate human brain age estimation (Kaufmann et al., 2019; Schulz et al., 2020). Nonetheless, we showed that given the limited monkey brain MRIs (dozens to lower hundreds), transfer learning brings higher accuracy and stability to AI models than without transfer learning. Our best-performing model also identified biologically meaningful age-related features and showed successful adaptation of the pre-trained model to the small sample target population as reflected by different saliency regions highlighted by the human pre-trained model and transferred monkey model which are consistent with previous studies.

Despite the limitations, we showed that transfer learning has a great potential to mitigate the small sample size problem in AI-driven brain age estimation using monkey MRIs and identified key features of aging in the frontal WM regions. Given the AI brain age estimator's wide success in human brain research, this first DL-based, transfer-learning-powered, lifespan-compatible, highly accurate, and publicly-released monkey brain age estimator could provide a simple but holistic measure of how brain changes in aging. Our study will also open wide opportunities to study monkey and other laboratory animal models under a variety of normal and pathological conditions, such as to help differentiate animals of the same chronological age but demonstrating varying levels of cognitive impairments.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found within the Supplementary material.

Ethics statement

The studies involving humans were approved by Boston Children's Hospital and Harvard Medical School. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements. The animal study was approved by the Boston University Institutional Animal Care and Use Committee.

Author contributions

Study design: SH, B-BK, and YO. Data analysis: SH, YG, CHC, B-BK, and YO. Manuscript writing: SH, YG, CHC, TM, JL, RK, DR, B-BK, and YO. All authors contributed to the article and approved the submitted version.

Funding

The study was supported, in part, by Charles A. King Trust Fellowship (SH), Massachusetts Life Science Center Bits to Bytes grants (YO), NIH/NICHD R03 HD104891, NIH/NICHD R03 HD107124, NIH/NINDS R21 NS121735, NIH/NCATS R21 TR004265, and NIH/NINDS R61 NS126792. The work was supported also, in part, by NIH/NIA RF1 AG062831-01, RF1-AG043640-06, and R01 AG042512 (DR), R01-AG071230 and R01-AG059028 (JL, YG, and B-BK), NIH/NIA R01 AG068168 (TM), NIH/NIA R56 AG059693, NIH/NINDS R21-NS102991, NIH/NINDS U01-NS076474, and NIH/NIA R01 AG043478 (TM and B-BK), R01AG055948 and GW210034 (B-BK and YG), and 1RF1AG083206 (B-BK, YG, and CHC).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of

their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Abrol, A., Fu, Z., Salman, M., Silva, R., Du, Y., Plis, S., et al. (2021). Deep learning encodes robust discriminative neuroimaging representations to outperform standard machine learning. *Nat. Commun.* 12, 1–17. doi: 10.1038/s41467-020-20655-6
- Bashyam, V. M., Erus, G., Doshi, J., Habes, M., Nasrallah, I. M., Truelove-Hill, M., et al. (2020). Mri signatures of brain age and disease over the lifespan based on a deep brain network and 14 468 individuals worldwide. *Brain* 143, 2312–2324. doi: 10.1093/brain/awaa160
- Bowley, M. P., Cabral, H., Rosene, D. L., and Peters, A. (2010). Age changes in myelinated nerve fibers of the cingulate bundle and corpus callosum in the rhesus monkey. *J. Comparat. Neurol.* 518, 3046–3064. doi: 10.1002/cne.22379
- Chen, X., Errangi, B., Li, L., Glasser, M. F., Westlye, L. T., Fjell, A. M., et al. (2013). Brain aging in humans, chimpanzees (pan troglodytes), and rhesus macaques (macaca mulatta): magnetic resonance imaging studies of macro-and microstructural changes. *Neurobiol. Aging* 34, 2248–2260. doi: 10.1016/j.neurobiolaging.2013.03.028
- Chung, Y., Addington, J., Bearden, C. E., Cadenhead, K., Cornblatt, B., Mathalon, D. H., et al. (2018). Use of machine learning to determine deviance in neuroanatomical maturity associated with future psychosis in youths at clinically high risk. *JAMA Psychiatry* 75, 960–968. doi: 10.1001/jamapsychiatry.2018.1543
- Cole, J. H., and Franke, K. (2017). Predicting age using neuroimaging: innovative brain ageing biomarkers. *Trends Neurosci.* 40, 681–690. doi: 10.1016/j.tins.2017.10.001
- Cole, J. H., Leech, R., Sharp, D. J., and Initiative, A. D. N. (2015). Prediction of brain age suggests accelerated atrophy after traumatic brain injury. *Ann. Neurol.* 77, 571–581. doi: 10.1002/ana.24367
- Cole, J. H., Ritchie, S. J., Bastin, M. E., Hernández, V., Mu noz Maniega, S., Royle, N., et al. (2018). Brain age predicts mortality. *Mol. Psychiatry* 23, 1385–1392. doi: 10.1038/mp.2017.62
- Daugherty, A. M., and Raz, N. (2017). A virtual water maze revisited: Two-year changes in navigation performance and their neural correlates in healthy adults. *Neuroimage* 146, 492–506. doi: 10.1016/j.neuroimage.2016.09.044
- Davatzikos, C., Genc, A., Xu, D., and Resnick, S. M. (2001). Voxel-based morphometry using the ravens maps: methods and validation using simulated longitudinal atrophy. *Neuroimage* 14, 1361–1369. doi: 10.1006/nimg.2001.0937
- de Lange, A.-M. G., and Cole, J. H. (2020). Commentary: Correction procedures in brain-age prediction. *NeuroImage: Clinical* 26, 102229. doi: 10.1016/j.nicl.2020.102229
- Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., et al. (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol. Psychiatry* 19, 659–667. doi: 10.1038/mp.2013.78
- Doshi, J., Erus, G., Ou, Y., Gaonkar, B., and Davatzikos, C. (2013). Multi-atlas skull-stripping. *Acad. Radiol.* 20, 1566–1576. doi: 10.1016/j.acra.2013.09.010
- Evans, A. C., Group, B. D. C., et al. (2006). The NIH MRI study of normal brain development. *Neuroimage* 30, 184–202. doi: 10.1016/j.neuroimage.2005.09.068
- Feng, X., Lipton, Z. C., Yang, J., Small, S. A., Provenzano, F. A., Initiative, A. D. N., et al. (2020). Estimating brain age based on a uniform healthy population with deep learning and structural magnetic resonance imaging. *Neurobiol. Aging* 91, 15–25. doi: 10.1016/j.neurobiolaging.2020.02.009
- Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., et al. (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat. Neurosci.* 18, 1664–1671. doi: 10.1038/nn.4135
- Franke, K., Clarke, G. D., Dahnke, R., Gaser, C., Kuo, A. H., Li, C., et al. (2017). Premature brain aging in baboons resulting from moderate fetal undernutrition. *Front. Aging Neurosci.* 9, 92. doi: 10.3389/fnagi.2017.00092
- Franke, K., Gaser, C., Manor, B., and Novak, V. (2013). Advanced brainage in older adults with type 2 diabetes mellitus. *Front. Aging Neurosci.* 5, 90. doi: 10.3389/fnagi.2013.00090
- Franke, K., Ziegler, G., Klöppel, S., Gaser, C., Initiative, A. D. N., 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, 883–892. doi: 10.1016/j.neuroimage.2010.01.005
- Ge, Y., Grossman, R. I., Babb, J. S., Rabin, M. L., Mannon, L. J., and Kolson, D. L. (2002). Age-related total gray matter and white matter changes in normal adult brain. Part i: volumetric mr imaging analysis. *Am. J. Neuroradiol.* 23, 1327–1333.
- Goyal, M. S., Blazey, T. M., Su, Y., Couture, L. E., Durbin, T. J., Bateman, R. J., et al. (2019). Persistent metabolic youth in the aging female brain. *Proc. Nat. Acad. Sci.* 116, 3251–3255. doi: 10.1073/pnas.1815917116
- Gray, D. T., and Barnes, C. A. (2019). Experiments in macaque monkeys provide critical insights into age-associated changes in cognitive and sensory function. *Proc. Nat. Acad. Sci.* 116, 26247–26254. doi: 10.1073/pnas.1902279116
- Guan, Y., Ebrahimzadeh, S. A., Cheng, C.-h., Chen, W., Leung, T., Bigornia, S., et al. (2022). Association of diabetes and hypertension with brain structural integrity and cognition in the boston puerto rican health study cohort. *Neurology* 98, e1534–e1544. doi: 10.1212/WNL.000000000000200120
- Guggenmos, M., Schmack, K., Sekutowicz, M., Garbusow, M., Sebold, M., Sommer, C., et al. (2017). Quantitative neurobiological evidence for accelerated brain aging in alcohol dependence. *Transl. Psychiatry* 7, 1–7. doi: 10.1038/s41398-017-0037-y
- Hara, Y., Rapp, P. R., and Morrison, J. H. (2012). Neuronal and morphological bases of cognitive decline in aged rhesus monkeys. *Age* 34, 1051–1073. doi: 10.1007/s11357-011-9278-5
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Las Vegas, NV), 770–778.
- He, S., Feng, Y., Grant, P. E., and Ou, Y. (2022). Deep relation learning for regression and its application to brain age estimation. *IEEE Trans. Med. Imaging* 41, 2304–2317. doi: 10.1109/TMI.2022.3161739
- He, S., Grant, P. E., and Ou, Y. (2021a). Global-local transformer for brain age estimation. *IEEE Trans. Med. Imaging* 41, 213–224. doi: 10.1109/TMI.2021.3108910
- He, S., Pereira, D., Perez, J. D., Gollub, R. L., Murphy, S. N., Prabhu, S., et al. (2021b). Multi-channel attention-fusion neural network for brain age estimation: accuracy, generality, and interpretation with 16,705 healthy mris across lifespan. *Med. Image Anal.* 72, 102091. doi: 10.1016/j.media.2021.102091
- Hollon, T. C., Pandian, B., Adapa, A. R., Urias, E., Save, A. V., Khalsa, S. S. S., et al. (2020). Near real-time intraoperative brain tumor diagnosis using stimulated raman histology and deep neural networks. *Nat. Med.* 26, 52–58. doi: 10.1038/s41591-019-0715-9
- Hu, D., Wu, Z., Lin, W., Li, G., and Shen, D. (2019). Hierarchical rough-to-fine model for infant age prediction based on cortical features. *IEEE J. Biomed. Health Informat.* 24, 214–225. doi: 10.1109/JBHI.2019.2897020
- Jeon, S., Ranard, K. M., Neuringer, M., Johnson, E. E., Renner, L., Kuchan, M. J., et al. (2018). Lutein is differentially deposited across brain regions following formula or breast feeding of infant rhesus macaques. *J. Nutr.* 148, 31–39. doi: 10.1093/jn/nxx023
- Jung, B., Taylor, P. A., Seidlitz, J., Sponheim, C., Perkins, P., Ungerleider, L. G., et al. (2021). A comprehensive macaque fmri pipeline and hierarchical atlas. *Neuroimage* 235, 117997. doi: 10.1016/j.neuroimage.2021.117997
- Kather, J. N., Pearson, A. T., Halama, N., Jäger, D., Krause, J., Loosen, S. H., et al. (2019). Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat. Med.* 25, 1054–1056. doi: 10.1038/s41591-019-0462-y
- Kaufmann, T., van der Meer, D., Doan, N. T., Schwarz, E., Lund, M. J., Agartz, I., et al. (2019). Common brain disorders are associated with heritable patterns of apparent aging of the brain. *Nat. Neurosci.* 22, 1617–1623.
- Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C., Liang, H., Baxter, S. L., et al. (2018). Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* 172, 1122–1131. doi: 10.1016/j.cell.2018.02.010
- Kim, J., Jung, Y., Barcus, R., Bachevalier, J. H., Sanchez, M. M., Nader, M. A., et al. (2020). Rhesus macaque brain developmental trajectory: a longitudinal analysis using tensor-based structural morphometry and diffusion tensor imaging. *Cerebral Cortex* 30, 4325–4335. doi: 10.1093/cercor/bhaa015
- Kingma, D. P., and Ba, J. (2014). Adam: a method for stochastic optimization. *arXiv [Preprint]*. arXiv:1412.6980. doi: 10.48550/arXiv.1412.6980
- Koo, B.-B., Calderazzo, S., Bowley, B. G., Kolli, A., Moss, M. B., Rosene, D. L., et al. (2018). Long-term effects of curcumin

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnagi.2023.1249415/full#supplementary-material>

- in the non-human primate brain. *Brain Res. Bull.* 142, 88–95. doi: 10.1016/j.brainresbull.2018.06.015
- Koo, B.-B., Oblak, A. L., Zhao, Y., Farris, C. W., Bowley, B., Rosene, D. L., et al. (2013). Hippocampal network connections account for differences in memory performance in the middle-aged rhesus monkey. *Hippocampus* 23, 1179–1188. doi: 10.1002/hipo.22156
- Koo, B.-B., Schettler, S. P., Murray, D. E., Lee, J.-M., Killiany, R. J., Rosene, D. L., et al. (2012). Age-related effects on cortical thickness patterns of the rhesus monkey brain. *Neurobiol. Aging* 33, 200–e23. doi: 10.1016/j.neurobiolaging.2010.07.010
- Kubicki, M., Baxi, M., Pasternak, O., Tang, Y., Karmacharya, S., Chunga, N., et al. (2019). Lifespan trajectories of white matter changes in rhesus monkeys. *Cerebral Cortex* 29, 1584–1593. doi: 10.1093/cercor/bhy056
- Kuchan, M. J., Ranard, K. M., Dey, P., Jeon, S., Sasaki, G. Y., Schimpf, K. J., et al. (2020). Infant rhesus macaque brain α -tocopherol stereoisomer profile is differentially impacted by the source of α -tocopherol in infant formula. *J. Nutr.* 150, 2305–2313. doi: 10.1093/jn/nxaa174
- Lacourse, A., and Herndon, J. (2009). *Nonhuman Primate Models of Cognitive Aging in Animal Models of Human Cognitive Aging*. New York, NY: Humana Press, 29–58.
- LaMontagne, P. J., Keefe, S., Lauren, W., Xiong, C., Grant, E. A., Moulder, K. L., et al. (2018). Oasis-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer's disease. *Alzheimer's Dement.* 14, P1097. doi: 10.1101/2019.12.13.19014902
- Lee, J., Burkett, B. J., Min, H.-K., Senjem, M. L., Lundt, E. S., Botha, H., et al. (2022). Deep learning-based brain age prediction in normal aging and dementia. *Nature Aging* 2, 412–424. doi: 10.1038/s43587-022-00219-7
- Liang, H., Zhang, F., and Niu, X. (2019). *Investigating Systematic Bias in Brain Age Estimation With Application To Post-Traumatic Stress Disorders. Technical report*. Hoboken, NJ: Wiley Online Library. doi: 10.1002/hbm.24588
- Liem, F., Varoquaux, G., Kynast, J., Beyer, F., Masouleh, S. K., Huntenburg, J. M., et al. (2017). Predicting brain-age from multimodal imaging data captures cognitive impairment. *Neuroimage* 148, 179–188. doi: 10.1016/j.neuroimage.2016.11.005
- Luders, E., Cherbuin, N., and Gaser, C. (2016). Estimating brain age using high-resolution pattern recognition: Younger brains in long-term meditation practitioners. *Neuroimage* 134, 508–513. doi: 10.1016/j.neuroimage.2016.04.007
- Luebeck, J., Barbas, H., and Peters, A. (2010). Effects of normal aging on prefrontal area 46 in the rhesus monkey. *Brain Res. Rev.* 62, 212–232. doi: 10.1016/j.brainresrev.2009.12.002
- Malkova, L., Heuer, E., and Saunders, R. (2006). Longitudinal magnetic resonance imaging study of rhesus monkey brain development. *Eur. J. Neurosci.* 24, 3204–3212. doi: 10.1111/j.1460-9568.2006.05175.x
- Mattison, J. A., and Vaughan, K. L. (2017). An overview of nonhuman primates in aging research. *Exp. Gerontol.* 94, 41–45. doi: 10.1016/j.exger.2016.12.005
- Mellus, E. L. (1905). A study of the location and arrangement of the giant cells in the cortex of the right hemisphere of the bonnet monkey (*macacus sinicus*). *Am. J. Anat.* 4, 405–408. doi: 10.1002/aja.1000040403
- Mellus, E. L. (1907). Relations of the frontal lobe in the monkey. *Am. J. Anat.* 7, 227–243. doi: 10.1002/aja.1000070205
- Milham, M. P., Ai, L., Koo, B., Xu, T., Amiez, C., Balezeau, F., et al. (2018). An open resource for non-human primate imaging. *Neuron* 100, 61–74. doi: 10.1016/j.neuron.2018.08.039
- Moore, J. H., Raghavachari, N., et al. (2019). Artificial intelligence based approaches to identify molecular determinants of exceptional health and life span-an interdisciplinary workshop at the national institute on aging. *Front. Artif. Intellig.* 2, 12. doi: 10.3389/frai.2019.00012
- Moss, M., Moore, T., Schettler, S., Killiany, R., Rosene, D., and Riddle, D. (2007). *Brain aging: Models, Methods, and Mechanisms*. Florida: CRC Press, 21–38.
- Nastase, S. A., Liu, Y.-F., Hillman, H., Zadbood, A., Hasenfratz, L., Keshavarzian, N., et al. (2021). Narratives: fMRI data for evaluating models of naturalistic language comprehension. *bioRxiv*. 2020, 12. doi: 10.1038/s41597-021-01033-3
- Ning, K., Zhao, L., Matloff, W., Sun, F., and Toga, A. W. (2020). Association of relative brain age with tobacco smoking, alcohol consumption, and genetic variants. *Sci. Rep.* 10, 1–10. doi: 10.1038/s41598-019-56089-4
- Nooner, K. B., Colcombe, S., Tobe, R., Mennes, M., Benedict, M., Moreno, A., et al. (2012). The NKI-Rockland sample: a model for accelerating the pace of discovery science in psychiatry. *Front. Neurosci.* 6, 152. doi: 10.3389/fnins.2012.00152
- Ou, Y., Akbari, H., Bilello, M., Da, X., and Davatzikos, C. (2014). Comparative evaluation of registration algorithms in different brain databases with varying difficulty: results and insights. *IEEE Trans. Med. Imaging* 33, 2039–2065. doi: 10.1109/TMI.2014.2330355
- Ou, Y., Gollub, R. L., Retzeppi, K., Reynolds, N., Pienaar, R., Pieper, S., et al. (2015). Brain extraction in pediatric ADC maps, toward characterizing neuro-development in multi-platform and multi-institution clinical images. *Neuroimage* 122, 246–261. doi: 10.1016/j.neuroimage.2015.08.002
- Ou, Y., Sotiras, A., Paragios, N., and Davatzikos, C. (2011). DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting. *Med. Image Anal.* 15, 622–639. doi: 10.1016/j.media.2010.07.002
- Ou, Y., Zöllei, L., Da, X., Retzeppi, K., Murphy, S. N., Gerstner, E. R., et al. (2018). Field of view normalization in multi-site brain MRI. *Neuroinformatics* 16, 431–444. doi: 10.1007/s12021-018-9359-z
- Ou, Y., Zöllei, L., Retzeppi, K., Castro, V., Bates, S. V., Pieper, S., et al. (2017). Using clinically acquired mri to construct age-specific adc atlases: Quantifying spatiotemporal adc changes from birth to 6-year old. *Hum. Brain Mapp.* 38, 3052–3068. doi: 10.1002/hbm.23573
- Park, J., Carp, J., Kennedy, K. M., Rodrigue, K. M., Bischof, G. N., Huang, C.-M., et al. (2012). Neural broadening or neural attenuation? Investigating age-related dedifferentiation in the face network in a large lifespan sample. *J. Neurosci.* 32, 2154–2158. doi: 10.1523/JNEUROSCI.4494-11.2012
- Peng, H., Gong, W., Beckmann, C. F., Vedaldi, A., and Smith, S. M. (2021). Accurate brain age prediction with lightweight deep neural networks. *Med. Image Anal.* 68, 101871. doi: 10.1016/j.media.2020.101871
- Peters, A. (2009). The effects of normal aging on myelinated nerve fibers in monkey central nervous system. *Front. Neuroanat.* 11, 9. doi: 10.3389/fnro.2009.05.011.2009
- Peters, A., and Kemper, T. (2012). A review of the structural alterations in the cerebral hemispheres of the aging rhesus monkey. *Neurobiol. Aging* 33, 2357–2372. doi: 10.1016/j.neurobiolaging.2011.11.015
- Poddar, J., Pradhan, M., Ganguly, G., and Chakrabarti, S. (2019). Biochemical deficits and cognitive decline in brain aging: Intervention by dietary supplements. *J. Chem. Neuroanat.* 95, 70–80. doi: 10.1016/j.jchemneu.2018.04.002
- Rakic, P., Bourgeois, J.-P., Eckenhoff, M. F., Zecevic, N., and Goldman-Rakic, P. S. (1986). Concurrent overproduction of synapses in diverse regions of the primate cerebral cortex. *Science* 232, 232–235. doi: 10.1126/science.3952506
- Rogenmoser, L., Kernbach, J., Schlaug, G., and Gaser, C. (2018). Keeping brains young with making music. *Brain Struct. Funct.* 223, 297–305. doi: 10.1007/s00429-017-1491-2
- Rohlfing, T., Zahr, N. M., Sullivan, E. V., and Pfefferbaum, A. (2010). The SRI24 multichannel atlas of normal adult human brain structure. *Hum. Brain Mapp.* 31, 798–819. doi: 10.1002/hbm.20906
- Roth, G. S., Mattison, J. A., Ottinger, M. A., Chachich, M. E., Lane, M. A., and Ingram, D. K. (2004). Aging in rhesus monkeys: relevance to human health interventions. *Science* 305, 1423–1426. doi: 10.1126/science.1102541
- Saad, Z. S., Glen, D. R., Chen, G., Beauchamp, M. S., Desai, R., and Cox, R. W. (2009). A new method for improving functional-to-structural mri alignment using local pearson correlation. *Neuroimage* 44, 839–848. doi: 10.1016/j.neuroimage.2008.09.037
- Schulz, M.-A., Yeo, B., Vogelstein, J. T., Mourao-Miranada, J., Kather, J. N., Kording, K., et al. (2020). Different scaling of linear models and deep learning in ukbiobank brain images versus machine-learning datasets. *Nat. Commun.* 11, 1–15. doi: 10.1038/s41467-020-18037-z
- Scott, J. A., Grayson, D., Fletcher, E., Lee, A., Bauman, M. D., Schumann, C. M., et al. (2016). Longitudinal analysis of the developing rhesus monkey brain using magnetic resonance imaging: birth to adulthood. *Brain Struct. Funct.* 221, 2847–2871. doi: 10.1007/s00429-015-1076-x
- Seidlitz, J., Sponheim, C., Glen, D., Frank, Q. Y., Saleem, K. S., Leopold, D. A., et al. (2018). A population mri brain template and analysis tools for the macaque. *Neuroimage* 170, 121–131. doi: 10.1016/j.neuroimage.2017.04.063
- Shi, Y., Short, S. J., Knickmeyer, R. C., Wang, J., Coe, C. L., Niethammer, M., et al. (2013). Diffusion tensor imaging-based characterization of brain neurodevelopment in primates. *Cerebral Cortex* 23, 36–48. doi: 10.1093/cercor/bhr372
- Sierpowska, J., Bryant, K. L., Janssen, N., Blazquez Freches, G., Römkens, M., Mangnus, M., et al. (2022). Comparing human and chimpanzee temporal lobe neuroanatomy reveals modifications to human language hubs beyond the frontotemporal arcuate fasciculus. *Proc. Nat. Acad. Sci.* 119, e2118295119. doi: 10.1073/pnas.2118295119
- Simmons, H. (2016). Age-associated pathology in rhesus macaques (*macaca mulatta*). *Vet. Pathol.* 53, 399–416. doi: 10.1177/0300985815620628
- Simonyan, K., and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv [Preprint]*. arXiv:1409.1556. doi: 10.48550/arXiv.1409.1556
- Smith, S. M., Elliott, L. T., Alfaro-Almagro, F., McCarthy, P., Nichols, T. E., Douaud, G., et al. (2020). Brain aging comprises many modes of structural and functional change with distinct genetic and biophysical associations. *Elife* 9, e52677. doi: 10.7554/eLife.52677
- Sotardi, S., Gollub, R. L., Bates, S. V., Weiss, R., Murphy, S. N., Grant, P. E., et al. (2021). Voxelwise and regional brain apparent diffusion coefficient changes on mri from birth to 6 years of age. *Radiology* 298, 415. doi: 10.1148/radiol.2020202279
- Souder, D. C., Dreischmeier, I. A., Smith, A. B., Wright, S., Martin, S. A., Sagar, M. A. K., et al. (2021). Rhesus monkeys as a translational model for late-onset Alzheimer's disease. *Aging Cell* 20, e13374. doi: 10.1111/acel.13374

- Steffener, J., Habeck, C., O'Shea, D., Razlighi, Q., Bherer, L., and Stern, Y. (2016). Differences between chronological and brain age are related to education and self-reported physical activity. *Neurobiol. Aging* 40, 138–144. doi: 10.1016/j.neurobiolaging.2016.01.014
- Ström, P., Kartasalo, K., Olsson, H., Solorzano, L., Delahunt, B., Berney, D. M., et al. (2020). Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. *Lancet Oncol.* 21, 222–232. doi: 10.1016/S1470-2045(19)30738-7
- Sundararajan, M., Taly, A., and Yan, Q. (2017). "Axiomatic attribution for deep networks," in *International Conference on Machine Learning*, 3319–3328. doi: 10.48550/arXiv.1703.01365
- Taylor, J. R., Williams, N., Cusack, R., Auer, T., Shafto, M. A., Dixon, M., et al. (2017). The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) data repository: structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample. *Neuroimage* 144, 262–269. doi: 10.1016/j.neuroimage.2015.09.018
- Tigges, J., Gordon, T. P., McClure, H. M., Hall, E. C., and Peters, A. (1988). Survival rate and life span of rhesus monkeys at the yerkes regional primate research center. *Am. J. Primatol.* 15, 263–273. doi: 10.1002/ajp.1350150308
- Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., et al. (2010). N4itk: improved n3 bias correction. *IEEE Trans. Med. Imaging* 29, 1310–1320. doi: 10.1109/TMI.2010.2046908
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). "Attention is all you need," in *Advances in Neural Information Processing Systems* (Long Beach, CA), 30.
- Wei, D., Zhuang, K., Ai, L., Chen, Q., Yang, W., Liu, W., et al. (2018). Structural and functional brain scans from the cross-sectional Southwest University adult lifespan dataset. *Sci. Data* 5, 180134. doi: 10.1038/sdata.2018.134
- Wisco, J. J., Killiany, R. J., Guttman, C. R., Warfield, S. K., Moss, M. B., and Rosene, D. L. (2008). An mri study of age-related white and gray matter volume changes in the rhesus monkey. *Neurobiol. Aging* 29, 1563–1575. doi: 10.1016/j.neurobiolaging.2007.03.022
- Zhang, C., Adeli, E., Wu, Z., Li, G., Lin, W., and Shen, D. (2018). Infant brain development prediction with latent partial multi-view representation learning. *IEEE Trans. Med. Imaging* 38, 909–918. doi: 10.1109/TMI.2018.2874964