

## RESEARCH ARTICLE

# Multimodal biological brain age prediction using magnetic resonance imaging and angiography with the identification of predictive regions

Pauline Mouches<sup>1,2,3</sup>  | Matthias Wilms<sup>1,2,4</sup>  | Deepthi Rajashekar<sup>1,2,3</sup>  |  
Sönke Langner<sup>5</sup> | Nils D. Forkert<sup>1,2,4</sup> 

<sup>1</sup>Department of Radiology, University of Calgary, Calgary, Alberta, Canada

<sup>2</sup>Hotchkiss Brain Institute, University of Calgary, Calgary, Alberta, Canada

<sup>3</sup>Biomedical Engineering Program, University of Calgary, Calgary, Alberta, Canada

<sup>4</sup>Alberta Children's Hospital Research Institute, University of Calgary, Calgary, Alberta, Canada

<sup>5</sup>Institute for Diagnostic Radiology and Neuroradiology, Rostock University Medical Center, Rostock, Germany

## Correspondence

Pauline Mouches, Department of Radiology, Faculty of Medicine, University of Calgary, 3330 Hospital Drive NW, Calgary, AB T2N 4N1, Canada.

Email: [pauline.mouches@ucalgary.ca](mailto:pauline.mouches@ucalgary.ca)

## Funding information

Canada Research Chairs; Canadian Open Neuroscience Platform; Natural Sciences and Engineering Research Council of Canada; River Fund at Calgary Foundation

## Abstract

Biological brain age predicted using machine learning models based on high-resolution imaging data has been suggested as a potential biomarker for neurological and cerebrovascular diseases. In this work, we aimed to develop deep learning models to predict the biological brain age using structural magnetic resonance imaging and angiography datasets from a large database of 2074 adults (21–81 years). Since different imaging modalities can provide complementary information, combining them might allow to identify more complex aging patterns, with angiography data, for instance, showing vascular aging effects complementary to the atrophic brain tissue changes seen in T1-weighted MRI sequences. We used saliency maps to investigate the contribution of cortical, subcortical, and arterial structures to the prediction. Our results show that combining T1-weighted and angiography MR data led to a significantly improved brain age prediction accuracy, with a mean absolute error of 3.85 years comparing the predicted and chronological age. The most predictive brain regions included the lateral sulcus, the fourth ventricle, and the amygdala, while the brain arteries contributing the most to the prediction included the basilar artery, the middle cerebral artery M2 segments, and the left posterior cerebral artery. Our study proposes a framework for brain age prediction using multimodal imaging, which gives accurate predictions and allows identifying the most predictive regions for this task, which can serve as a surrogate for the brain regions that are most affected by aging.

## KEYWORDS

brain aging, deep learning, magnetic resonance angiography, magnetic resonance imaging

## 1 | INTRODUCTION

Normal brain aging is generally associated with morphological changes of the cortical, subcortical (Potvin et al., 2016; Potvin, Dieumegarde, & Duchesne, 2017), and cerebrovascular structures (Bullitt et al., 2010).

More precisely, aging was found to be associated with a decrease of brain tissue (atrophy) and a blood vessel loss in adults, although it is not clear if the vessels actually disappear or if they become invisible on magnetic resonance angiography due to decreased blood flow (Amin-Hanjani, Du, Pandey, Thulborn, & Charbel, 2015) or

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Human Brain Mapping* published by Wiley Periodicals LLC.

atherosclerosis reducing the inner lumen. Across the various brain and vascular structures, these changes are often reported to have a nonlinear trend with respect to age (Peters, 2006), with some structures undergoing increasing atrophy rate with increasing age, such as the hippocampus, or following a U-shape volume change with aging, such as the caudate (Potvin et al., 2016). However, these morphological changes do not only occur due to normal brain aging but are also associated with neurological or cerebrovascular diseases such as Alzheimer's disease, transient ischemic attacks, or stenoses (Foteno, Snyder, Girton, Morris, & Buckner, 2005; Ritz, Denswil, Stam, van Lieshout, & Daemen, 2014; Wardlaw et al., 2013) with significantly accelerated brain aging effects. For clinical diagnosis and decision making, it is important to improve our understanding of normal brain aging to enable a precise diagnosis of any pathological deviation (i.e., accelerated or abnormal aging) at an early time point when potential treatments and interventions are arguably the most effective.

The morphology of cortical and subcortical structures is typically assessed using structural high-resolution T1-weighted magnetic resonance imaging (MRI) scans, while the diameter and density of cerebrovascular structures can be measured, for example, using time-of-flight magnetic resonance angiography (TOF MRA). TOF MRA imaging provides a good blood-to-background contrast even without any exogenous contrast agent (Forkert et al., 2013). However, extracting artery-related quantitative features from TOF MRA datasets, such as the artery diameter or density, requires a lot of processing steps, including generating a segmentation of the cerebrovascular system, that can bias the measurements. The same is also true for the segmentation of high-resolution T1-weighted datasets, although to a lesser extent. Therefore, being able to analyze the raw T1-weighted MRI and TOF MRA datasets directly has the potential to significantly reduce the data preprocessing time needed when manually extracting measurements and allows the model to automatically select the most relevant features from the images.

In the brain aging literature, it has been suggested that the biological brain age might be a sensitive biomarker for several neurological diseases associated with accelerated or abnormal brain aging (Cole & Franke, 2017; Rokicki et al., 2020). More precisely, abnormal aging can be identified and quantified by calculating the difference between the chronological and the biological brain age of a person. The biological brain age is typically estimated using predictive models based on neuroimaging data as input. Such predictive models are typically trained using data from healthy subjects, assuming that the biological brain age is equal to the chronological age in a healthy individual. Various statistical and machine learning methods and types of input data have been used for biological brain age prediction so far. Among others, previous research has aimed to predict biological brain age based on numerical features extracted from T1-weighted MRI scans (Valizadeh, Hänggi, Méritat, & Jäncke, 2017), raw T1-weighted MRI scans (Cole et al., 2017; Peng, Gong, Beckmann, Vedaldi, & Smith, 2021), multimodal neuroimaging data often combining various numerical features extracted from structural and functional MRI modalities (Cole, 2020; de Lange et al., 2020; Niu, Zhang, Kounios, & Liang, 2020), and cerebral blood flow information using arterial spin

labeling data (MacDonald et al., 2020; Rokicki et al., 2020). While brain age prediction using T1-weighted MRI data has been of high interest in past studies, only one study (to the best of our knowledge) has attempted to predict the biological brain age using TOF MRA datasets (Nam et al., 2020), although vascular changes due to aging are well documented (Mouches, Langner, Domin, Hill, & Forkert, 2021). However, the contributions of the different arteries to the age prediction result have not been investigated yet. Additionally, the TOF MRA datasets were used as the only imaging modality in this case, thereby, potentially missing the highly informative complementary value of T1-weighted MRI datasets. Finally, most studies use data collected at different sites, acquired on different scanners, and with varying scanning parameters (Levakov, Rosenthal, Shelef, Raviv, & Avidan, 2020). While combining data from different sources results in larger training populations, it can also introduce different biases related to scanning protocols and hardware (Krugger, Turner, & Muftuler, 2010). These biases are known to negatively affect brain age prediction accuracies. Hence, specific model architectures were proposed to reduce the impact of, for example, scanning parameters on the predictions (Dinsdale, Jenkinson, & Namburete, 2021). Furthermore, various data harmonization techniques have been proposed in the past to remove site- and scanner-related effects (Pomponio et al., 2020; Wrobel et al., 2020).

Recent brain age estimation approaches that use imaging data directly without requiring any explicit feature engineering mostly employ convolutional neural networks (CNNs) with T1-weighted MRI data as the only input (Bashyam et al., 2020; Cole et al., 2017; Peng et al., 2021). Briefly, CNNs are a class of deep artificial neural networks that use imaging data as inputs and transform them into the desired outputs through layers of convolution operations with many different convolution kernels. The convolution kernels are similar to feature extractors in more traditional machine learning models, but they are automatically learned/optimized during model training instead of being hand-crafted by the user (Lo Vercio et al., 2020). In the context of brain aging, this allows the CNN to learn to use optimal features directly from the image to predict the biological brain age more accurately (O'Shea & Nash, 2015). However, such models are usually deemed black boxes as the user does not know which parts of the input image particularly contributed to make the prediction. This is in stark contrast to traditional machine learning models for which several methods can be applied for assessing the feature importance (Lo Vercio et al., 2020). To overcome this issue and to add explainability to CNN models, saliency map techniques (Samek, Montavon, Lapuschkin, Anders, & Müller, 2021; Simonyan, Vedaldi, & Zisserman, 2013) can be used to identify the regions of the input image that contribute the most to the brain age prediction. Saliency maps provide an established approach in the machine learning domain for visualizing the importance of each input image voxel for the prediction made by a trained CNN. Technically, this is achieved by computing the gradient of the function parameterized by the CNN with respect to the individual voxels (i.e., answering the question "How would changing this voxel affect the predicted age?"). For this specific task, such techniques were, for example, used in Levakov et al. (2020)

to identify the most important regions in a brain age prediction method based on T1-weighted MRI datasets. However, this explainability method has not been investigated for the use with multimodal imaging inputs to understand the relative contribution of each brain region in each imaging modality.

The primary aim of this work is to develop and evaluate a multimodal CNN-based biological brain age prediction model that jointly uses T1-weighted MRI and TOF MRA scans. The major technical contributions of this article are as follows: (a) the validation of an existing brain age prediction model architecture and (b) the use of saliency maps in the context of multimodal biological brain age prediction. The clinically relevant contributions include (a) the investigation of the added benefit of combining two modalities that provide complementary brain tissue and artery information, (b) the age-specific identification of important brain regions and arteries contributing most to the brain age estimation, and (c) the use of a large database of adults without any known brain pathologies or neurological diseases (21–82 years old), collected on a single scanner with the same scanning parameters, which avoids data acquisition-induced biases often encountered when combining data from different centers and/or different studies.

## 2 | MATERIAL AND METHODS

### 2.1 | Database

All datasets used for this study were acquired within the Study of Health in Pomerania (SHIP). All SHIP participants were randomly selected in the region of Pomerania in Germany, with the aim of obtaining a representative general population sample (Völzke et al., 2011). Thus, the database includes predominantly Caucasian participants with no known pathologies in brain MRI scans or presence of any neurodegenerative diseases. For this secondary study, T1-weighted MRI and TOF MRA datasets from 2,118 adults (1,029 males and 1,089 females), aged between 21 and 82 years (mean:  $51 \pm 14$ ) were used.

### 2.2 | Data acquisition and preprocessing

All scans for each participant were acquired on a single 1.5T system (Magnetom Avanto; Siemens Medical Solutions, Erlangen, Germany) with the following acquisition parameters: T1-weighted MRI: TR = 1900 ms, TE = 3.4 ms, flip angle =  $15^\circ$ , spacing =  $1.0 \times 1.0 \times 1.0 \text{ mm}^3$ ; TOF MRA: TR = 23 ms, TE = 7 ms, flip angle =  $25^\circ$ , spacing =  $0.7 \times 0.7 \times 0.7 \text{ mm}^3$ .

All datasets were corrected for bias field inhomogeneities using the N4 algorithm (Tustison et al., 2010), skull stripped (Isensee et al., 2019) to remove any nonbrain tissues, and affinely registered to the MNI brain atlas (Mazziotta et al., 2001). Using an affine registration allows to align the images while preserving participant-specific brain morphology, as opposed to nonlinear registration (Dinsdale,

Bluemke, 2021). Finally, and following common practice in machine learning, the intensities of all images were rescaled to have zero mean and unit variance to ensure that all datasets have a similar intensity range. Figure 1 shows an example of preprocessed data, as well as the maximum intensity projection of a TOF MRA dataset showing the arteries visible in the dataset.

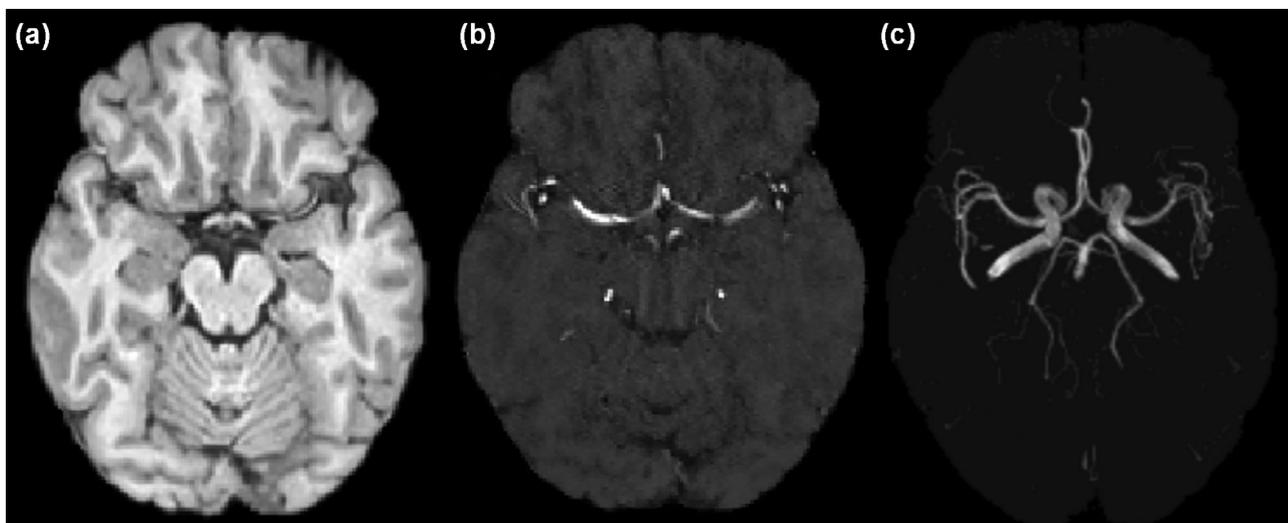
For the TOF MRA datasets, a custom brain mask that also includes the arteries at the base of the brain (the petrous segment of the internal carotid arteries) was used to perform the skull stripping. More precisely, the custom brain mask was delineated in the MNI brain atlas space, using the cerebrovascular atlas from Mouches and Forkert (2019), and transformed to each TOF MRA dataset. This was achieved by computing a nonlinear transformation between the MNI brain atlas and each T1-weighted MRI dataset and a rigid transformation between each T1-weighted MRI dataset and its corresponding TOF MRA dataset. After this, the two transformations were concatenated and used to transform the brain mask to each TOF MRA dataset. The same principle of concatenating registration transformations was used to affinely register the TOF MRA datasets to the MNI brain atlas space. All registrations were performed using the Advanced Normalization Tools (ANTs) toolkit (Avants, Tustison, & Johnson, 2009).

Finally, all affinely aligned datasets were cropped to remove background voxels. Considering that the field of view in the craniocaudal direction was considerably smaller for the TOF MRA datasets than for the T1-weighted MRI datasets, the final image size was  $155 \times 190 \times 50$  for the TOF MRA datasets and  $144 \times 192 \times 160$  for the T1-weighted MRI datasets.

All preprocessed TOF MRA datasets were visually inspected to exclude datasets in which the image acquisition was incomplete and did not cover the full cropping mask extent in the craniocaudal direction. Consequently, 44 of the initial 2,118 participants were excluded resulting in a total sample size of 2,074 participants.

### 2.3 | Brain age prediction

Three different models predicting the brain age were developed in this work. In a first step, two models with the same architecture, inspired by the Simple Fully Convolutional Neural Network (SFCN) model (Peng et al., 2021), were trained using different inputs: T1-weighted MRI ( $\text{CNN}_{\text{T1}}$ ) and TOF MRA ( $\text{CNN}_{\text{TOF}}$ ) datasets. This architecture was chosen as it was specifically designed for the brain age prediction task and is one of the best performing models on the UK Biobank data (Sudlow et al., 2015). The model architecture consists of six blocks. The first four blocks include one 3-dimensional convolutional layer with  $(3 \times 3 \times 3)$  kernels, one batch normalization layer, and one  $(2 \times 2 \times 2)$  max pooling layer followed by a ReLU nonlinear activation (Nair & Hinton, 2010). The convolutional layers for the first four blocks have 32, 64, 128, and 256 filters, respectively. The fifth block consists of one 3-dimensional convolutional layer with a  $(1 \times 1 \times 1)$  kernel and 64 filters, followed by a batch normalization layer and ReLU activation. The model ends with a sixth block



**FIGURE 1** Example of preprocessed datasets used as input for the brain age prediction models. (a) T1-weighted MRI dataset; (b) TOF MRA dataset; (c) maximum intensity projection of the TOF MRA dataset in cranio-caudal direction showing the arteries included in the TOF MRA dataset

consisting of an averaging pooling layer, a dropout layer with 0.5 dropout rate, and a dense layer with linear activation outputting the age prediction (Figure 2b). The models were optimized during training using the mean squared error between the predicted and chronological brain age as loss function.

The CNN models were trained from scratch using the available data. The data were split into 65% training data (1,340 datasets), 15% validation data (334 datasets), and 20% testing data (400 datasets), with an equivalent percentage of participants of each age in each split, as illustrated in Figure 2a. A data augmentation strategy applying random  $\pm 5^\circ$  rotations and  $\pm 10$  voxel translations on 50% of the datasets in each batch was employed to reduce the risk of model overfitting, which led to slightly improved results (Table S1, Supporting Information). The two models were trained using the Adam optimizer (Kingma & Ba, 2017) with a learning rate of 0.001, which was empirically determined, a weight decay of 0.0003, and a batch size of 8 using TensorFlow (Abadi et al., 2016). The validation set was used for early stopping of the optimization process.

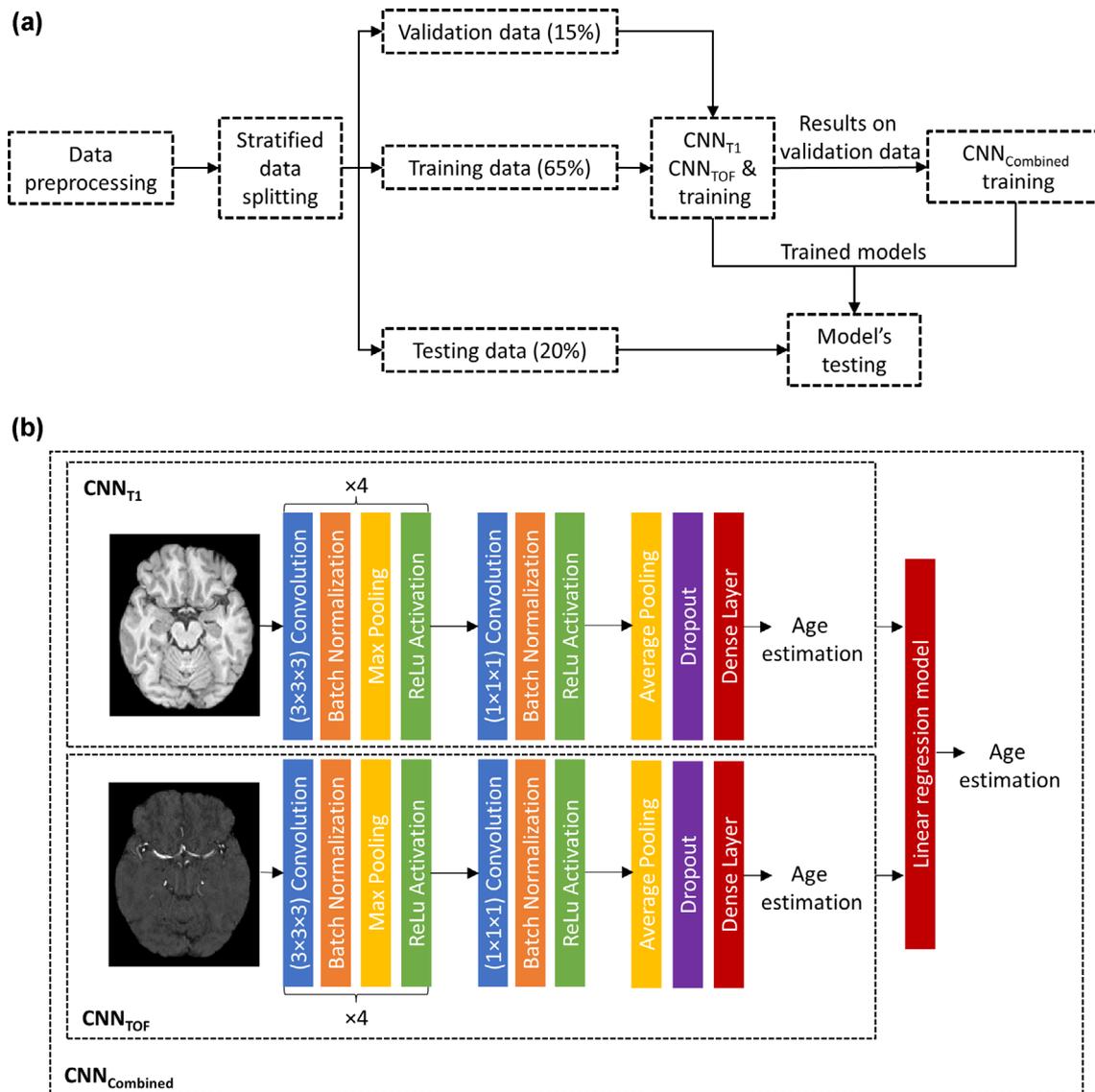
Next, the outputs from the  $\text{CNN}_{T1}$  and  $\text{CNN}_{TOF}$  models were combined using a multiple linear regression model, following the approach described in Jonsson et al. (2019). Therefore, the biological brain age of the 334 validation datasets was estimated using the trained  $\text{CNN}_{T1}$  and  $\text{CNN}_{TOF}$  models. Then, a multiple linear regression model with the brain age prediction from the  $\text{CNN}_{T1}$  and  $\text{CNN}_{TOF}$  models as predictor variables and the chronological age as target variable was trained using the 334 validation datasets (see Figure 2a). Once trained, the linear regression model was concatenated with the previously trained  $\text{CNN}_{T1}$  and  $\text{CNN}_{TOF}$  models to form the  $\text{CNN}_{\text{Combined}}$  model, which is used to predict the brain age of the test data and to generate saliency maps as described in the following. Combining both modality-specific models at the decision level allows to generate saliency maps representing the relative contribution of each modality and circumvent problems previously observed when

combining modalities at the input level (Jin, Li, & Hamarneh, 2021). Figure 2b illustrates the three different models and their architecture.

## 2.4 | Saliency maps generation

Individual saliency maps were computed for each participant of the test set using the  $\text{CNN}_{\text{Combined}}$  model. To do so, the SmoothGrad algorithm (Smilkov, Thorat, Kim, Viégas, & Wattenberg, 2017) was adapted to this regression problem. This method was specially chosen as it is simple, has previously shown to lead to good results in the brain age prediction context (Levakov et al., 2020), and directly outputs voxel-wise gradient values. In comparison, methods using class activation maps, such as Grad-Cam (Selvaraju et al., 2020), should be applied on the last convolutional layer of a CNN model, which in the current model contains highly downsampled activation maps compared to the input image size and would result in spatially inaccurate saliency maps. Briefly described, this method applies random noise from a normal distribution  $N(0,0.1)$  to each test dataset and passes the noisy image as input through the trained model. Then, the partial derivative of the loss function with respect to the noisy input image is backpropagated through the model, resulting in one value per image voxel representing the voxel importance. This process is repeated 10 times for each dataset and the output maps are averaged. For the  $\text{CNN}_{\text{Combined}}$  model, the gradient going through the linear regression model used to combine the two modality-specific models was included in the calculations. In this multimodal input setting, the noise was applied to both inputs and the output consists of one saliency map per modality per test participant, in which the gradient values account for the relative importance of each modality.

As saliency maps have been shown to be not reliable in some cases, it is important to ensure their robustness against the model weights, label randomization, as well as their repeatability and



**FIGURE 2** Flowchart of the proposed approach (a) and model (b). CNN<sub>T1</sub> and CNN<sub>TOF</sub> are two sub-models using 3-dimensional T1-weighted MRI and TOF MRA datasets as input, respectively, and outputting the estimated brain age. CNN<sub>Combined</sub> combines the estimations from the two single modality models and outputs the final estimated brain age

localization relevance (Arun et al., 2021). Therefore, sanity checks were conducted, following Adebayo et al. (2018). The first sanity check consists of instantiating a model with the proposed architecture but random weights and generating the saliency maps on the test data with this untrained model. By doing so, it is possible to check if the saliency maps are sensitive to the model weights. The second sanity check consists of training a new model with the proposed architecture but using randomized input data where the age information is randomly permuted across samples. This second sanity check ensures that the maps truly reflect the dependency between the input data and the outcome. Finally, generating maps for a large number of participants demonstrate the repeatability aspect and comparing the regions showing high importance to the literature ensures localization relevance.

## 2.5 | Regions of importance identification

In the next step, z-score maps of the T1-weighted MRI and TOF MRA saliency maps of each participant in the test set were computed conjointly to preserve the information about the relative importance of each modality. The z-score maps were then nonlinearly transformed to the MNI brain atlas space (using the already available registration-derived transformations; see section 2.2) and averaged for all participants per age decade. This process resulted in six age-specific average saliency maps corresponding to participants aged below 30, between 30 and 40, between 40 and 50, between 50 and 60, between 60 and 70, and over 70. The six age-specific average saliency maps were finally used as the input to the popular probabilistic threshold free cluster enhancement algorithm (Spisák et al., 2019) with the aim of

determining significant clusters of predictive voxels ( $p < .05$ , corrected for multiple comparisons) within each map.

To identify the most important brain regions for the brain age prediction task, brain atlas regions of interest (ROIs) defined in the MNI space were used. For the T1-weighted MRI data, the cortical and subcortical regions described in the CerebrA atlas (Manera, Dadar, Fonov, & Collins, 2020), which is a corrected and improved version of the Mindboggle-101 atlas (Klein & Tourville, 2012), and the cortical regions described in the human finer grain cerebral cortex atlas (Glasser et al., 2016), were used as parcellations. These specific atlases were chosen as the CerebrA atlas contains the cerebrospinal fluid filled structures, which were previously identified as important markers of brain aging (Lemaître et al., 2005; Levakov et al., 2020), and the human cerebral cortex atlas provides a fine parcellation of biologically relevant regions of the cortex. For the TOF MRA datasets, masks of the main arteries, including the basilar artery (BA), the anterior cerebral arteries (ACA) A1 and A2 segments, the middle cerebral arteries (MCA) M1 and M2 segments, the posterior cerebral arteries (PCA), and the internal carotid arteries (ICA) were manually defined using the cerebrovascular statistical atlas described in Mouches and Forkert (2019). These ROIs were used in combination with the clusters extracted from the age-specific average maps to identify the ROIs in which the clusters are located.

### 3 | RESULTS

#### 3.1 | Brain age prediction

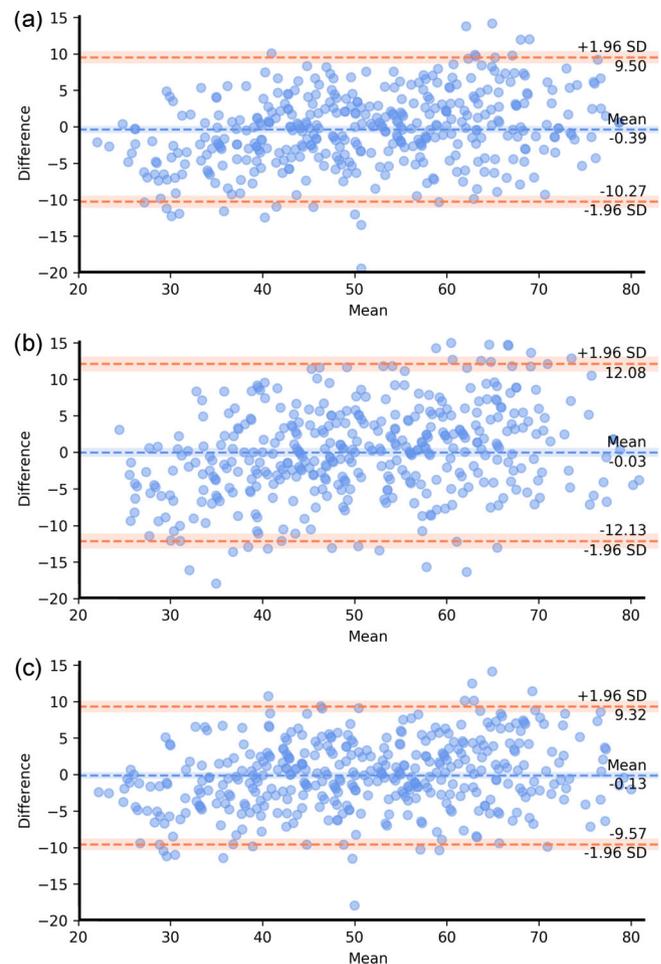
Table 1 shows the prediction accuracies of the  $\text{CNN}_{\text{T1}}$ ,  $\text{CNN}_{\text{TOF}}$ , and  $\text{CNN}_{\text{Combined}}$  biological brain age prediction models in terms of mean absolute error and  $R^2$  comparing the chronological and estimated biological age. The combination of both imaging modalities ( $\text{CNN}_{\text{Combined}}$ ) results in the best age prediction performance (mean absolute error [MAE] = 3.85 years,  $R^2 = 0.88$ ) and shows a significant absolute error improvement, using a 2-tailed paired  $t$  test, compared to the models using a single imaging modality (T1-weighted MRI:  $p < .05$ ; TOF MRA:  $p < .0001$ ). The model training and validation loss curves (Figure S1) indicate no overfitting (i.e., increased validation loss over the epochs) but a gap between the training and validation loss, which was also previously observed with the SFCN model architecture (Peng et al., 2021).

Figure 3 illustrates the performance of the different models. The Bland–Altman plot for the  $\text{CNN}_{\text{Combined}}$  model (Figure 3c) shows a tighter limit of agreement ( $-9.57; 9.32$ ) compared to the two other

**TABLE 1** Biological brain age prediction results of the different models computed for the 400 test datasets

Model input	Mean absolute error (SD)	$R^2$
T1-weighted MRI ( $\text{CNN}_{\text{T1}}$ )	4.01 (3.08)*	0.872
TOF MRA ( $\text{CNN}_{\text{TOF}}$ )	4.91 (3.75)**	0.805
All ( $\text{CNN}_{\text{Combined}}$ )	3.85 (2.90)	0.882

Note: Significant difference with the  $\text{CNN}_{\text{Combined}}$  model is indicated.  
\* $p < .05$ ; \*\* $p < .01$ .



**FIGURE 3** Bland–Altman plots comparing the participants chronological age and predicted age from the (a)  $\text{CNN}_{\text{T1}}$ , (b)  $\text{CNN}_{\text{TOF}}$ , and (c)  $\text{CNN}_{\text{Combined}}$  model. The x-axis shows the mean of the chronological and predicted age (in years) and the y-axis shows the difference (chronological age–predicted age)

plots (Figure 3a,b), again demonstrating the benefits of combining the two imaging modalities. This observation is further supported by the plots directly comparing the predicted and true chronological age, which are presented in Figure S2. All models show small mean differences between the chronological and predicted age, with the highest difference found for the  $\text{CNN}_{\text{T1}}$  model ( $-0.39$ ) indicating no systematic biases although the three models seem to slightly overpredict the age of young adults and underpredict the age of elderly, which has also been previously observed (de Lange & Cole, 2020; Le et al., 2018; Liang, Zhang, & Niu, 2019; Treder et al., 2021).

#### 3.2 | Saliency maps and regions of importance

Figure 4 shows the age-specific average saliency maps for each imaging modality of the combined model ( $\text{CNN}_{\text{Combined}}$ ). Furthermore, significant clusters detected within each map can be found in Figure S6. For the T1-weighted age-specific average saliency maps, Figures 5 and S7 show

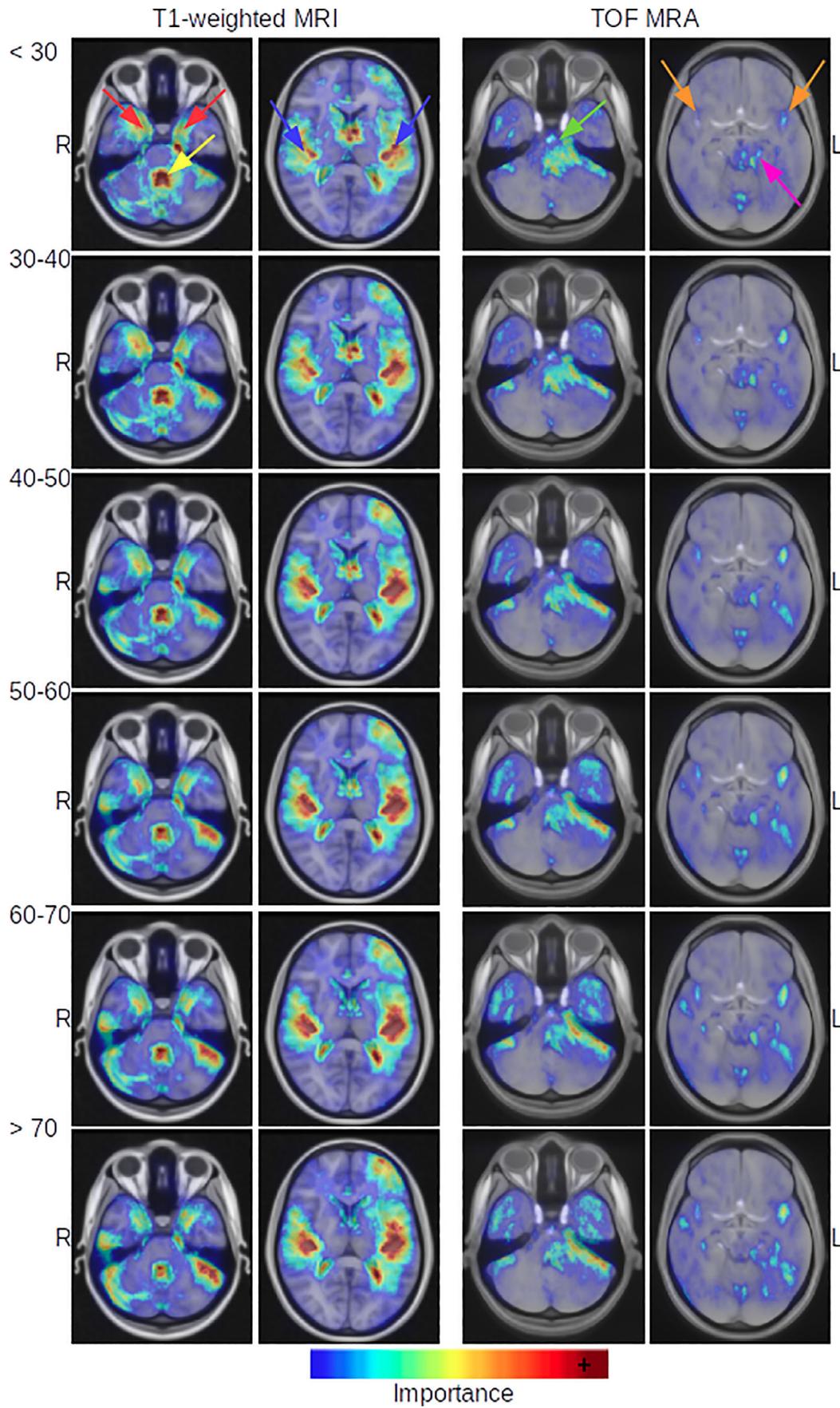


FIGURE 4 Legend on next page.

heatmaps illustrating the number of significant clusters overlapping with the different ROIs when using either the CerebrA or the finer human cerebral cortex atlases. The heatmaps only display the brain regions with at least one cluster of important voxels. Overall, three brain areas located around (a) the fourth ventricle, (b) the amygdala, and (c) the lateral sulci contained clusters of voxels with high importance. Figures 5 and S7 also show fewer ROIs containing clusters in older participants, as well as varying ROIs containing clusters of important voxels across the different age ranges. For instance, the cerebellar vermis and white matter are indicated in the saliency maps of young participants but not in those generated from older participants. This shift in focus of the CNN model is also visible in Figure 4 where the importance in these regions decreases with aging. For the brain arteries, no cluster was overlapping with any of the defined ROIs, and the few significant clusters were very small (see Figure S6). However, the age-specific average saliency maps (Figure 4) show important regions around the BA, MCA M2, and PCA, although the importance in the TOF MRA age-specific average saliency maps was overall lower than that in the T1-weighted age-specific average saliency maps.

Finally, the results of the two saliency map sanity checks that were performed are presented along with the raw saliency maps obtained using the proposed model in Figure S4 for randomly sampled participants at different ages. The first sanity check resulted in noisy maps with regions of high contrast being more important while the second sanity check resulted in random brain regions being important, thus suggesting that the saliency maps generated using the original data are informative and trustworthy. When averaging the saliency maps generated as part of the sanity checks per age range (Figure S5), the averaged maps show no differences between the age ranges and are very different from the averaged maps obtained with the proposed model (Figure 4), again supporting the trustworthiness of the age dependences observed with the proposed model.

## 4 | DISCUSSION

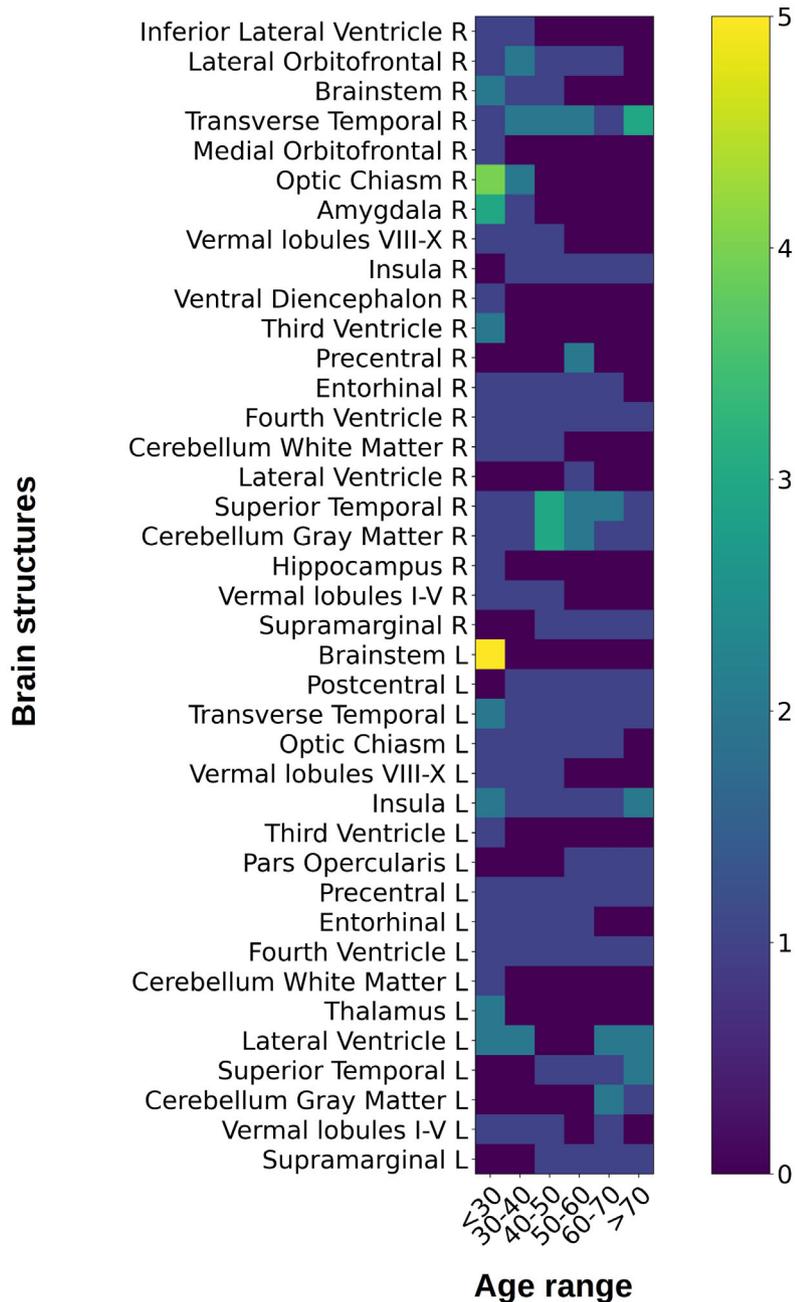
The primary aim of this work was to investigate the benefits of combining the information from cortical, subcortical, and cerebrovascular structures to solve the biological brain age prediction task by combining T1-weighted MRI and TOF MRA datasets. Therefore, the technical and clinical contributions of this work include (a) The validation of the state-of-the-art brain age prediction SFCN model (Peng et al., 2021) in a multimodal setting, and with different imaging modalities (originally implemented using T1-weighted MRI datasets only), which translated to an improved accuracy when using T1-weighted MRI and TOF MRA modalities together and (b) the use of saliency maps in a multimodal setting that eventually enable the identification of brain regions and arteries that are most strongly associated with aging.

### 4.1 | Brain age prediction

In this study, the biological brain age prediction model shows significantly improved results when combining T1-weighted MRI and TOF MRA datasets. The T1-weighted MRI scans, containing macrostructural brain tissue information, provide a better age prediction than the TOF MRA scans, in which the brain tissue contrast is low, but arteries are enhanced. Interestingly, despite the rather low anatomical brain tissue contrast, the TOF MRA scans alone allow predicting the brain age with a reasonable MAE of 4.9 years, which is slightly better compared to results described in literature where a MAE of 5.23 years was reported when using similar data (Nam et al., 2020). This result shows that the used model architecture is robust when applied for brain age prediction based on different imaging modalities although it was originally developed for T1-weighted MRI data (Peng et al., 2021). Moreover, it is important to note that the TOF MRA image acquisition does not include the full brain, but only the inferior part, where the major arteries are located (see Figure 1c), while the superior part (above the lateral ventricles) is typically missing in the datasets due to the field of view of the acquisition. Thus, the fact that the TOF MRA datasets include less information than the T1-weighted MRI datasets might partly explain the difference in accuracy between the  $CNN_{T1}$  and  $CNN_{TOF}$  models. When combining both modalities, the resulting mean absolute error significantly improves and is comparable to the results described in literature, with reported MAEs mostly varying between 3 to 5 years when using T1-weighted MRI datasets (Bashyam et al., 2020; Cole et al., 2017; Jonsson et al., 2019; Levakov et al., 2020; Mouches, Wilms, Rajashekar, Langner, & Forkert, 2021; Peng et al., 2021; Wilms et al., 2020), but often using considerably more training data, and participants with a much narrower age range, which hinders a direct comparison of the results (de Lange et al., 2021). Moreover, Bashyam et al. (2020) who previously trained a deep learning brain age prediction model using more than 11,000 datasets and tested it on the SHIP database reported a MAE of 4.12 years, showing that this database is rather challenging for the brain age prediction task. Recent studies attempted to explain the observed difference between the chronological and predicted brain age by using clinical and life behavioral factors differences between participants. For example, previous studies found that the difference between the chronological and predicted brain age is correlated with blood pressure, smoking status, and alcohol consumption (Cole, 2020; de Lange et al., 2020). Although this kind of analysis using TOF MRA data has not been described in the literature yet, such correlations can be expected as brain artery measurements extracted from TOF MRA data were previously shown to be associated with several of these factors (Mouches et al., 2021). However, this hypothesis needs to be investigated in future work.

**FIGURE 4** Age-specific average T1-weighted MRI and TOF MRA saliency maps overlaid onto the MNI T1-weighted brain template and the cerebrovascular statistical atlas described in Mouches and Forkert (2019), respectively. T1-weighted MRI arrows: blue, lateral sulcus; red, amygdala/entorhinal cortex/optic chiasm; yellow, fourth ventricle. TOF MRA arrows: orange, middle cerebral artery (MCA) M2 segment; pink, left posterior cerebral artery (PCA); green, basilar artery (BA)

### Number of clusters overlapping with each brain region, for each age-specific averaged saliency map



**FIGURE 5** Heatmap of the number of clusters overlapping with each brain region, as described in the CerebrA atlas (Manera et al., 2020), for each age-specific T1-weighted MRI average saliency map. Only brain regions with at least one cluster of important voxels are displayed

The observed age prediction improvement when using multiple modalities is also in line with previous results from multimodal brain age prediction studies. For example, in Cole (2020), hand-crafted features from six different modalities were used, whereas combining them resulted in a substantially improved MAE (T1-weighted only: 4.14, all modalities: 3.52). Similarly, Rokicki et al. (2020) found that combining hand-crafted features from T1- and T2-weighted MRI and cerebral blood flow information from arterial spin labeling imaging led to significantly improved brain age predictions. Finally, Jonsson et al. (2019) used a single modality, T1-weighted MRI, but generated four types of images from it and trained four 3D CNN models on each

of them. Combining the output of all the CNNs resulted in the best accuracy, supporting the relevance of ensembling multiple CNN results.

#### 4.2 | Brain structures importance

The generated saliency maps indicate brain regions that are important for the model for the brain age prediction task. These regions bring complementary and nonredundant information from both imaging sequences allowing the CNN<sub>Combined</sub> model to perform better than

the single modality models. Thus, the regions identified can be assumed to be impacted by aging in a unique and informative way.

In the T1-weighted datasets, three brain areas were primarily used by the model to make the prediction: the lateral sulcus (transverse temporal and insula), the fourth ventricle, and the medial temporal lobe (amygdala and entorhinal cortex, expanding to the optic chiasm). The lateral sulcus was previously identified as one of the regions most affected by cerebrospinal fluid volume increase in elderly (Lemaître et al., 2005). This finding is in line with the higher importance of the voxels around this area observed in the age-specific average saliency maps of older adults. This observed increased importance could also be associated with the accelerated reduction of cortical thickness after the age of 60 years, which is widely reported in the brain aging community (Potvin et al., 2017). The fourth ventricle previously showed low volume change with age (Fjell et al., 2009), but was identified as a region of interest by the model. The saliency maps also highlight surrounding structures, including the cerebellum, which has been previously identified displaying an accelerated volume change with age (Han, An, Carass, Prince, & Resnick, 2020). Interestingly, the lateral sulcus and fourth ventricles were also identified as important in Levakov et al. (2020) where a different brain age prediction model architecture and a population with a wider age range (4–94 years old) were used. Finally, the area around the amygdala was also identified as important, whereas this structure was previously found to have a decreasing volume with aging (Potvin et al., 2016).

For the TOF MRA datasets as an input, the important regions identified in the average age-specific saliency maps are sparser, possibly due to the intersubject variability in vessel morphology, and only very small clusters were identified as significant. This finding is to be expected given the highly variable vascular system consisting of small arteries, rather than large brain regions, that can differ quite a bit even between healthy subjects (Mouches & Forkert, 2019). Overall, the cisterns, spaces filled with cerebrospinal fluid, seem to be the most important regions for the model, with the arteries located within the cisterns (the MCA M2 perforating the lateral sulcus and the PCA in the ambient and quadrigeminal cistern) showing the highest importance among the arteries in the age-specific average saliency maps. Additionally, the region around the BA also shows some importance, although none of the significant clusters were found in these areas. Indeed, these arteries were previously reported to be affected by aging, with age having a significant effect on their diameter and on the vessel number in their associated flow territories (Bullitt et al., 2010; Mouches, Langner, et al., 2021). In Mouches, Langner, et al. (2021), the SHIP database was also used to extract artery-related measurements from TOF MRA datasets. The reported associations between these measurements and age are generally in line with the findings of this study. For example, the arteries in the PCA blood flow territory were found to be negatively impacted by age whereas the saliency maps generated in this work also identified predictive areas in this region for brain age prediction. Additionally, the thickness of the BA and the MCA M2 segments was found to decrease with age in Mouches, Langner, et al. (2021), whereas these arteries were also highlighted as important regions in the saliency maps generated in this

work. Based on the age-specific average saliency maps, it appears that the model uses more information from the MCA M2 segment to predict the age of older participants. This finding could be indicative of greater changes in these arteries in older participants, and in the cisterns in which the arteries lie, as this finding was also observed in the lateral sulcus region in T1-weighted MRI saliency maps.

Moreover, it is important to note that many previously published brain aging studies used brain measurements that were derived using advanced image processing methods, such as regional volume or thickness measurements, to identify the regional morphological changes associated with aging. In contrast, CNN models can extract novel, optimal features by combining texture and shape information and, thus, may reveal regions of interest not identified in previous works. Moreover, in CNNs, voxel-wise importance can be directly assessed by retrieving the spatial localization of the important features in the input space, as opposed to methods using hand-crafted features in which the importance is assessed in the feature space. Therefore, future work is needed to investigate the similarities and differences in brain regions identified as important when predicting brain age using features extracted by a CNN versus hand-crafted features.

### 4.3 | Limitations

The results of this study show that the proposed approach is generalizable, could easily be extended to more imaging modalities, and requires minimal preprocessing steps. The results also show that the presented approach leads to robust and clinically plausible results when applied to modalities other than T1-weighted MRI, for which it was originally developed (Peng et al., 2021). However, some limitations remain. First, while using a single database increases the consistency of the results and reduces biases, leading to more robust explanations, it also results in a model that is less robust to varying scanning parameters, and limits the amount of data available. Therefore, the model prediction accuracy would benefit from training using a larger sample size, as previously demonstrated in the context of brain age prediction (de Lange et al., 2021), and data collected from different centers, especially when using deep learning models, which are known to be data hungry (Marcus, 2018). Nevertheless, based on the excellent results of the SFCN architecture on the highly diverse PAC2019 brain age prediction data reported in Peng et al. (2021), we assume that the general findings of this study will hold true even for multicenter datasets, especially when proper harmonization strategies are implemented to remove possible confounding biases. However, this assumption should be experimentally confirmed in future work. Second, although the image preprocessing steps are minimal compared to, for instance, the steps necessary to extract conventional morphological or artery features, the effect of each preprocessing step on the prediction accuracy was not assessed. While the preprocessing steps used in this work follow common practice, their effect on the brain age prediction accuracy should be investigated in future work. Third, a simple linear regression model was used to

combine both single modalities model outputs. However, more complex nonlinear fusion methods could be used, but might require more data to be trained. Finally, the saliency maps provide some insight of brain regions affected by aging. However, areas providing redundant information might have been ignored by the model. Therefore, the regions identified as important constitute a subset of regions allowing to accurately predict the biological brain age, but regions not highlighted on the saliency maps cannot be inferred as nonimportant or not associated with age. These maps also do not provide information about the directionality of the relationship between the brain regions and the age, which could be improved in future work by using different explainability methods such as layer-wise relevance propagation (Bach et al., 2015) or Shapley additive explanations (Lundberg & Lee, 2017).

## 5 | CONCLUSION

This study provides a technical contribution by applying and validating an existing state-of-the-art model architecture originally developed for the analysis of a single modality image in a multimodal setting, as well as using saliency maps in the same context. From a medical and neuroscience perspective, the proposed framework allows identifying brain regions and arteries contributing the most to the model prediction. The results show that combining brain tissue and artery information significantly improves the brain age prediction, whereas the lateral sulcus, fourth ventricle, and medial temporal lobe brain regions were identified as especially important morphological features. The artery-related information showed an overall lower contribution but still improved the brain age prediction. Overall, it is important to highlight that the proposed approach can be easily extended to other imaging modalities to improve our understanding of their relative contribution to biological brain age prediction.

## ACKNOWLEDGMENTS

This work was supported by the Canada Research Chairs program, the River Fund at Calgary Foundation, the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Canadian Open Neuroscience Platform (CONP).

## CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the Study of Health in Pomerania study upon reasonable request (<https://www2.medizin.uni-greifswald.de/cm/fv/ship/>). The code used in this study is available at: <https://github.com/pmouches/Multi-modal-biological-brain-age-prediction>.

## ETHICS STATEMENT

All participants provided written informed consent and the Study of Health in Pomerania study was approved by the local ethics

commission of the University of Greifswald (BB 39/08, 19.06.2008). The scans were completely anonymized for this secondary study so that no additional ethics approval was required.

## ORCID

Pauline Mouches  <https://orcid.org/0000-0002-0304-5584>

Matthias Wilms  <https://orcid.org/0000-0001-8845-360X>

Deepthi Rajashekar  <https://orcid.org/0000-0003-2009-764X>

Nils D. Forkert  <https://orcid.org/0000-0003-2556-3224>

## REFERENCES

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... Zheng, X. (2016). TensorFlow: A system for large-scale machine learning. In *12th USENIX symposium on operating systems design and implementation (OSDI 16)*, 265–283.
- Adebayo, J., Gilmer, J., Muelly, M., Goodfellow, I., Hardt, M. & Kim, B. (2018). Sanity checks for saliency maps. ArXiv:1810.03292. Retrieved from <http://arxiv.org/abs/1810.03292>
- Amin-Hanjani, S., Du, X., Pandey, D. K., Thulborn, K. R., & Charbel, F. T. (2015). Effect of age and vascular anatomy on blood flow in major cerebral vessels. *Journal of Cerebral Blood Flow & Metabolism*, 35(2), 312–318. <https://doi.org/10.1038/jcbfm.2014.203>
- Arun, N., Gaw, N., Singh, P., Chang, K., Aggarwal, M., Chen, B., ... Kalpathy-Cramer, J. (2021). Assessing the trustworthiness of saliency maps for localizing abnormalities in medical imaging. *Radiology: Artificial Intelligence*, 3, 6. <https://doi.org/10.1148/ryai.2021200267>
- Avants, B. B., Tustison, N., & Johnson, H. (2009). Advanced normalization tools (ANTS). *Insight Journal*, 2(365), 1–25.
- Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K.-R., & Samek, W. (2015). On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLoS One*, 10(7), e0130140. <https://doi.org/10.1371/journal.pone.0130140>
- Bashyam, V. M., Erus, G., Doshi, J., Habes, M., Nasrallah, I. M., Truelove-Hill, M., ... Davatzikos, C. (2020). MRI signatures of brain age and disease over the lifespan based on a deep brain network and 14 468 individuals worldwide. *Brain*, 143(7), 2312–2324. <https://doi.org/10.1093/brain/awaa160>
- Bullitt, E., Zeng, D., Mortamet, B., Ghosh, A., Aylward, S. R., Lin, W., ... Smith, K. (2010). The effects of healthy aging on intracerebral blood vessels visualized by magnetic resonance angiography. *Neurobiology of Aging*, 31(2), 290–300. <https://doi.org/10.1016/j.neurobiolaging.2008.03.022>
- Cole, J. H. (2020). Multimodality neuroimaging brain-age in UK biobank: Relationship to biomedical, lifestyle, and cognitive factors. *Neurobiology of Aging*, 92, 34–42. <https://doi.org/10.1016/j.neurobiolaging.2020.03.014>
- Cole, J. H., & Franke, K. (2017). Predicting age using neuroimaging: Innovative brain ageing biomarkers. *Trends in Neurosciences*, 40(12), 681–690. <https://doi.org/10.1016/j.tins.2017.10.001>
- Cole, J. H., Poudel, R. P. K., Tsagkrasoulis, D., Caan, M. W. A., Steves, C., Spector, T. D., & Montana, G. (2017). Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *NeuroImage*, 163, 115–124. <https://doi.org/10.1016/j.neuroimage.2017.07.059>
- de Lange, A.-M. G., Anatórk, M., Rokicki, J., Han, L. K. M., Franke, K., Alnæs, D., ... Cole, J. H. (2021). Mind the gap: Performance metric evaluation in brain-age prediction. *BioRxiv*. <https://doi.org/10.1101/2021.05.16.444349>
- de Lange, A.-M. G., Anatórk, M., Suri, S., Kaufmann, T., Cole, J. H., Griffanti, L., ... Ebmeier, K. P. (2020). Multimodal brain-age prediction and cardiovascular risk: The Whitehall II MRI sub-study. *NeuroImage*, 222, 117292. <https://doi.org/10.1016/j.neuroimage.2020.117292>

- de Lange, A.-M. G., & Cole, J. H. (2020). Commentary: Correction procedures in brain-age prediction. *NeuroImage: Clinical*, 26, 102229. <https://doi.org/10.1016/j.nicl.2020.102229>
- Dinsdale, N. K., Bluemke, E., Smith, S. M., Arya, Z., Vidaurre, D., Jenkinson, M., & Namburete, A. I. L. (2021). Learning patterns of the ageing brain in MRI using deep convolutional networks. *NeuroImage*, 224, 117401. <https://doi.org/10.1016/j.neuroimage.2020.117401>
- Dinsdale, N. K., Jenkinson, M., & Namburete, A. I. L. (2021). Deep learning-based unlearning of dataset bias for MRI harmonisation and confound removal. *NeuroImage*, 228, 117689. <https://doi.org/10.1016/j.neuroimage.2020.117689>
- Fjell, A. M., Westlye, L. T., Amlie, I., Espeseth, T., Reinvang, I., Raz, N., ... Walhovd, K. B. (2009). Minute effects of sex on the aging brain: A multisample magnetic resonance imaging study of healthy aging and Alzheimer's disease. *Journal of Neuroscience*, 29(27), 8774–8783. <https://doi.org/10.1523/JNEUROSCI.0115-09.2009>
- Forkert, N. D., Schmidt-Richberg, A., Fiehler, J., Illies, T., Möller, D., Säring, D., ... Ehrhardt, J. (2013). 3D cerebrovascular segmentation combining fuzzy vessel enhancement and level-sets with anisotropic energy weights. *Magnetic Resonance Imaging*, 31(2), 262–271. <https://doi.org/10.1016/j.mri.2012.07.008>
- Fotinos, A. F., Snyder, A. Z., Gorton, L. E., Morris, J. C., & Buckner, R. L. (2005). Normative estimates of cross-sectional and longitudinal brain volume decline in aging and AD. *Neurology*, 64(6), 1032–1039. <https://doi.org/10.1212/01.WNL.0000154530.72969.11>
- Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., ... Van Essen, D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, 536(7615), 171–178. <https://doi.org/10.1038/nature18933>
- Han, S., An, Y., Carass, A., Prince, J. L., & Resnick, S. M. (2020). Longitudinal analysis of regional cerebellum volumes during normal aging. *NeuroImage*, 220, 117062. <https://doi.org/10.1016/j.neuroimage.2020.117062>
- Isensee, F., Schell, M., Pflueger, I., Brugnara, G., Bonekamp, D., Neuberger, U., ... Kickingereder, P. (2019). Automated brain extraction of multisequence MRI using artificial neural networks. *Human Brain Mapping*, 40(17), 4952–4964. <https://doi.org/10.1002/hbm.24750>
- Jin, W., Li, X., & Hamarneh, G. (2021). One map does not fit all: Evaluating saliency map explanation on multi-modal medical images. *ArXiv*. <http://arxiv.org/abs/2107.05047>
- Jonsson, B. A., Bjornsdottir, G., Thorgeirsson, T. E., Ellingsen, L. M., Walters, G. B., Gudbjartsson, D. F., ... Ulfarsson, M. O. (2019). Brain age prediction using deep learning uncovers associated sequence variants. *Nature Communications*, 10(1), 1–10. <https://doi.org/10.1038/s41467-019-13163-9>
- Kingma, D. P., & Ba, J. (2017). Adam: A method for stochastic optimization. *ArXiv*. <http://arxiv.org/abs/1412.6980>
- Klein, A., & Tourville, J. (2012). 101 labeled brain images and a consistent human cortical labeling protocol. *Frontiers in Neuroscience*, 6, 171. <https://doi.org/10.3389/fnins.2012.00171>
- Kruggel, F., Turner, J., & Muftuler, L. T. (2010). Impact of scanner hardware and imaging protocol on image quality and compartment volume precision in the ADNI cohort. *NeuroImage*, 49(3), 2123–2133. <https://doi.org/10.1016/j.neuroimage.2009.11.006>
- Le, T. T., Kuplicki, R. T., McKinney, B. A., Yeh, H.-W., Thompson, W. K., Paulus, M. P., ... Victor, T. A. (2018). A nonlinear simulation framework supports adjusting for age when analyzing BrainAGE. *Frontiers in Aging Neuroscience*, 10, 317. <https://doi.org/10.3389/fnagi.2018.00317>
- Lemaître, H., Crivello, F., Grasiot, B., Alperovitch, A., Tzourio, C., & Mazoyer, B. (2005). Age- and sex-related effects on the neuroanatomy of healthy elderly. *NeuroImage*, 26(3), 900–911. <https://doi.org/10.1016/j.neuroimage.2005.02.042>
- Levakov, G., Rosenthal, G., Shelef, I., Raviv, T. R., & Avidan, G. (2020). From a deep learning model back to the brain—Identifying regional predictors and their relation to aging. *Human Brain Mapping*, 41(12), 3235–3252. <https://doi.org/10.1002/hbm.25011>
- Liang, H., Zhang, F., & Niu, X. (2019). Investigating systematic bias in brain age estimation with application to post-traumatic stress disorders. *Human Brain Mapping*, 40(11), 3143–3152. <https://doi.org/10.1002/hbm.24588>
- Lo Vercio, L., Amador, K., Bannister, J. J., Crites, S., Gutierrez, A., MacDonald, M. E., ... Forkert, N. D. (2020). Supervised machine learning tools: A tutorial for clinicians. *Journal of Neural Engineering*, 17(6), 062001. <https://doi.org/10.1088/1741-2552/abbff2>
- Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan & R. Garnett (Eds.), *Advances in Neural Information Processing Systems 30* (pp. 4765–4774). Red Hook, NY: Curran Associates, Inc.
- MacDonald, M. E., Williams, R. J., Rajashekar, D., Stafford, R. B., Hanganu, A., Sun, H., ... Pike, G. B. (2020). Age-related differences in cerebral blood flow and cortical thickness with an application to age prediction. *Neurobiology of Aging*, 95, 131–142. <https://doi.org/10.1016/j.neurobiolaging.2020.06.019>
- Manera, A. L., Dadar, M., Fonov, V., & Collins, D. L. (2020). Cerebra, registration and manual label correction of Mindboggle-101 atlas for MNI-ICBM152 template. *Scientific Data*, 7(1), 237. <https://doi.org/10.1038/s41597-020-0557-9>
- Marcus, G. (2018). Deep learning: A critical appraisal. *ArXiv*. <http://arxiv.org/abs/1801.00631>
- Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., ... Mazoyer, B. (2001). A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 356(1412), 1293–1322. <https://doi.org/10.1098/rstb.2001.0915>
- Mouches, P., & Forkert, N. D. (2019). A statistical atlas of cerebral arteries generated using multi-center MRA datasets from healthy subjects. *Scientific Data*, 6(1), 1–8. <https://doi.org/10.1038/s41597-019-0034-5>
- Mouches, P., Langner, S., Domin, M., Hill, M. D., & Forkert, N. D. (2021). Influence of cardiovascular risk-factors on morphological changes of cerebral arteries in healthy adults across the life span. *Scientific Reports*, 11(1), 12236. <https://doi.org/10.1038/s41598-021-91669-3>
- Mouches, P., Wilms, M., Rajashekar, D., Langner, S., & Forkert, N. (2021). Unifying brain age prediction and age-conditioned template generation with a deterministic autoencoder. *Medical Imaging with Deep Learning*, 143, 497–506.
- Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In *ICML'10: Proceedings of the 27th International Conference on International Conference on Machine Learning* (pp. 807–814). Madison, WI: Omnipress.
- Nam, Y., Jang, J., Lee, H. Y., Choi, Y., Shin, N. Y., Ryu, K.-H., ... Kim, B. (2020). Estimating age-related changes in in vivo cerebral magnetic resonance angiography using convolutional neural network. *Neurobiology of Aging*, 87, 125–131. <https://doi.org/10.1016/j.neurobiolaging.2019.12.008>
- Niu, X., Zhang, F., Kounios, J., & Liang, H. (2020). Improved prediction of brain age using multimodal neuroimaging data. *Human Brain Mapping*, 41(6), 1626–1643. <https://doi.org/10.1002/hbm.24899>
- O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *ArXiv preprint*. <https://arxiv.org/abs/1511.08458>
- Peng, H., Gong, W., Beckmann, C. F., Vedaldi, A., & Smith, S. M. (2021). Accurate brain age prediction with lightweight deep neural networks. *Medical Image Analysis*, 68, 101871. <https://doi.org/10.1016/j.media.2020.101871>
- Peters, R. (2006). Ageing and the brain. *Postgraduate Medical Journal*, 82(964), 84–88. <https://doi.org/10.1136/pgmj.2005.036665>
- Pomponio, R., Erus, G., Habes, M., Doshi, J., Srinivasan, D., Mamourian, E., ... Davatzikos, C. (2020). Harmonization of large

- MRI datasets for the analysis of brain imaging patterns throughout the lifespan. *NeuroImage*, 208, 116450. <https://doi.org/10.1016/j.neuroimage.2019.116450>
- Potvin, O., Dieumegarde, L., & Duchesne, S. (2017). Normative morphometric data for cerebral cortical areas over the lifetime of the adult human brain. *NeuroImage*, 156, 315–339. <https://doi.org/10.1016/j.neuroimage.2017.05.019>
- Potvin, O., Mouiha, A., Dieumegarde, L., Duchesne, S., & Alzheimer's Disease Neuroimaging Initiative. (2016). Normative data for subcortical regional volumes over the lifetime of the adult human brain. *NeuroImage*, 137, 9–20. <https://doi.org/10.1016/j.neuroimage.2016.05.016>
- Ritz, K., Denswil, N. P., Stam, O. C. G., van Lieshout, J. J., & Daemen, M. J. A. P. (2014). Cause and mechanisms of intracranial atherosclerosis. *Circulation*, 130(16), 1407–1414. <https://doi.org/10.1161/CIRCULATIONAHA.114.011147>
- Rokicki, J., Wolfers, T., Nordhøy, W., Tesli, N., Quintana, D. S., Alnæs, D., ... Westlye, L. T. (2020). Multimodal imaging improves brain age prediction and reveals distinct abnormalities in patients with psychiatric and neurological disorders. *Human Brain Mapping*, 42(6), 1714–1726. <https://doi.org/10.1002/hbm.25323>
- Samek, W., Montavon, G., Lapuschkin, S., Anders, C. J., & Müller, K.-R. (2021). Explaining deep neural networks and beyond: A review of methods and applications. *Proceedings of the IEEE*, 109(3), 247–278. <https://doi.org/10.1109/JPROC.2021.3060483>
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE international conference on computer vision*, 618–626. <https://doi.org/10.1109/ICCV.2017.74>
- Simonyan, K., Vedaldi, A. & Zisserman, A. (2013). Deep inside convolutional networks: Visualising image classification models and saliency maps. *ArXiv*. <https://arxiv.org/abs/1312.6034v2>
- Smilkov, D., Thorat, N., Kim, B., Viégas, F. & Wattenberg, M. (2017). SmoothGrad: Removing noise by adding noise. *ArXiv*. <http://arxiv.org/abs/1706.03825>
- Spisák, T., Spisák, Z., Zunhammer, M., Bingel, U., Smith, S., Nichols, T., & Kincses, T. (2019). Probabilistic TFCE: A generalized combination of cluster size and voxel intensity to increase statistical power. *NeuroImage*, 185, 12–26. <https://doi.org/10.1016/j.neuroimage.2018.09.078>
- Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., ... Collins, R. (2015). UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Medicine*, 12(3), e1001779. <https://doi.org/10.1371/journal.pmed.1001779>
- Treder, M. S., Shock, J. P., Stein, D. J., du Plessis, S., Seedat, S., & Tsvetanov, K. A. (2021). Correlation constraints for regression models: Controlling bias in brain age prediction. *Frontiers in Psychiatry*, 12, 25. <https://doi.org/10.3389/fpsy.2021.615754>
- Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 bias correction. *IEEE Transactions on Medical Imaging*, 29(6), 1310–1320. <https://doi.org/10.1109/TMI.2010.2046908>
- Valizadeh, S. A., Hänggi, J., Mérillat, S., & Jäncke, L. (2017). Age prediction on the basis of brain anatomical measures. *Human Brain Mapping*, 38(2), 997–1008. <https://doi.org/10.1002/hbm.23434>
- Völzke, H., Alte, D., Schmidt, C. O., Radke, D., Lorbeer, R., Friedrich, N., ... Hoffmann, W. (2011). Cohort profile: The study of health in Pomerania. *International Journal of Epidemiology*, 40(2), 294–307. <https://doi.org/10.1093/ije/dyp394>
- Wardlaw, J. M., Smith, E. E., Biessels, G. J., Cordonnier, C., Fazekas, F., Frayne, R., ... STandards for Reporting Vascular changes on nEuroimaging (STRIVE v1). (2013). Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *Lancet. Neurology*, 12(8), 822–838. [https://doi.org/10.1016/S1474-4422\(13\)70124-8](https://doi.org/10.1016/S1474-4422(13)70124-8)
- Wilms, M., Bannister, J. J., Mouches, P., MacDonald, M. E., Rajashekar, D., Langner, S., & Forkert, N. D. (2020). Bidirectional modeling and analysis of brain aging with normalizing flows. In S. M. Kia, H. Mohy-ud-Din, A. Abdulkadir, C. Bass, M. Habes, J. M. Rondina et al. (Eds.), *Machine learning in clinical neuroimaging and radiogenomics in neuro-oncology* (Vol. 12449, pp. 22–33). Cham, Switzerland: Springer. [https://doi.org/10.1007/978-3-030-66843-3\\_3](https://doi.org/10.1007/978-3-030-66843-3_3)
- Wrobel, J., Martin, M. L., Bakshi, R., Calabresi, P. A., Elliot, M., Roalf, D., ... Goldsmith, J. (2020). Intensity warping for multisite MRI harmonization. *NeuroImage*, 223, 117242. <https://doi.org/10.1016/j.neuroimage.2020.117242>

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Mouches, P., Wilms, M., Rajashekar, D., Langner, S., & Forkert, N. D. (2022). Multimodal biological brain age prediction using magnetic resonance imaging and angiography with the identification of predictive regions. *Human Brain Mapping*, 43(8), 2554–2566. <https://doi.org/10.1002/hbm.25805>