bioRxiv preprint doi: https://doi.org/10.1101/2021.05.10.443469; this version posted May 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Decentralized Brain Age Estimation using MRI Data

Sunitha Basodi[1], Rajikha Raja[1,2], Bhaskar Ray[1,3], Harshvardhan Gazula[4], Jingyu Liu[1,3], Eric Verner[1], and Vince D. Calhoun[1,3,5]

[1]Tri-institutional Center for Translational Research in Neuroimaging and Data Science, Georgia State University, Georgia Institute of Technology, Emory University, Atlanta, GA, USA
[2]Department of Radiology, University of Arkansas for Medical Sciences, Little Rock, AR, USA
[3]Department of Computer Science, Georgia State University, GA, USA
[4]Princeton Neuroscience Institute, Princeton, NJ, USA
[5]Department of Psychology, Georgia State University, GA, USA

Abstract

Recent studies have demonstrated that neuroimaging data can be used to predict brain age, as it captures information about the neuroanatomical and functional changes the brain undergoes during development and the aging process. However, researchers often have limited access to neuroimaging data because of its challenging and expensive acquisition process, thereby limiting the effectiveness of the predictive model. Decentralized models provide a way to build more accurate and generalizable prediction models, bypassing the traditional data-sharing methodology. In this work, we propose a decentralized method for brain age estimation and evaluate it on three different feature sets, including both volumetric and voxelwise structural MRI data as well as resting functional MRI data. The results demonstrate that a decentralized brain age model can achieve similar performance compared to the models trained with all the data in one location.

Keywords: Brain Age, Decentralized, Federated, COINSTAC

1 Introduction

Brain age estimation (BAE) from magnetic resonance imaging (MRI) data has become widely popular in recent years. Computed as the difference between chronological age and estimated brain age, the brain age gap is helpful in early identification of various neurodegenerative diseases [1, 2, 3] and was also found to be wider in patients with dementia and autism [4]. Several studies have used traditional machine learning [5, 6] and deep learning approaches [7] to develop promising prediction models on large datasets that have either been collected at one location or sourced from different locations. Although having a large training dataset can help in achieving robust models, there are limitations on the amount of data that can be gathered for such large-scale analyses. Examples of such limitations include the cost-prohibitiveness of MRI data collection, varying institutional data-sharing policies, constrained data-usage agreements, data-privacy concerns, and many more [8].

One efficient approach to bypass these limitations is to use decentralized algorithms, which do not require pooling the data in one location. Decentralized learning algorithms have become popular in recent years due to the demand for collaborative studies [8, 9]. Such approaches are particularly important when there is a need to perform a large-N analysis involving heterogeneous datasets without worrying about data transmission or violating privacy.

In this work, we have developed a decentralized brain age prediction algorithm using support vector regression and performed detailed experiments on three different feature sets (types of datasets) using varying sampling methods and demonstrate the robustness of our approach. We implement our approach within the Collaborative Informatics and Neuroimaging Suite Toolkit for Anonymous Computation (COINSTAC) [9] framework (more on this later). Results show that the proposed decentralized method attained performance on par with a centralized approach.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.10.443469; this version posted May 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The remainder of this paper is organized as follows. The background on brain age, current prediction approaches, decentralized learning, and COINSTAC are presented in section 2. The decentralized BAE method is presented in section 3, followed by a discussion of the results in section 4 and conclusion in section 5.

2 Background

2.1 Brain age

The brain age (of an individual) is the observable age of the brain in contrast to the chronological (actual) age. As brain age cannot be directly measured, predictive models are typically trained with the chronological age of cognitively normal subjects. Any discrepancy between the estimated brain age and chronological age is considered to be the brain age gap. Studies have shown that patients suffering from neurodegenerative diseases (such as Alzheimer's, Huntington's, Parkinson's, and Multiple Sclerosis) have an increased brain age gap or brain-predicted age [4]. In contrast, there are studies that show activities such as meditation [10] and physical exercise decrease brain age [11]. These and other studies [12, 13, 2] suggest BAE may serve as an important biomarker in understanding the aging process of the brain and its relationship to brain health and disorder.

2.2 Current prediction approaches

Predicting brain age on the basis of MRI is now a widely used approach. There are broadly three different approaches that have been previously used, namely, voxel-based, region-based, and surface-based approaches for structural MRI[14]. In a voxel-based approach, MR images are first segmented into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF), and then this voxel-level information is used to fit the models. In a region-based approach, an atlas is used to summarize the data for estimating prediction models and brain age. Surface-based methods generally compute a triangulated mesh using WM/GM boundaries or GM/CSF boundaries and then extract surface-based features for prediction. In this work, we also introduce a component-based approach for functional MRI (fMRI) data that uses the timecourses from independent components or rather their cross-correlations, called functional network connectivity (FNC), as features [15]. The different features extracted from each of these approaches can be used with either traditional machine learning or deep learning models. Deep learning models have shown promising results in BAE. In [12], BAE was computed using both raw MR images and preprocessed MR images, and both showed consistent results. The authors of [16] employ single and multimodal brain imaging data, including structural MRI (sMRI), diffusion tensor imaging (DTI), and resting state fMRI (rs-fMRI), and evaluate performance on many models. Recently, graph neural networks have also been used for BAE [17] using data from the UK Biobank [18]. However, there is no clear understanding about which models perform the best; though it is clear that having large amounts of training data usually helps in achieving robust models.

2.3 Decentralized learning

Decentralized learning has recently gained attention as it allows privacy-preserving, large-scale, machine learning analysis [19]. Also referred to as federated learning in literature, decentralized machine learning is focused on training models on data that are not shared but that researchers want to be analyzed together. In this method, all the participating sites start with the same model, i.e., a local model is trained on its own data. In each iteration, gradients from local models are sent back to the main site for aggregation, after which updated gradients are returned to local sites to update their models. This is repeated over many iterations until the model achieves its stopping criteria, such as the desired performance. There are several challenges in designing decentralized algorithms, including heterogeneity across different sites, transmission, maintaining synchronization during training iterations, and preserving data privacy. Heterogeneity arises from the different devices (hardware and software) involved during the training process, and one cannot overlook this issue, as it can effect the overall

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.10.443469; this version posted May 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

performance of a decentralized model. Though there is no raw data exchange, decentralized learning involves transmission of model gradients/parameters that utilize significant bandwidth, especially when training deep learning models. The main site also needs to ensure that the aggregation is performed on the correct iteration data from all the sites when there is repeated communication between the local and main site. For a detailed summary of the machine learning algorithms, challenges, and architectures for decentralization, we refer the readers to [20, 21].

2.4 COINSTAC

COINSTAC [8, 9, 22] is a platform that enables decentralized analysis of (neuroimaging) data without the need to pool the data at one location. It was created to enable collaborative research by removing the barriers to traditional data-centric approaches. As there is no pooling of data, COINSTAC protects the privacy of individual datasets and solves the problem of researchers wanting to collaborate but being unable to because of data sharing restrictions [23]. Gazula et. al [24] used this application to perform a decentralized regression on sMRI data preprocessed using voxel-based morphometry to analyze the structural changes in the brain as linked to age, body mass index, and smoking. This study also emphasized the benefits of large-scale neuroimaging analysis. COINSTAC implements a wide and growing range of decentralized neuroimaging pipelines and supports such large-scale analysis of decentralized data with results on par with results from pooled data.

3 Methods

Though decentralized learning has been applied in some domains, this paper, to our knowledge, presents the first approach for decentralized brain age analysis. In this work, we perform brain age prediction using a decentralized approach. Let us assume there are $N+1$ participating sites, each gathering data from a different set of participants.

In centralized models, data from all the local sites is combined at a central location, where it is used for training a model, and therefore a centralized model has better performance than any model trained only on a subset of that data. In decentralized studies, instead of transferring the original data, only the information needed to train a model is shared, which not only improves the prediction performance of the model but also keeps the data secure at the local sites. Information sent from local sites is collected at the main site to build an aggregated model. The challenge is to reduce the performance gap between decentralized and centralized models without sharing the original data. Such a decentralized training approach can be used to train any prediction model; however, the type of information transferred between local and main sites highly depends on the type of the prediction model used (as different machine learning algorithms have different parameters and approaches to reach the optimal solution).

Formally, let N+1 be the total participating sites, where $site-0$ is considered as main site and

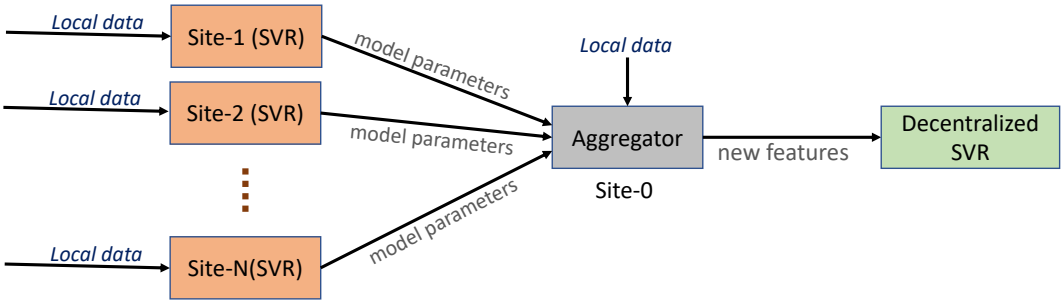


Figure 1: Overall flow of model parameters from locally trained SVR models that are sent to the main site for aggregation. The aggregated model parameters are used to transform the data at the main site into new features, which are used for training a new decentralized SVR model.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.10.443469; this version posted May 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the remaining sites are local sites. Let the matrix $X_i^{m\times n} \in R$ be data of m subjects with n features available at a site $i \in \{0, 1, 2, .., N\}$, with the matrix $Y_i^{m\times 1} \in R$ representing their chronological age. At each local site i, a regression model M_i is constructed, and the corresponding model parameters $P_i^{n\times k}$ are sent to the main site, where k is the dimension of the parameters generated by the model.

$$Y_i^{pred} = M_i(X_i^{m\times n}, Y_i^{m\times 1}) \tag{1}$$

At main site ($site - 0$), these model parameters are aggregated and used to transform the original data $X_0^{m\times n} \in R$. The data are transformed using the parameter matrix with the $\bigotimes$ operation. In our case, it is a simple matrix multiplication due to a linear regression model, although in general it can be any matrix operation. The transformed features $Xnew_0^{m\times k}$ are used for BAE.

$$P_0^{n\times k} = Average(P_i^{n\times k}), i \in \{1, 2, 3, .., N\}$$

$$Xnew_0^{m\times k} = X_0^{m\times n} \bigotimes P_i^{n\times k}$$

$$Y_0^{pred} = M_0(Xnew_0^{m\times k}, Y_0^{m\times 1})$$

Here M_0 represents the aggregated model containing information from models $M_i, i \in \{1, 2, .., N\}$.

Algorithm 1: Decentralized brainage prediction

Data: $\mathbf{X}_i^{m\times n} \in R$ be data of m subjects, with n features each, available at site $i \in \{0, 1, 2, .., N\}$, and $\mathbf{Y}_i^{m\times 1} \in R$ representing their chronological age.

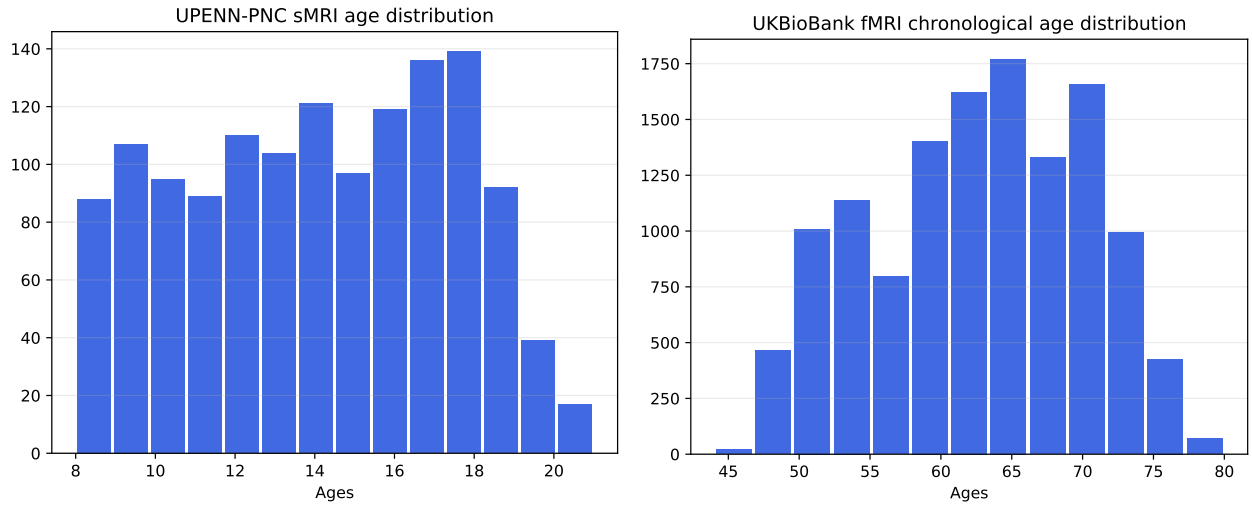
```
/* Local sites                                                    */
for all sites i in 1 .. N do
    Construct a regression model M_i with (X_i^{m×n} as features and Y_i^{m×1}) as age estimator
    Send model parameters P_i^{n×k} to the aggregator site
end
/* aggregator site                                                */
if site i == 0 then
    Gather parameters from all local sites i=1..N
    Combine parameters P_0^{n×k} = Average(P_i^{n×k}), i ∈ {1, 2, 3, .., N}
    Transform data X_0 to Xnew_0^{m×k} = X_0^{m×n} ⊗ P_i^{n×k}
    Construct a decentralized model M_0 with Xnew_0^{m×k} as features and Y_0^{m×1}) as age
      estimator
end
```

In our initial setup, all the local sites are provided with the details of the prediction model and the type of information to be shared with the main site after training their local models. Of all the participating sites, N sites are used as local sites, and the remaining site is used as the main site. In this study, we use support vector regression (SVR) as the prediction model and apply decentralization by employing a training strategy similar to [19, 25]. As a first step, all the local sites train an SVR model locally with their data and transfer the weight vectors of the locally learned models to the main site. The main site averages these vectors and uses the average weight vector to transform its data into a new feature space. This modified data is then used to train a decentralized SVR model (see Fig. 1).

4 Results

The primary goal of this work is to develop a decentralized BAE model and contrast its performance against a centralized model. For this purpose, we performed three experiments with features extracted from two different datasets. All experiments were performed in COINSTAC [9] with six sites (one as main and the remaining five as local nodes).

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.10.443469; this version posted May 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

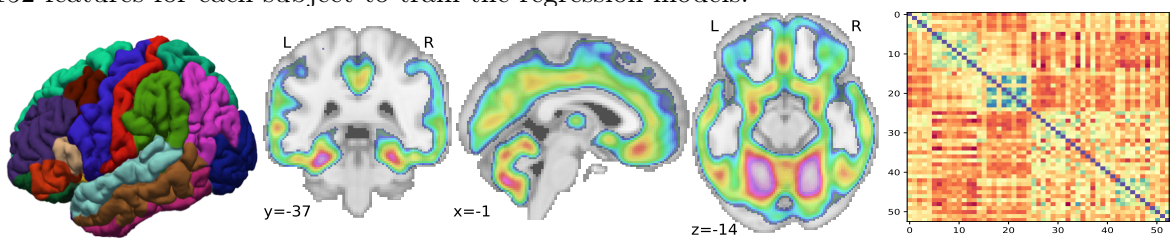


(a) UPENN-PNC (b) UKBioBank

Figure 2: Histograms showing the chronological age distribution of healthy subjects of UPENN-PNC (left) and UKBiobank (right) datasets.

Datasets: There are two MRI datasets that are analyzed in this work, *viz*, UPENN-PNC [26] and UKBiobank [27]. The UPENN-PNC data consists of T1-weighted sMRI of 1417 health participants with chronological ages ranging from 8 to 21 years during acquisition. This data is jointly spatially normalized and segmented and then smoothed by a 6-mm full-width at half maximum (FWHM) Gaussian kernel. Segmentation was performed in SPM12 [28]. The second dataset, UKBiobank, has fMRI images of 11,754 subjects with chronological ages in the range of 44 to 80 years. The data were preprocessed using a combination of FSL [29] and SPM12. This included distortion correction, rigid body motion correction, normalization to Montreal Neurological Institute (MNI) space and smoothing using a Gaussian kernel with an FWHM of 6 mm. Next, we ran a fully-automated independent component analysis (ICA) pipeline that can capture corresponding functional network features while retaining more single-subject variability [30]. This pipeline has been successfully applied to multiple studies to identify a wide range of connectivity abnormalities in numerous brain diseases [30]. Group ICA was first performed on two large healthy control datasets to create spatial network (component) priors following which a spatially constrained ICA algorithm was applied to back-reconstruct spatial maps and time-courses (TCs) for each subject [31]. Fig. 2 shows the age range distribution of the subjects in these datasets. Fig. 3 shows the three different features extracted using these datasets as described below.

Prediction using FreeSurfer features: We used FreeSurfer (version v5.3) [32] to extract brain structural features from UPENN-PNC data. A standard `aseg.stats` file that has features corresponding to total Intracranial Volume (eTIV), left hemisphere (lh) and right hemisphere (rh) subcortical regions is generated. Additionally, features corresponding to cortical thicknesses and volumes of the parcellated regions in surface GM in both left and right hemispheres are also extracted. In total, we use 152 features for each subject to train the regression models.



(a) FreeSurfer features (b) Graymatter ROI (c) FNC Component maps

Figure 3: Different features extracted from sMRI and fMRI datasets for BAE.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.10.443469; this version posted May 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Prediction using GM features: We use the Automatic Anatomical Labeling (AAL) [33] brain atlas of 116 brain parcellations to extract the mean GM density of each region of interest (ROIs)from UPENN-PNC data for 1417 subjects. The extracted mean GM density of 116 ROIs are used as GM features to train the regression models.

Prediction using FNC features: From the UKBiobank, we applied a fully automated spatially constrained ICA approach to fMRI data from 11,754 subjects, yielding 53 nonartifactual ICNs with 490 time points each (for more details, see Datasets in Sec. 4). Therefore, for each subject we have 53×53 correlation matrices as features. Because these matrices are symmetric, we extract the nondiagonal upper triangular matrix values (1378 features) for each subject to train the regression models.

Sampling methods and metrics: As the data need to be distributed across six sites, we employ three strategies to split the data equally. We employ a 90%-10% train-test split at each site. In the first approach, we randomly partition the data into these sites, and within each site, the data is further randomly split into training and testing datasets. This approach is referred to as *random sampling*. In the second approach, we use stratified sampling based on subject age to partition the data into six sites. Within each site, the training and testing datasets were randomly split. We call this *age-stratified sampling*. In the third approach, we group the subjects into different bins based on their age ranges, label these bins and use these labels to perform stratified sampling. We refer to this method as *age-bin-stratified sampling*. All experiments were repeated five times, including splitting the data and training models. To build a centralized model, training data from all the sites is pooled into one training dataset, and testing data from all the sites is combined to create the testing dataset.

We use root mean square error (RMSE) and mean absolute error (MAE) as metrics to measure the performance of the fitted regression models. RMSE is a standard measure used to analyze the performance of regression models. It is a measure of the distance between the predicted values and actual values and is computed as follows:

Table 1: Performance comparison of decentralized and centralized models for BAE.

$Input-data$	$Sampling$	$Method$	$RMSE$		MAE	
			train	test	train	test
FreeSurfer	random	Decentralized	3.673 ± 0.045	3.624 ± 0.242	3.139 ± 0.068	3.083 ± 0.221
		Centralized	3.518 ± 0.025	3.845 ± 0.149	2.863 ± 0.017	3.244 ± 0.161
	age stratified	Decentralized	3.701 ± 0.052	3.7 ± 0.184	3.156 ± 0.051	3.176 ± 0.172
		Centralized	3.501 ± 0.012	3.911 ± 0.137	2.85 ± 0.01	3.312 ± 0.159
	age-bin stratified	Decentralized	3.652 ± 0.033	3.646 ± 0.071	3.131 ± 0.038	3.102 ± 0.03
		Centralized	3.522 ± 0.01	3.862 ± 0.15	2.863 ± 0.013	3.24 ± 0.115
GM	random	Decentralized	2.652 ± 0.132	2.459 ± 0.156	2.172 ± 0.134	1.983 ± 0.147
		Centralized	2.138 ± 0.015	2.213 ± 0.141	1.686 ± 0.012	1.769 ± 0.109
	age stratified	Decentralized	2.588 ± 0.035	2.489 ± 0.141	2.063 ± 0.031	2.029 ± 0.105
		Centralized	2.151 ± 0.01	2.185 ± 0.081	1.691 ± 0.009	1.761 ± 0.055
	age-bin stratified	Decentralized	2.625 ± 0.061	2.613 ± 0.087	2.131 ± 0.061	2.12 ± 0.056
		Centralized	2.149 ± 0.004	2.2 ± 0.062	1.694 ± 0.005	1.752 ± 0.054
FNC	random	Decentralized	7.497 ± 0.076	7.486 ± 0.074	6.285 ± 0.078	6.284 ± 0.066
		Centralized	7.21 ± 0.005	7.857 ± 0.084	5.722 ± 0.009	6.535 ± 0.06
	age stratified	Decentralized	7.502 ± 0.034	7.494 ± 0.073	6.298 ± 0.04	6.284 ± 0.058
		Centralized	7.203 ± 0.01	7.968 ± 0.123	5.709 ± 0.018	6.623 ± 0.138
	age-bin stratified	Decentralized	7.511 ± 0.056	7.501 ± 0.047	6.318 ± 0.045	6.279 ± 0.038
		Centralized	7.194 ± 0.005	7.997 ± 0.096	5.708 ± 0.01	6.624 ± 0.059

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.10.443469; this version posted May 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$$RMSE = \sqrt{\frac{\Sigma_{i=1}^{n_{samples}} \left(y_{actual} - y_{predicted}\right)^2}{n_{samples}}}$$

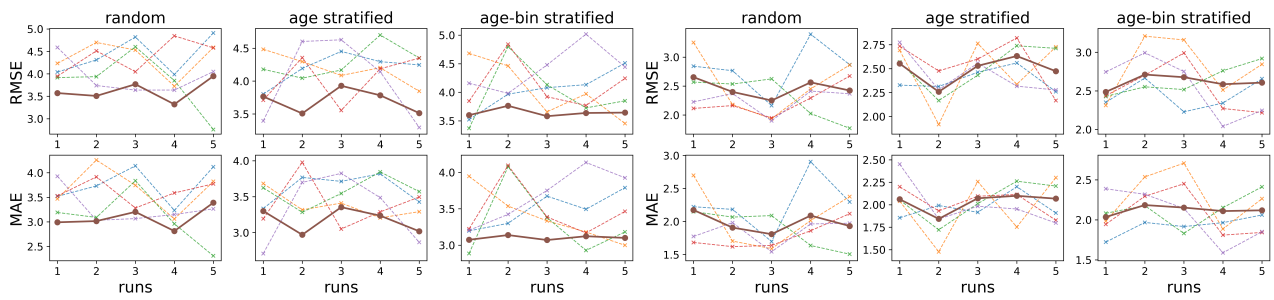
We also compute MAE as it reflects the performance of brain age gaps and is commonly used in literature. It is an average measure of the magnitude of prediction errors without considering their error directions:

$$MAE = \frac{\Sigma_{i=1}^{n_{samples}} |y_{actual} - y_{predicted}|}{n_{samples}}$$

The results show that the decentralized models have similar performance compared to centralized models (see Table 1). For models trained using FreeSurfer features, the decentralized models showed higher RMSE and MAE values on the training data across all sampling strategies, whereas the centralized models demonstrated slightly higher values on the test data. For models trained with GM features, we observed slightly higher error values for decentralized models compared to centralized models on both the training and testing sets across all sampling strategies. For models trained with FNC features, the observations were similar to the ones found on models trained with FreeSurfer features. The decentralized models have slightly higher error values compared to centralized models during training, whereas the testing metrics are slightly improved compared to centralized values.

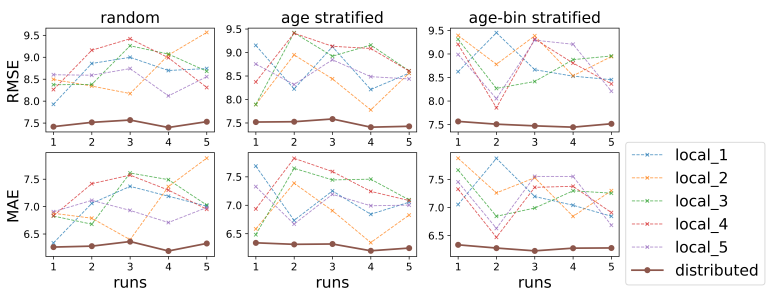
Fig. 4 shows the performance of the final (decentralized) model compared to the local models across five runs for different feature sets. Fig. 4(a) and 4(b) show the performance of distributed and local models for FreeSurfer and GM models for 1417 subjects, and Fig. 4(c) shows the performance of the FNC dataset with 11754 subjects. These results show that as the amount of training data increases, the distributed model achieves better accuracy compared to any of the locally trained models.

We statistically compare the performance of centralized and decentralized models for both measures using a Wilcoxon signed-rank test [34]. This is a nonparametric pairwise comparison test, with no assumptions on the data distribution and a null hypothesis that the differences between two samples have a distribution centered about zero. For each metric (RMSE and MAE), we compare the training scores of decentralized and centralized models for different sampling methods separately using different input data features. We repeat this setup and also compare the testing scores for these metrics. In all the cases, this test fails to reject the null hypothesis indicating that the decentralized models achieved



(a) FreeSurfer

(b) GM



(c) FNC

Figure 4: Performance of decentralized models compared to local models across five runs (repetitions) for different feature sets.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.10.443469; this version posted May 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

performance similar to that of centralized models. This is a very interesting observation as we tested decentralized models with samples of over one thousand and over 11 thousands,and in all the cases decentralized models perform similar to centralized models. This indicates the de-centralized models work consistently with centralized models for smaller or larger data samples.

5 Conclusion

In this work, we employ decentralized machine learning for brain age estimation and compare the results with a centralized model. The decentralized model is trained by utilizing information from locally trained models at different sites and involves no data sharing. Results from models trained on three different feature sets with three different data splitting strategies show that the predictive performance in the decentralized model is consistently better than individual sites separately, and as good as the centralized model. The key benefit of decentralization is that it does not require data sharing and therefore encourages collaboration by allowing different research groups to readily participate in larger studies without worrying about their data-sharing policies or data transmission. Future work will include developing a multishot version of this pipeline coupled with differential privacy strategies.

Acknowledgments

This work was funded by the National Institutes of Health (R01DA040487), National Institute on Drug Abuse (R01DA049238) and the National Institute of Mental Health (R01MH121246).

Author Contributions

All the authors helped improve the manuscript. SB designed decentralized models for FNC features, performed the data analysis for all the models and wrote the initial manuscript. RR and HG designed the decentralized model for FreeSurfer and GM features. BR designed feature extraction strategies. JL provided guidance about feature extraction and decentralized model design. EV manages the COINSTAC project and helped write the paper. VDC supervised all the stages of the project and also funded the project.

References

[1] J. H. Cole, R. E. Marioni, S. E. Harris, and I. J. Deary, "Brain age and other bodily 'ages': implications for neuropsychiatry," *Molecular psychiatry*, vol. 24, no. 2, pp. 266–281, 2019.

[2] K. Franke and C. Gaser, "Ten years of brainage as a neuroimaging biomarker of brain aging: what insights have we gained?," *Frontiers in neurology*, vol. 10, p. 789, 2019.

[3] M. L. Elliott, D. W. Belsky, A. R. Knodt, D. Ireland, T. R. Melzer, R. Poulton, S. Ramrakha, A. Caspi, T. E. Moffitt, and A. R. Hariri, "Brain-age in midlife is associated with accelerated biological aging and cognitive decline in a longitudinal birth cohort," *Molecular psychiatry*, pp. 1–10, 2019.

[4] A. Reeve, E. Simcox, and D. Turnbull, "Ageing and parkinson's disease: why is advancing age the biggest risk factor?," *Ageing research reviews*, vol. 14, pp. 19–30, 2014.

[5] J. H. Cole, S. J. Ritchie, M. E. Bastin, M. V. Hernández, S. M. Maniega, N. Royle, J. Corley, A. Pattie, S. E. Harris, Q. Zhang, *et al.*, "Brain age predicts mortality," *Molecular psychiatry*, vol. 23, no. 5, pp. 1385–1392, 2018.

[6] F. Liem, G. Varoquaux, J. Kynast, F. Beyer, S. K. Masouleh, J. M. Huntenburg, L. Lampe, M. Rahim, A. Abraham, R. C. Craddock, *et al.*, "Predicting brain-age from multimodal imaging data captures cognitive impairment," *Neuroimage*, vol. 148, pp. 179–188, 2017.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.10.443469; this version posted May 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[7] B. A. Jónsson, G. Bjornsdottir, T. Thorgeirsson, L. M. Ellingsen, G. B. Walters, D. Gudbjartsson, H. Stefansson, K. Stefansson, and M. Ulfarsson, "Brain age prediction using deep learning uncovers associated sequence variants," *Nature communications*, vol. 10, no. 1, pp. 1–10, 2019.

[8] T. White, E. Blok, and V. D. Calhoun, "Data sharing and privacy issues in neuroimaging research: Opportunities, obstacles, challenges, and monsters under the bed," *Human Brain Mapping*, 2020.

[9] S. M. Plis, A. D. Sarwate, D. Wood, C. Dieringer, D. Landis, C. Reed, S. R. Panta, J. A. Turner, J. M. Shoemaker, K. W. Carter, *et al.*, "Coinstac: a privacy enabled model and prototype for leveraging and processing decentralized brain imaging data," *Frontiers in neuroscience*, vol. 10, p. 365, 2016.

[10] E. Luders, N. Cherbuin, and C. Gaser, "Estimating brain age using high-resolution pattern recognition: Younger brains in long-term meditation practitioners," *Neuroimage*, vol. 134, pp. 508–513, 2016.

[11] J. Steffener, C. Habeck, D. O'Shea, Q. Razlighi, L. Bherer, and Y. Stern, "Differences between chronological and brain age are related to education and self-reported physical activity," *Neurobiology of aging*, vol. 40, pp. 138–144, 2016.

[12] J. H. Cole, R. P. Poudel, D. Tsagkrasoulis, M. W. Caan, C. Steves, T. D. Spector, and G. Montana, "Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker," *NeuroImage*, vol. 163, pp. 115–124, 2017.

[13] L. Yang, C. Cao, E. D. Kantor, L. H. Nguyen, X. Zheng, Y. Park, E. L. Giovannucci, C. E. Matthews, G. A. Colditz, and Y. Cao, "Trends in sedentary behavior among the us population, 2001-2016," *Jama*, vol. 321, no. 16, pp. 1587–1597, 2019.

[14] H. Sajedi and N. Pardakhti, "Age prediction based on brain mri image: a survey," *Journal of medical systems*, vol. 43, no. 8, p. 279, 2019.

[15] M. J. Jafri, G. D. Pearlson, M. Stevens, and V. D. Calhoun, "A method for functional network connectivity among spatially independent resting-state components in schizophrenia," *Neuroimage*, vol. 39, no. 4, pp. 1666–1681, 2008.

[16] X. Niu, F. Zhang, J. Kounios, and H. Liang, "Improved prediction of brain age using multimodal neuroimaging data," *Human brain mapping*, vol. 41, no. 6, pp. 1626–1643, 2020.

[17] K. Stankevičiūtė, T. Azevedo, A. Campbell, R. A. Bethlehem, and P. Liò, "Population graph gnns for brain age prediction," *bioRxiv*, 2020.

[18] C. Sudlow, J. Gallacher, N. Allen, V. Beral, P. Burton, J. Danesh, P. Downey, P. Elliott, J. Green, M. Landray, *et al.*, "Uk biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age," *Plos med*, vol. 12, no. 3, p. e1001779, 2015.

[19] A. D. Sarwate, S. M. Plis, J. A. Turner, M. R. Arbabshirani, and V. D. Calhoun, "Sharing privacy-sensitive access to neuroimaging and genetics data: a review and preliminary validation," *Frontiers in neuroinformatics*, vol. 8, p. 35, 2014.

[20] M. Aledhari, R. Razzak, R. M. Parizi, and F. Saeed, "Federated learning: A survey on enabling technologies, protocols, and applications," *IEEE Access*, vol. 8, pp. 140699–140725, 2020.

[21] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 50–60, 2020.

[22] "COINSTAC." http://coinstac.trendscenter.org.

[23] J. Ming, E. Verner, A. Sarwate, R. Kelly, C. Reed, T. Kahleck, R. Silva, S. Panta, J. Turner, S. Plis, *et al.*, "Coinstac: Decentralizing the future of brain imaging analysis," *F1000Research*, vol. 6, 2017.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.10.443469; this version posted May 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[24] H. Gazula, B. Holla, Z. Zhang, J. Xu, E. Verner, R. Kelly, G. Schumann, and V. D. Calhoun, "Decentralized multi-site vbm analysis during adolescence shows structural changes linked to age, body mass index, and smoking: A coinstac analysis," *bioRxiv*, p. 846386, 2019.

[25] K. Chaudhuri, C. Monteleoni, and A. D. Sarwate, "Differentially private empirical risk minimization.," *Journal of Machine Learning Research*, vol. 12, no. 3, 2011.

[26] T. D. Satterthwaite, M. A. Elliott, K. Ruparel, J. Loughead, K. Prabhakaran, M. E. Calkins, R. Hopson, C. Jackson, J. Keefe, M. Riley, *et al.*, "Neuroimaging of the philadelphia neurodevelopmental cohort," *Neuroimage*, vol. 86, pp. 544–553, 2014.

[27] K. L. Miller, F. Alfaro-Almagro, N. K. Bangerter, D. L. Thomas, E. Yacoub, J. Xu, A. J. Bartsch, S. Jbabdi, S. N. Sotiropoulos, J. L. Andersson, *et al.*, "Multimodal population brain imaging in the uk biobank prospective epidemiological study," *Nature neuroscience*, vol. 19, no. 11, pp. 1523–1536, 2016.

[28] J. Ashburner, G. Barnes, C.-C. Chen, J. Daunizeau, G. Flandin, K. Friston, S. Kiebel, J. Kilner, V. Litvak, R. Moran, *et al.*, "Spm12 manual," *Wellcome Trust Centre for Neuroimaging, London, UK*, vol. 2464, 2014.

[29] S. Smith, M. Woolrich, T. Behrens, C. Beckmann, D. Flitney, M. Jenkinson, P. Bannister, S. Clare, M. De Luca, P. Hansen, *et al.*, "Fmrib software library."

[30] Y. Du, G. D. Pearlson, J. Liu, J. Sui, Q. Yu, H. He, E. Castro, and V. D. Calhoun, "A group ica based framework for evaluating resting fmri markers when disease categories are unclear: application to schizophrenia, bipolar, and schizoaffective disorders," *Neuroimage*, vol. 122, pp. 272–280, 2015.

[31] Y. Du, Z. Fu, J. Sui, S. Gao, Y. Xing, D. Lin, M. Salman, M. A. Rahaman, A. Abrol, J. Chen, *et al.*, "Neuromark: a fully automated ica method to identify effective fmri markers of brain disorders," *medRxiv*, p. 19008631, 2019.

[32] B. Fischl, "Freesurfer," *Neuroimage*, vol. 62, no. 2, pp. 774–781, 2012.

[33] E. T. Rolls, C.-C. Huang, C.-P. Lin, J. Feng, and M. Joliot, "Automated anatomical labelling atlas 3," *Neuroimage*, vol. 206, p. 116189, 2020.

[34] R. Woolson, "Wilcoxon signed-rank test," *Wiley encyclopedia of clinical trials*, pp. 1–3, 2007.