2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Federation of Brain Age Estimation in Structural Neuroimaging 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]

Abstract— **Brain age estimation is a widely used approach to evaluate the impact of various neurological or psychiatric brain disorders on the brain developmental or aging process. Current studies show that neuroimaging data can be used to predict brain age, as it captures structural and functional changes that the brain undergoes during development and the aging process. A robust brain age prediction model not only has the potential in assisting early diagnosis of brain disorders but also helps in monitoring and evaluating effects of a treatment. Although access to large amounts of data helps build better models and validate their effectiveness, researchers often have limited access to brain data because of its challenging and expensive acquisition process. This data is not always sharable due to privacy restrictions. Decentralized models provide a way which does not require data exchange between the multiple involved groups. In this work, we propose a decentralized approach for brain age prediction and evaluate our models using features extracted from structural MRI data. Results demonstrate that our decentralized brain age model achieves similar performance compared to the models trained with all the data in one location.**

Keywords: Brain Age, Decentralized, Federated, COINSTAC, Neuroimaging

I. Introduction

Brain age estimation (BAE) from magnetic resonance imaging (MRI) images has become widely used in recent years. The brain age gap, computed as the difference between chronological age and estimated brain age, is helpful in early identification of various neurodegenerative diseases, such as Alzheimer's, Huntington's, and Parkinson's diseases, and is known to be present in patients with disorders such as dementia and autism[1], [2]. There are several studies using machine learning and deep learning algorithms with promising prediction models[3], [4]. However, there is no clear understanding on which models perform the best. It is a known fact that having a large training dataset helps in achieving robust models, but due to the expensive data collection process and lack of availability of many subjects with MRI data, one is limited by the amount of data that can be gathered for analysis. This is especially true for evaluating the 'brain age gap' which is often used to evaluate patient groups. One approach to alleviate this problem is to work with different datasets collected worldwide. Researchers can greatly benefit if a larger study can be performed using these worldwide datasets.

One approach of pooling data is to collect all the data in a centralized location and perform analysis. This approach is inefficient for neuroimaging-based studies as MRI data is usually large and saving all the data in one central location not only has high transmission costs and delays, but also requires large amounts of redundant storage. In addition, data cannot be readily shared (which represents a substantial amount of the existing data) due to privacy restrictions.

A better approach to bypass these problems is to use decentralized (or federated) algorithms, which do not require assembling the data in one central location. Decentralized algorithms are particularly important when there is a need to perform analysis on large data that involve diverse worldwide datasets without worrying about data transmission or violating privacy. Decentralized algorithms have been successfully used in different domains and are an active research area.

Prior studies using BAE perform analysis require all the data to be available at one location. In this work, we develop and apply a decentralized brain age prediction algorithm using a regression model and show that decentralized prediction models achieve performance similar to that of centralized models. We perform detailed experiments on surface parcellation features extracted from structural MRI data using various sampling methods and demonstrate the robustness of our approach. We implement our approach within the Collaborative Informatics and Neuroimaging Suite Toolkit for Anonymous Computation[5], [6] framework for our decentralized analysis, which emphasizes decentralized approaches for neuroimaging analysis.

II. Background

A. Brain age Estimation and current approaches

Brain age Estimation (BAE) is helpful in early identification of some neurodegenerative diseases(Alzheimer, Parkinson and others) and identifying brain age gaps (a significant symptom in patients with dementia and autism)[2]. 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 measured directly, models are typically trained with chronological age of the healthy subjects to predict the brain age and the difference between the estimated brain age and chronological age is considered to be the brain age gap. The chronological age of an individual is usually close to the brain age; however, studies show that brain age gap can increase due to neurodegenerative diseases (such as

[1]Tri-institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), 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, Atlanta, GA, USA
[4]Princeton Neuroscience Institute, Princeton, NJ, USA
[5]Department of Psychology, Georgia State University, Atlanta, GA, USA

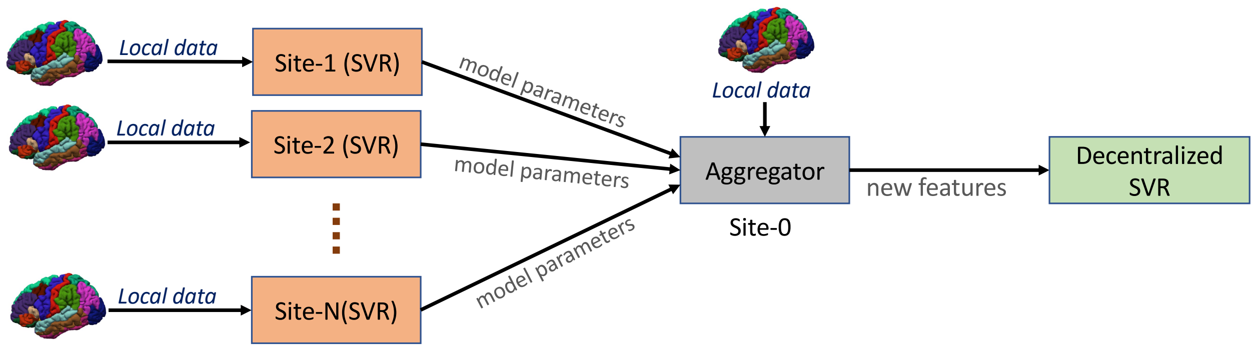


Fig. 1. Overall flow of model parameters from locally trained SVR models with Freesurfer features extracted from sMRI data that are sent to the main site for aggregation. The aggregated model parameters are used to transform the data at the main site into new features, which are used for training a new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

Alzheimer's, Huntington's, Parkinson's, and multiple sclerosis) or decrease with activities such as meditation[7] and physical exercise[8].

Recently, MRI images are widely used for brain age prediction. There are broadly three different approaches that have been previously used, namely, voxel-based, region-based, and surface-based approaches[9]. Deep learning models have shown promising results in BAE. Authors of [10] compute BAE using both raw MR images and preprocessed MR images with deep learning models, and both showed consistent results. The authors of [11] 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[12] using data from the UK Biobank[13]. 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.

B. Decentralized learning

Decentralized learning (also referred to as federated learning in literature) provides a way for worldwide research groups to collaborate and build more accurate and generalizable prediction models while providing solutions to address data transfer or their data-sharing policies[14]. Most of the research in decentralized machine learning algorithms is focused on deep learning models due to their flexibility and high performance in a wide number of domains. In this method, all the participating sites start with the same model, and each 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 reaches its stopping criteria, such as the desired performance. Sometimes, instead of averaging the gradients, the locally trained model parameters are sent to the main site where all the local model parameters are aggregated (known as model averaging). The new models are then sent to the participating sites for the next training iteration and it continues until a desired performance is achieved. There are several challenges in designing decentralized algorithms, including heterogeneity across different sites, transmission, maintaining synchronization during training iterations, and preserving data privacy. The authors of[15], [16] provide a detailed summary of the machine learning algorithms, challenges, available architectures for decentralization.

III. METHODS

Decentralized learning has been applied in some domains but, this paper, to our knowledge, presents the first approach for decentralized brain age analysis. As stated earlier, in a centralized approach, models are built with the data from all the sites pooled at a central location, and therefore a centralized model has better performance than any model trained only on a subset of that data.

In this work, brain age prediction is done using a decentralized approach where 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).

Our decentralized algorithm is described in Alg. 1. Let us assume there are $N+1$ participating sites, each gathering data from a different set of participants. In our initial setup, all the local sites (*sites from 1,2,..,N)* are provided with the details of the prediction model and the type of information to be shared with the main site(*site-0*) 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)[17] as the prediction model and apply decentralization by employing a training strategy similar to[14], [18]. As a first step, all the local sites train an SVR model locally with their data and

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 ($\mathbf{X}_i^{m \times n}$ as features and $\mathbf{Y}_i^{m \times 1}$) as age estimator

  Send model parameters $\mathbf{P}_i^{n \times k}$ to the aggregator site

end

/* aggregator site */

if *site i == 0* **then**

  Gather parameters from all local sites *i=1..N*

  Combine parameters $\mathbf{P}_0^{n \times k} = Average(\mathbf{P}_i^{n \times k}), i \in \{1, 2, 3, .., N\}$

  Transform data $\mathbf{X}_0$ to $\mathbf{Xnew}_0^{m \times k} = \mathbf{X}_0^{m \times n} \bigotimes \mathbf{P}_i^{n \times k}$

  Construct a decentralized model $\mathbf{M}_0$ with $\mathbf{Xnew}_0^{m \times k}$ as features and $\mathbf{Y}_0^{m \times 1}$) as age estimator

end

transfer the weight vectors of the locally learned models to the main site. The main site (*site-0*) 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).

IV. Results

The main emphasis of this work is to develop a decentralized BAE model that has comparable performance with its corresponding centralized model. For this purpose, we perform multiple experiments with Freesurfer ROI features extracted from UPENN-PNC sMRI dataset with six sites (one as main and the remaining five as local sites) using COINSTAC[5], [6], [19], a platform that enables decentralized analysis of (neuroimaging) data without the need to pool the data at one location. 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.

The UPENN-PNC[20] data consists of sMRI of 1591 healthy subjects 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[21]. We use FreeSurfer[22] (Fig. 1) 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.

To distribute the data across six sites, we employ three strategies to split the data equally, with 90%-10% train-test split at each site. In the first approach (*random sampling*), data is randomly partitioned across these sites, and within each site, the data is further randomly split into training and testing datasets. In the second approach named as *age-stratified sampling*, we use stratified sampling based on subject age to partition the data into six sites. Within each site, the training and testing datasets were split randomly. In the third approach referred as *age-bin-stratified sampling*, subjects are grouped into different bins based on their age ranges, label these bins and use these labels to perform stratified sampling. In the case of *age-bin-stratified sampling*, data is grouped into 8 bins based on the age and these 8 labelled classes are used to divide the data. 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.

Root mean square error (RMSE) and mean absolute error (MAE) are the metrics generally used in literature to measure the performance of the fitted regression models. RMSE is a standard measure used to analyze the performance of regression models and MAE is an average measure of the magnitude of absolute difference between the actual and predicted values (generally used to measure brain age gaps). Clearly, lower the values of these measures better the prediction model. Both the metrics are computed as follows:

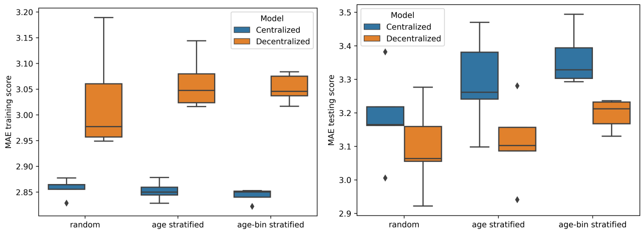
$$RMSE = \sqrt{\frac{\Sigma_{i=1}^{n_{samples}} \left(y_{actual} - y_{predicted} \right)^2}{n_{samples}}}$$

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

Fig. 2 shows the performance of the decentralized models compared to their corresponding centralized models. 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. To statistically compare the performance of centralized and decentralized models for both measures, we used a Wilcoxon signed-rank test[23]. 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. We compare the training and testing scores for both RMSE and MAE for all sampling methods, and this test fails to reject the null hypothesis in all the cases (with $\alpha = 0.05$ as shown in Fig. 3). Results thus confirm that the decentralized models achieve performance similar to that of centralized models.

V. Conclusions

In this work, decentralized machine learning models are used for brain age estimation compared to existing cen-



(a) MAE Training scores (b) MAE Testing scores (c) RMSE Training scores (d) RMSE Testing scores

Fig. 2. Performance comparison of centralized and decentralized models using Freesurfer features of sMRI data with varying sampling methods. Statistical nonparametric pairwise comparison test shows that decentralized models has performance similar to that of centralized models.

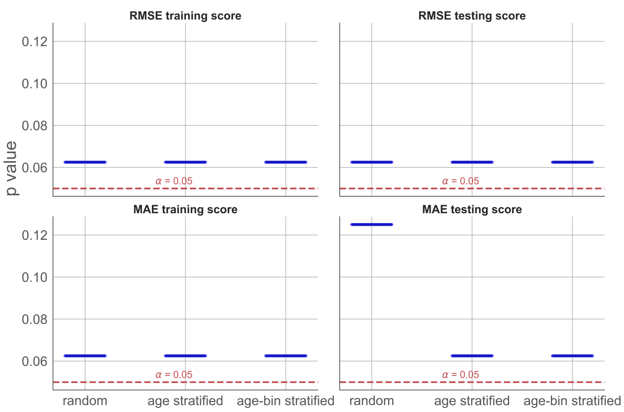


Fig. 3. Results of Wilcoxon signed-rank test for different metrics. p-values greater than reference $\alpha = 0.05$ (red line) show that they fail to reject the null hypothesis showing that decentralized and centralized models have similar performance for these metrics.

tralized approaches. In this approach, decentralized models are built using the information of locally trained models at different sites and do not involve actual data sharing. We compare performance using features extracted from sMRI data with three different data splitting strategies and showed that decentralized models have similar performance to their centralized counterparts. The key benefit of decentralization is that it encourages collaboration by allowing different research groups to readily participate in larger studies without worrying about their data-sharing policies or data transmission.

Acknowledgment

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).

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

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

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

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

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] H. Drucker, C. J. Burges, L. Kaufman, A. Smola, V. Vapnik, *et al.*, "Support vector regression machines," *Advances in neural information processing systems*, vol. 9, pp. 155–161, 1997.

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

[19] 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.

[20] 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.

[21] 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.

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

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