Investigating Voxel-level Brain Age Prediction as a Pretext Task for Brain MRI Segmentation

Tasneem Nasser[1,3], Roberto Souza[2,3], and Naser El-Sheimy[4]

[1] Department of Biomedical Engineering, University of Calgary, Calgary, AB, Canada
[2] Department of Electrical and Software Engineering, University of Calgary, Calgary, AB, Canada
[3] Hotchkiss Brain Institute, University of Calgary, Calgary, AB, Canada
[4] Department of Geomatics Engineering, University of Calgary, Calgary, AB, Canada

Abstract. To address the challenge of few annotated datasets for training brain magnetic resonance imaging (MRI) segmentation models, we propose to use voxel-level brain age prediction as a domain-specific pretext task for self-supervised learning before adapting models to a segmentation downstream task. We combined publicly available T1-weighted, normative brain MRI datasets to create a large (N = 1,710), representative dataset with a balanced distribution across age groups and sexes, minimizing potential biases in our model. We then compared three state-of-the-art architectures, Swin UNETR, UNETR, and UNET, on the voxel-level brain age prediction pretext task. Swin UNETR achieved the best performance with a mean absolute error (MAE) of 5.9 ± 4.4 years, outperforming UNETR (MAE: 7.2 ± 4.4 years) and UNET (MAE: 6.2 ± 4.2 years). Based on this performance, we selected Swin UNETR for a brain MRI segmentation downstream task to evaluate the effectiveness of the voxel-level brain age prediction as a self-supervised learning pretext task. We fine-tuned it and compared its performance against two baselines: (1) training from scratch and (2) fine-tuning a model pre-trained on an image inpainting task, a non-domain-specific pretext task. The Swin UNETR model pre-trained on voxel-level brain age prediction achieved the highest Dice coefficient on an out-of-distribution test set and performed comparably to the inpainting-pretrained model on an in-distribution test set. These results demonstrate the potential of voxel-level brain age prediction as a domain-specific pretext task for self-supervised learning in neuroimaging, improving segmentation performance, especially in challenging, low-data scenarios.

Keywords: Brain age prediction · Segmentation · Self-supervised learning · Swin UNETR.

1 Introduction

Data scarcity is a significant challenge in training Deep learning (DL) models, particularly in fields like medical imaging, where labelled data is often limited

Our code is available at: https://github.com/TasneemN/BrainAgePrediction_asPretextTask

due to the time-intensive nature and expertise required to annotate the images. Since most DL frameworks rely on supervised learning, they require substantial amounts of annotated data to learn meaningful representations effectively. Given the difficulty of manual annotation, self-supervised learning (SSL) has gained traction as an alternative paradigm for feature learning [26]. SSL leverages pretext tasks such as solving jigsaw puzzles, memorizing spatial relationships, predicting image rotations, and restoring missing regions to extract meaningful representations from unlabeled data. Recent advancements in SSL have focused on masking and restoring image patches as a learning strategy. Masked Autoencoder pretrains Vision Transformers (ViTs) by randomly masking image regions and predicting the missing content, a technique known as masked image modeling (MIM) [12]. Kim et al. [12] applied this concept to Swin Transformers, demonstrating their effectiveness in learning visual features without requiring extensive labeled datasets. While SSL pretext tasks such as inpainting have proven effective in computer vision for pre-training models to extract generalizable features, their direct application to 3D medical imaging remains challenging. To address this, designing domain-aware pre-training tasks is crucial to ensure models learn meaningful medical-specific features rather than irrelevant patterns [12]. Therefore, we propose leveraging voxel-level brain age prediction [18,8] as a pretext task. Since aging is a universal and inevitable process, and the widespread availability of age information in many normative neuroimaging datasets makes it a valuable proxy for learning meaningful brain representations. This approach builds upon the assumption that, for healthy individuals, brain age is equal to the chronological age [6]. However, this assumption does not apply to individuals with underlying neurological disorders such as Alzheimer's disease, depression, and schizophrenia, where an accelerated aging pattern is often observed and considered as an effective biomarker of developing neurological disease [8].

Aging is characterized by the gradual decline of physiological functions due to the accumulation of cellular damage over time [8]. Brain aging affects cognitive function and overall well-being, making it essential to study its patterns in healthy individuals and correlate them with the onset of mental disorders for early detection and intervention [8]. Unlike traditional pretext tasks such as image inpainting [12,5], brain age prediction focuses on detecting subtle structural changes in the brain associated with aging, which may enhance the model's ability to extract relevant features for other neuroimaging tasks. Therefore, we hypothesize that the voxel level brain age prediction can serve as an effective domain-specific pretext task for self-supervised learning in brain MRI segmentation tasks, enabling models to learn meaningful anatomical representations that enhance performance on downstream tasks.

Our objective is to investigate voxel-level brain age prediction as a domain-specific pretext task and analyze its adaptation to a brain segmentation downstream task, particularly in data-scarce scenarios. To achieve this, we explore established architectures, such as the UNET[19], and more modern transformer architectures that leverage self-attention mechanisms to capture long-range spatial dependencies [23,9]. Transformers have demonstrated strong performance in

various medical imaging applications [23,9], but not in the context of voxel-level brain age prediction. which has the potential to enhance feature extraction and improve the understanding of complex structural changes associated with brain aging. The main contribution of this paper is to show that voxel-level brain age prediction can serve as an effective SSL pretext task for brain MRI segmentation.

2 Methods

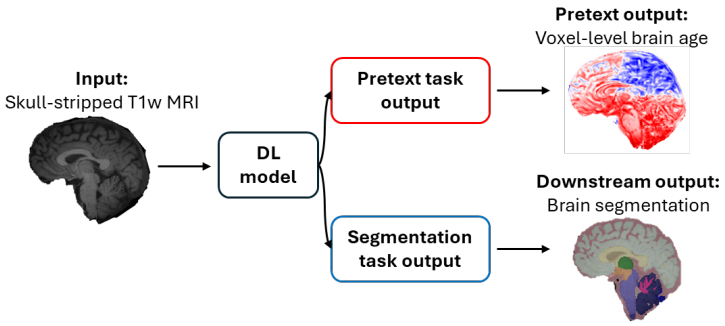

Fig. 1. Flowchart of the proposed methodology. The input to the DL model is a skull-stripped T1W MRI. The model is originally pre-trained on the pretext task of voxel-level brain age prediction. Then, the output layer of the pretrained model is replaced by a segmentation task prediction layer, and the model is fine-tuned.

2.1 Proposed Methodology

Our proposed methodology consists of two stages (Fig. 1): (1) pretaining the model on the voxel-level brain age prediction task; and (2) Replacing the model prediction layer, and fine-tuning the model for the brain MRI segmentation task.

Voxel-level brain age prediction pretext task: Voxel-level brain age prediction provides insights into the different aging patterns of various brain regions in healthy and diseased conditions, allowing for a more fine-grained analysis of neurodegenerative processes. To our knowledge, only two studies have implemented this approach. Popescu et al. [18] adapted a 3D U-Net, originally designed for segmentation, to predict localized brain age, capturing spatial patterns of brain aging. However, their model required white matter and gray matter segmentation as inputs to their model. Therefore, their approach is not suitable for SSL pretaining since it requires segmentation. Building upon this, Gianchandani et al. [8] extended the U-Net model to incorporate multitask learning, integrating segmentation and global-level brain age prediction tasks in addition to the voxel-level brain age prediction. Since their model also requires segmentation masks for training, it is not suitable for SSL pre-training. We propose

an approach similar to [8], but our model only has the voxel-level brain age prediction task. The rest of the model details follow the exact details as in [8].

Segmentation downstream task: To adapt the model pre-trained on the pretext task, we replaced the voxel-level brain age prediction output layer with a new output layer that accommodates the segmentation output requirements, and we fine-tuned the entire model. While investigating different ways to adapt the pre-trained model to the downstream task, we attempted to freeze the encoder and train only the decoder and output layers and train only the output layer. However, these approaches resulted in poorer performance than training the entire model with a low learning rate.

2.2 Baselines for Comparison

Training from scratch: This baseline for comparison corresponds to traditional supervised learning, where the model is trained from scratch using varying amounts of training data as it will be described in the experiments section.

Image inpainting as a pretext task: Self-supervised learning presents a promising solution to the challenges posed by limited annotated data in medical imaging. To ensure a fair comparison between a well-established, non-domain-specific task and our domain-specific voxel-level brain age prediction, we selected image inpainting, which is a widely used technique that enables models to reconstruct missing or occluded regions by leveraging contextual information from surrounding areas. This process enhances the model's understanding of image structures, improving its performance in downstream tasks such as segmentation[4].

In this study, we build upon previous works [4,22] by implementing a structured inpainting strategy with a random sampling approach for mask generation, ensuring diverse missing regions across training samples. To simulate missing voxel regions, we applied coarse dropout, introducing 12 "holes" of size $32 \times 32 \times 32$ within each input volume. This dropout covered 15% of the total input space, with input image dimensions set to be patches of dimensions $128 \times 160 \times 128$. By reconstructing these occluded regions, the model learns to capture spatial dependencies more effectively, facilitating robust feature extraction from brain MRI scans. To enhance reconstruction quality, we incorporated perceptual loss using AlexNet [25] as our loss function. This approach leverages feature-space representation to refine inpainting predictions, improving the model's ability to reconstruct missing regions more accurately. These enhancements contribute to more meaningful and informative representations, which are particularly valuable in medical imaging applications, where precise feature learning is critical for downstream tasks such as segmentation and classification.

2.3 Datasets

We compiled a dataset of normative T1-weighted MRI scans from individuals aged 18 to 80 obtained from multiple publicly available datasets. These include the Consortium for Reliability and Reproducibility (CORR) (https://fcon_

1000.projects.nitrc.org/indi/CoRR/html/) [27], the Information eXtraction from Images (IXI) dataset (https://brain-development.org/ixi-dataset/), the Autism Brain Imaging Data Exchange (ABIDE I and II) (https://fcon_1000.projects.nitrc.org/indi/abide/) [3,2], the Open Access Series of Imaging Studies (OASIS-1) (https://sites.wustl.edu/oasisbrains/home/oasis-1/) [14], the Alzheimer's Disease Neuroimaging Initiative (ADNI)[11,16,17] (https://adni.loni.usc.edu/), the Cambridge Centre for Ageing and Neuroscience (Cam-CAN) (https://camcan-archive.mrc-cbu.cam.ac.uk/dataaccess/) [20,24], Calgary Campinas dataset (https://sites.google.com/view/calgary-campinas-dataset/home) [21], and the Calgary Normative Study [15]. To ensure that the models during pretext training learn appropriate representations across all age and sex groups, we carefully curated the dataset to be well-balanced, maintaining an equal distribution of demographic variables across the training, validation, and testing sets. This step was critical in preventing model bias toward any particular group, ensuring fair and reliable predictions across the entire population.

The dataset used to train the pretext task, either voxel-level brain age prediction or image inpainting, consists of 1,710 MRI scans. These were divided into 1,025 samples for training, 428 for validation, and 257 for testing. For the segmentation downstream task, a second dataset combination was used, where training was conducted on three different subsets containing 78, 52, and 26 samples, respectively, while the validation and testing sets each contained 52 samples. Similar to the data used for training the pretext task, we ensured that this dataset maintained a balanced distribution across age and sex groups, allowing the segmentation model to generalize well across different demographics. To further evaluate the segmentation model's robustness, we tested it on an independent dataset not used during training, consisting of 127 MRI scans from OASIS-2 (https://sites.wustl.edu/oasisbrains/home/oasis-2/) [13]. This out-of-distribution test set was specifically included to assess the model's ability to generalize to previously unseen data from a different population and imaging protocol, providing insight into its real-world applicability.

Since some of the datasets used in this study contain longitudinal data, additional precautions were taken to prevent potential biases in model predictions. To mitigate the risk of data leakage and model overfitting, only a single time point was used per individual, ensuring that all models were trained and evaluated on independent samples. There was no data overlap across samples used for pretext and downstream tasks training and testing.

To ensure consistency across datasets, we preprocessed the data. All images were converted to NIfTI format, resized to a $256\times256\times256$ voxel grid and resampled to 1 mm isotropic. We used SynthStrip [10] to obtain skull-stripped images and SynthSeg [1] to obtain the brain MRI segmentation labels. Visual inspection was performed and samples with poor segmentation quality were excluded from the data. Table 1 summarizes the data used in this work.

Table 1. Summary of the datasets. For the segmentation downstream task, we had an additional out-of-distribution (OOD) test set.

Task	Train	Validation	Test	OOD Test
Pretext	1,025	428	257	-
Segmentation	78/52/26	52	52	127

2.4 Experiments

Ablation study: We initially performed an ablation study to search for the best architecture for the voxel-level brain age prediction task. We compared three commonly used architectures: UNET, UNETR, and Swin UNETR. UNET was chosen due to its extensive use in medical image analysis and prior application to this problem [8,7,18]. To explore the potential of transformer-based architectures, we included UNETR and Swin UNETR, which leverage self-attention mechanisms to capture long-range dependencies and hierarchical features [9,23,12]. Our goal was to identify the best-performing model for voxel-level brain age prediction as a pretext task, which would then be fine-tuned for segmentation to benefit from its pre-training. For training, we followed the procedure described in [8]. We used the Adam optimizer with a learning rate of 1e-3 and weight decay of 1e-4. A StepLR scheduler was applied, with a step size of 150 epochs for Swin UNETR and 70 epochs for UNET and UNETR, both using a decay factor of $\gamma = 0.6$. All models were trained for 600 epochs. Mean absolute error (MAE) and the coefficient of determination (R^2) were used to evaluate the results of the different architectures.

Segmentation Task: We trained Swin UNETR, the best-performing model from the voxel-level brain age prediction task (see Table 2), for segmentation using ground truth labels of 32 brain structures generated by SynthSeg[1] as a proof of concept. To ensure reliability, visual inspection was conducted, and poor segmentations were excluded. To evaluate the impact of pre-training, we compared a model trained from scratch, a model pre-trained on inpainting, and a model pre-trained on voxel-level brain age prediction as a domain-specific pretext task. The analysis was conducted across different training set sizes (78, 52, and 26 volumes). We validated and tested the model on 52 samples each from the in-distribution dataset. To ensure that training the pretext task and segmentation models on this diverse combination of publicly available datasets leads to a robust and generalizable model, we included an independent (out-of-distribution) dataset of 127 samples for further evaluation. During fine-tuning, we applied a low learning rate to preserve the pre-trained features while allowing the model to adjust to the new task. Specifically, we used the Adam optimizer with a learning rate of 1e-3 and a weight decay of 1e-4, along with a StepLR scheduler configured with a step size of 70 epochs and a decay factor of 0.6. The model was fine-tuned for 200 epochs with early stopping, leveraging the learned voxel-level brain age prediction or the inpainting representations to enhance segmentation performance. All models were trained using the cross-entropy loss function, and their performance was evaluated using the Dice score metric.

3 Results and Discussion

The results for the voxel-level brain age prediction ablation study are reported in Table 2. The Swin UNETR model obtained the lowest MAE error, followed by the UNET and the UNETR models. The Swin UNETR and UNET models obtained the same R^2 score. This is an interesting finding since past voxel-level brain age prediction works used the UNET architecture [18,8].

Table 2. Voxel-level brain age prediction task results.

Model	MAE (years)	R^2
UNET	6.15 ± 4.2	**0.84**
UNETR	7.17 ± 4.4	0.81
Swin UNETR	**5.86 ± 4.4**	**0.84**

For the segmentation downstream task, we selected the Swin UNETR architecture based on our ablation study results. The segmentation results for the Swin UNETR model trained from scratch, pre-trained on the inpainting or the voxel-level brain age prediction pretext task for different amounts of training data are summarized in Fig. 2, and sample segmentation masks are shown in Fig. 3. In the in-distribution test set, all models perform similarly when trained on 78 samples, but the model trained from scratch achieved a slightly higher Dice score. However, when the models were trained with 52 or 26 samples, both pre-trained models (inpainting and voxel-level brain age prediction) outperformed training from scratch. The model pre-trained on inpainting outperformed the model trained on voxel-level brain age prediction on the in-distribution test set when the model was trained on 52 and 78 samples. In contrast, with only 26 training samples, the model pre-trained on voxel-level brain age prediction achieved a slightly higher Dice score than the model pre-trained on the inpainting, highlighting the advantage of domain-specific pre-training in low-data scenarios.

All models were evaluated on an out-of-distribution test set. While the model pre-trained on the inpainting task performed the worst at 78 training samples, the model trained from scratch achieved the highest score. However, as the training set size decreased to 52 and 26 samples, the brain age pre-trained model consistently outperformed all other models on the out-of-distribution test set. These results underscore the effectiveness of domain-specific pre-training in low-data scenarios and demonstrate that training on a diverse, publicly available normative T1-weighted dataset enhances the model's generalization capability.

One limitation of this study is the use of SynthSeg-generated segmentation labels as ground truth. To mitigate this, we visually inspected and excluded poor-quality segmentations. Despite this limitation, SynthSeg labels were suitable for this proof-of-concept study, as our primary goal was to demonstrate the importance of domain-specific pre-training. In future work, we plan to use expert-annotated segmentation labels to validate our results on more accurate ground truth data. Additionally, we aim to develop new domain-specific pretext tasks

and compare their effectiveness to voxel-level brain age prediction in improving segmentation. This study underscores the need for task-specific SSL approaches in medical imaging, particularly in data-limited settings.

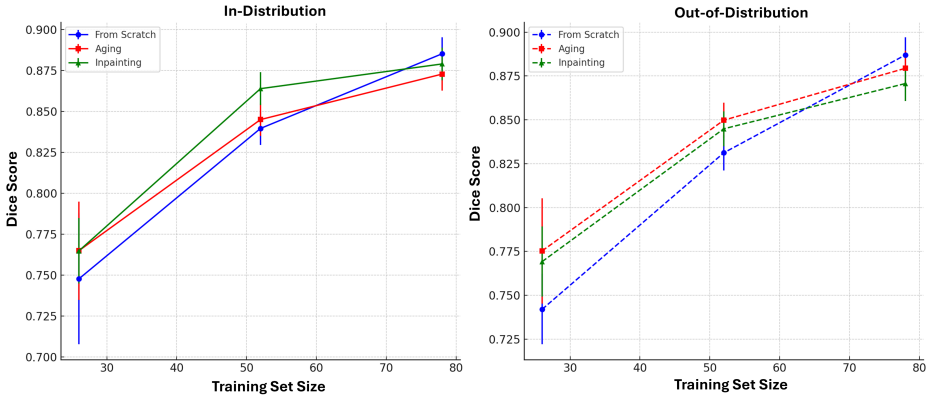

Fig. 2. Segmentation performance (Dice score) across different training set sizes for models trained from scratch, pre-trained on inpainting, and voxel-level brain age prediction. In-distribution (left) and out-of-distribution (right) test set results.

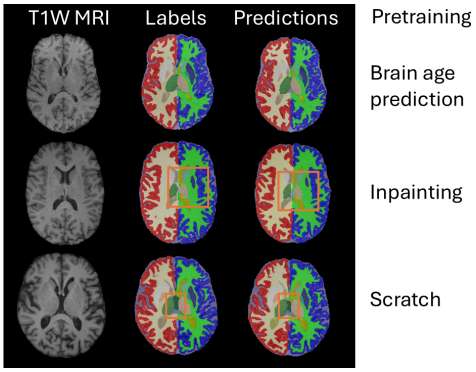

Fig. 3. Comparison of segmentation outputs for representative samples of the in-distribution test set for the models developed with 52 samples.

4 Conclusion

This study demonstrated the effectiveness of voxel-level brain age prediction as a domain-specific pretext task for self-supervised learning in neuroimaging. Our

results show that Swin UNETR, pre-trained on voxel-level brain age prediction, outperforms models trained from scratch and those pre-trained on inpainting in brain segmentation tasks, particularly in low-data scenarios and out-of-distribution data. These findings highlight the potential of brain age prediction for improving downstream medical imaging tasks.

Acknowledgments. This work was supported by research funding from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canada Research Chairs Program. We also acknowledge the Digital Research Alliance of Canada for providing computational infrastructure support.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Billot, B., Greve, D.N., Puonti, O., Thielscher, A., Van Leemput, K., Fischl, B., Dalca, A.V., Iglesias, J.E., et al.: SynthSeg: Segmentation of brain MRI scans of any contrast and resolution without retraining. Medical image analysis **86**, 102789 (2023)
2. Di Martino, A., O'connor, D., Chen, B., Alaerts, K., Anderson, J.S., Assaf, M., Balsters, J.H., Baxter, L., Beggiato, A., Bernaerts, S., et al.: Enhancing studies of the connectome in autism using the autism brain imaging data exchange II. Scientific Data **4**(1), 1–15 (2017)
3. Di Martino, A., Yan, C.G., Li, Q., Denio, E., Castellanos, F.X., Alaerts, K., Anderson, J.S., Assaf, M., Bookheimer, S.Y., Dapretto, M., et al.: The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Molecular psychiatry **19**(6), 659–667 (2014)
4. Dominic, J., Bhaskhar, N., Desai, A.D., Schmidt, A., Rubin, E., Gunel, B., Gold, G.E., Hargreaves, B.A., Lenchik, L., Boutin, R., et al.: Improving data-efficiency and robustness of medical imaging segmentation using inpainting-based self-supervised learning. Bioengineering 10 (2), 207 (2023) (2023)
5. Elharrouss, O., Almaadeed, N., Al-Maadeed, S., Akbari, Y.: Image inpainting: A review. Neural Processing Letters **51**, 2007–2028 (2020)
6. Franke, K., Gaser, C.: Ten years of BrainAGE as a neuroimaging biomarker of brain aging: what insights have we gained? Frontiers In Neurology **10**, 789 (2019)
7. Gianchandani, N., Dibaji, M., Bento, M., MacDonald, E., Souza, R.: Reframing the brain age prediction problem to a more interpretable and quantitative approach. ArXiv Preprint ArXiv:2308.12416 (2023)
8. Gianchandani, N., Ospel, J., MacDonald, E., Souza, R.: A multitask deep learning model for voxel-level brain age estimation. In: International Workshop on Machine Learning in Medical Imaging. pp. 283–292. Springer (2023)
9. Hatamizadeh, A., Tang, Y., Nath, V., Yang, D., Myronenko, A., Landman, B., Roth, H.R., Xu, D.: Unetr: Transformers for 3d medical image segmentation. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision. pp. 574–584 (2022)
10. Hoopes, A., Mora, J.S., Dalca, A.V., Fischl, B., Hoffmann, M.: SynthStrip: skull-stripping for any brain image. NeuroImage **260**, 119474 (2022)

11. Jack Jr, C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., L. Whitwell, J., Ward, C., et al.: The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine **27**(4), 685–691 (2008)
12. Kim, J., Kim, M., Park, H.: Domain aware multi-task pretraining of 3D Swin transformer for T1-weighted brain MRI. In: Computer Vision – ACCV 2024. pp. 121–141. Springer, Springer Nature Singapore (2025)
13. Marcus, D.S., Fotenos, A.F., Csernansky, J.G., Morris, J.C., Buckner, R.L.: Open access series of imaging studies: longitudinal MRI data in nondemented and demented older adults. Journal of cognitive neuroscience **22**(12), 2677–2684 (2010)
14. Marcus, D.S., Wang, T.H., Parker, J., Csernansky, J.G., Morris, J.C., Buckner, R.L.: Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. Journal of cognitive neuroscience **19**(9), 1498–1507 (2007)
15. McCreary, C.R., Salluzzi, M., Andersen, L.B., Gobbi, D., Lauzon, L., Saad, F., Smith, E.E., Frayne, R.: Calgary Normative Study: design of a prospective longitudinal study to characterise potential quantitative MR biomarkers of neurodegeneration over the adult lifespan. BMJ open **10**(8), e038120 (2020)
16. Mueller, S.G., Weiner, M.W., Thal, L.J., Petersen, R.C., Jack, C.R., Jagust, W., Trojanowski, J.Q., Toga, A.W., Beckett, L.: Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI). Alzheimer's & Dementia **1**(1), 55–66 (2005)
17. Petersen, R.C., Aisen, P.S., Beckett, L.A., Donohue, M.C., Gamst, A.C., Harvey, D.J., Jack Jr, C., Jagust, W.J., Shaw, L.M., Toga, A.W., et al.: Alzheimer's disease Neuroimaging Initiative (ADNI) clinical characterization. Neurology **74**(3), 201–209 (2010)
18. Popescu, S.G., Glocker, B., Sharp, D.J., Cole, J.H.: Local brain-age: a U-net model. Frontiers in Aging Neuroscience **13**, 761954 (2021)
19. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. pp. 234–241. Springer (2015)
20. Shafto, M.A., Tyler, L.K., Dixon, M., Taylor, J.R., Rowe, J.B., Cusack, R., Calder, A.J., Marslen-Wilson, W.D., Duncan, J., Dalgleish, T., et al.: The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. BMC neurology **14**, 1–25 (2014)
21. Souza, R., Lucena, O., Garrafa, J., Gobbi, D., Saluzzi, M., Appenzeller, S., Rittner, L., Frayne, R., Lotufo, R.: An open, multi-vendor, multi-field-strength brain MR dataset and analysis of publicly available skull stripping methods agreement. NeuroImage **170**, 482–494 (2018)
22. Stoean, C., Bacanin, N., Stoean, R., Ionescu, L., Alecsa, C., Hotoleanu, M., Atencia, M., Joya, G.: On using perceptual loss within the U-net architecture for the semantic inpainting of textile artefacts with traditional motifs. In: 2022 24th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC). pp. 276–283. IEEE (2022)
23. Tang, Y., Yang, D., Li, W., Roth, H.R., Landman, B., Xu, D., Nath, V., Hatamizadeh, A.: Self-supervised pre-training of swin transformers for 3d medical image analysis. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 20730–20740 (2022)

24. Taylor, J.R., Williams, N., Cusack, R., Auer, T., Shafto, M.A., Dixon, M., Tyler, L.K., Henson, R.N., et al.: The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) data repository: Structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample. neuroimage **144**, 262–269 (2017)
25. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 586–595 (2018)
26. Zheng, C.: Self-supervised pretext tasks for Alzheimer's Disease classification using 3D convolutional neural networks on large-scale synthetic neuroimaging dataset. ArXiv Preprint ArXiv:2406.14210 (2024)
27. Zuo, X.N., Anderson, J.S., Bellec, P., Birn, R.M., Biswal, B.B., Blautzik, J., Breitner, J., Buckner, R.L., Calhoun, V.D., Castellanos, F.X., et al.: An open science resource for establishing reliability and reproducibility in functional connectomics. Scientific data **1**(1), 1–13 (2014)