Efficient Atrophy Mapping: A Single-Step U-net Approach for Rapid Brain Change Estimation

1st Riccardo Raciti
Dept. of Math and Computer Science
University of Catania
Catania, Italy
riccardo.raciti@studium.unict.it

2nd Alessia Rondinella
Dept. of Math and Computer Science
University of Catania
Catania, Italy
alessia.rondinella@phd.unict.it

3rd Lemuel Puglisi
Dept. of Math and Computer Science
University of Catania
Catania, Italy
lemuel.puglisi@phd.unict.it

4th Francesco Guarnera
Dept. of Math and Computer Science
University of Catania
Catania, Italy
francesco.guarnera@unict.it

5th Daniele Ravì
School of Physics, Engineering and CS
University of Hertfordshire
Hatfield, UK
d.ravi@herts.ac.uk

6th Sebastiano Battiato
Dept. of Math and Computer Science
University of Catania
Catania, Italy
battiato@dmi.unict.it

***Abstract*—The estimation of brain atrophy is crucial for evaluating brain diseases and analyzing neurodegeneration. Existing methods for computing atrophy maps often suffer from lengthy processing times due to the computational cost of multi-step processing. In this work, we propose a novel technique for atrophy map calculation using a single U-net-based architecture. This approach consolidates multiple traditional medical imaging processing steps into a single process, aiming to accelerate the computational time required. Specifically, our method estimates structural changes by generating a flow map from two longitudinal Magnetic Resonance Imaging (MRI) scans of the same subject. We trained and evaluated our system on a dataset comprising 2000 T1-weighted MRI scans sourced from two different public datasets on Alzheimer's Disease. Experimental results demonstrate a considerable reduction in execution time while maintaining atrophy mapping performance comparable to state-of-the-art solutions. We believe that our pipeline could significantly benefit clinical applications for measuring brain atrophy, especially in scenarios requiring the evaluation of large cohorts, such as clinical trials. Our code is freely available at https://github.com/Efficient Atrophy Mapping: A Single-Step U-net Approach for Rapid Brain Change Estimation.**

***Index Terms*—Atrophy map, Medical Imaging, MRI**

I. Introduction

Brain atrophy refers to the gradual loss of brain tissue, causing a consequent decline in brain function [1]. While it can be caused by the normal aging process, it also plays a pivotal role in the pathophysiology and clinical manifestation of neurodegenerative diseases. This progressive loss of neuronal tissue can affect various brain regions, leading to diverse symptomatology depending on the specific disease and affected areas. In Alzheimer's disease, for instance, neuronal depletion in the hippocampus and cerebral cortex correlates strongly with memory deficits and cognitive decline [2], [3]. Parkinson's disease is characterized by the degeneration of dopaminergic neurons in the substantia nigra, resulting in movement disorders such as tremors and rigidity [4]. Multiple sclerosis, conversely, involves demyelination and neuronal loss, manifesting in symptoms like coordination problems, muscle weakness, and cognitive dysfunction [5], [6]. Given the significant impact of brain atrophy on patient outcomes, accurate measurement and quantification through MRI scans have become crucial for the diagnosis, monitoring, and management of neurodegenerative conditions [7]–[9].

Advanced neuroimaging techniques and analytical methods have been developed to address the need for atrophy quantification. Conventional approaches include voxel-based morphometry, which generates statistical parametric maps to assess regional differences in brain tissue between two time points, thereby precisely quantifying volumetric changes [10]. Other traditional computational methods, such as Deformation-Based Morphometry (DBM) [11] and Boundary Shift Integral (BSI) [12], compare image intensities to derive atrophy measures, but they often misinterpret these differences as changes in brain volume, leading to increased variance in measurements. Another method called Structural Image Evaluation using Normalization of Atrophy (SIENA) [13], is highly accurate but is limited by long processing times, which can hinder its practical usage. Recently, deep learning techniques, have been increasingly explored to enhance the accuracy of analyzing MRI scans for diagnostic purposes in this context [14], [15]. These innovative methods leverage the power of data-driven solutions to automatically learn intricate patterns from neuroimaging data, offering the potential to significantly improve our understanding of brain atrophy and its implications. Ultimately, advancements in brain atrophy quantification [16] could benefit clinical practice and patient

care by enabling earlier diagnosis, more precise monitoring of disease progression, and improved evaluation of therapeutic interventions [17].

Building upon recent advancements in deep learning, we propose a novel approach to address the limitations of traditional computational methods for atrophy quantification. Our methodology involves training a 3D U-Net architecture to estimate the flow field between two MRI scans of the same patient. Subsequently, we utilize the predicted flow field to quantify brain atrophy. While conventional techniques for flow field computation, such as those employed in the SIENA, are computationally intensive, our end-to-end prediction approach offers a substantial acceleration of the atrophy quantification process. We evaluate our method against SIENA, emphasizing the significant reduction in execution time. To assess the quality of the generated flow fields, we calculate the Percentage of Brain Volume Change (PBVC) using our estimated flow field and compare it to that derived from SIENA. To the best of our knowledge, this study represents the first implementation of an end-to-end 3D U-Net architecture specifically designed for efficient atrophy quantification. We posit that this enhancement will facilitate the atrophy quantification process in large-scale neuroimaging datasets, potentially accelerating research in neurodegenerative disorders.

The remainder of this paper is organized as follows: Section II reviews the state-of-the-art in the field, Section III describes the proposed approach with details about the dataset and model, Section IV presents qualitative and quantitative results, and finally, Section V concludes the paper with considerations about this work and future directions.

II. Related work

This section reviews methodologies and models relevant to our study, highlighting their strengths and limitations. Initial work in this area primarily focused on just diagnostic prediction using longitudinal data but they could not provide spatial maps or segmentations of affected regions, which are crucial for interpreting and understanding disease progression. For example, Recurrent Neural Networks (RNNs) and convolutional recurrent networks were employed for diagnostic prediction using MRI [18] or PET scans [19] in Alzheimer's Disease (AD). However, a main issue with these early approaches is their sensitivity to the irregular temporal spacing between instances, which is common in disease modeling.

To address this issue, mixed-effects models were proposed to explicitly describe each patient's progression, allowing sampling at any time point [20], [21]. Through temporal reparametrization, all patients are aligned on a common pathological timeline, with individual trajectories modeled as variations around a reference trajectory. While mixed-effects models have become a standard in longitudinal modeling, providing insights into disease progression, they typically rely on explicit labels, such as the patient's age at each visit.

In an effort to reduce the need for labeled data, several self-supervised approaches have been proposed. These approaches often combine deep learning techniques with mixed-effects models for longitudinal modeling of disease progression. One such approach [22] utilizes an RNN to produce the parameters of a mixed-effects model, describing patients' trajectories as straight lines in the latent space of a Variational Autoencoder (VAE). By leveraging the unsupervised learning capabilities of VAEs, this method can learn meaningful representations of disease progression without relying on explicit labels like the patient's age.

Other self-supervised methods, such as [23] and [24], employ VAE encoders to learn latent temporal and spatial variables, separating temporal progression from the patient's intrinsic characteristics. Longitudinal VAE architectures, like Gaussian Process VAEs (GP-VAE) [25] were also introduced to provide a more general prior for the posterior distribution in the form of a Gaussian Process dependent on age and covariates [26], [27]. Unlike the above methods that focus on modeling disease progression and trajectories, another category of approaches aims to generate spatial maps or segmentations of the affected regions, complementing the insights from longitudinal modeling. Diffeomorphic methods, such as those based on geodesic regression [28], [29], learn deformation maps to model the effect of time on a subject's images. However, these methods exhibit limited long-term predictive abilities compared to mixed-effects models, which aggregate information from subjects at different disease stages [30]. In the context of neuroimaging, where quantifying brain atrophy is crucial for understanding and monitoring neurodegenerative diseases like multiple sclerosis and Alzheimer's disease, SIENA [13] is a well-regarded computational method that calculates a flow map. This flow map shows the displacement of voxels over time, indicating areas of increased or decreased atrophy in the brain. While SIENA is not based on deep learning, its ability to generate this flow map sets it apart from other traditional methods. However, SIENA's reliance on computational modeling leads to long execution times, as the flow map calculation depends on preliminary results obtained during the model's execution.

This limitation has motivated the exploration of deep learning techniques, which have shown promise in medical image segmentation tasks. The introduction of the U-Net architecture [31] has been a significant development in this regard, with its distinctive "skip connections" allowing direct information flow from lower to higher levels, enabling the model to maintain crucial spatial details for accurate segmentation. Compared to other deep learning architectures, the U-Net is more compact, reducing the total number of parameters and improving computational efficiency and training speed.

Building upon the segmentation capabilities of [31], we propose to use a 3D U-Net to embed the results obtained by SIENA into an end-to-end deep neural network. Our model, along with SIENA, is among the few that allow for the calculation of this flow map, which is essential for quantifying brain atrophy. However, our approach circumvents

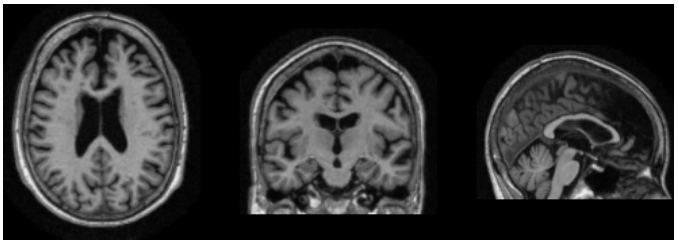

Fig. 1. An example of MRI scan from OASIS-3 [32]. From left to right, the figures show the axial, coronal and sagittal views.

the computational limitations of SIENA by eliminating the need for intermediate results in the flow map calculation. By leveraging the efficiency and accuracy of the 3D U-Net, our method aims to provide a more rapid and precise estimation of brain atrophy, overcoming the computational bottlenecks of traditional computational methods like SIENA.

III. Method

In this section, we first describe the dataset used in our experiments, then the preprocessing phase, and finally the implementation of the proposed architecture.

A. Dataset

We collected 3D T1-weighted MRIs from two public datasets, ADNI [33] and OASIS-3 [32]. We selected 1000 pairs (MRI_A and MRI_B), focusing on images with significant atrophy as measured by Eq. 1. This selection ensures that the dataset includes cases with substantial changes in brain structure, which is critical for training models aimed at detecting and analyzing neurodegenerative conditions. By selecting images with greater atrophy, we aim to enhance the model's ability to recognize and quantify significant pathological changes, ultimately improving its diagnostic accuracy and robustness.

$$d = \frac{|MRI_A - MRI_B|}{\max(MRI_A, MRI_B)} \tag{1}$$

The final dataset comprises patients aged from 55 to 97 years and is balanced in terms of gender (547 men and 453 women). In our experiment, we use 70% of the scans for the training set, 20% for the validation set, and 10% for the test set. An example of an input MRI is shown in Fig. 1.

To assess our pipeline, we use an intermediate output of SIENA as ground truth (GT). As mentioned above, SIENA is a well-known approach in neuroimaging that estimates the PBVC from MRI_A to MRI_B. Below is a brief summary of the steps implemented by SIENA:

- Brain Extraction Tool (BET): This method is used to perform brain extraction on both MRIs.
- Skull Extraction: This step is used for both skull removal and MRI alignment.
- Registration: This involves the use of FMRIB's Linear Image Registration Tool (FLIRT) for aligning the images.
- Masking: This step is used to obtain better-registered images by applying masks to the images.
- Change Analysis: This step extracts the motion field to analyze changes between the two MRIs.

An example of the output of SIENA is shown in Fig. 2. As shown in the figure, SIENA processes two MRI scans taken at different times and generates a flow field. The second MRI, compared to the first, reveals a deteriorated health condition. This is evidenced by the increased dark regions in the frontal lobe and the expansion of the dark areas in the central brain region, corresponding to the ventricles.

The flow field creates an atrophy map (our GT) showing how each voxel changes from MRI_A to MRI_B. Black areas in the GT indicate an increase in atrophy, while white areas indicate a decrease. In our study, we do not account for a decrease in atrophy, since we carry out its calculation considering two MRIs where the first precedes the second.

B. Preprocessing

In our study, a preprocessing step is applied to all MRI scans to ensure data consistency and eliminate irrelevant variations. The preprocessing pipeline comprises several key stages. Initially, all MRI scans undergo registration to the Montreal Neurological Institute (MNI) template, ensuring spatial alignment across different subjects for accurate inter-subject comparison. Following registration, non-brain tissues are removed from the images through skull stripping. Subsequently, intensity normalization is performed to standardize pixel values across all images within the range of -1 to 1. This normalization mitigates variations stemming from acquisition parameters or scanner discrepancies. Finally, each MRI is rescaled to $121 \times 145 \times 113$ voxels, reducing the size by half along each axis. This rescaling minimizes computational costs associated with processing whole scans and expedites data processing without significant loss of anatomical detail.

C. CNN architecture

In this section, we detail our pipeline. The objective of our solution is to estimate brain changes through a single-step method. To achieve this, we employed a 3D U-Net architecture (Fig. 3), using the implementation available from

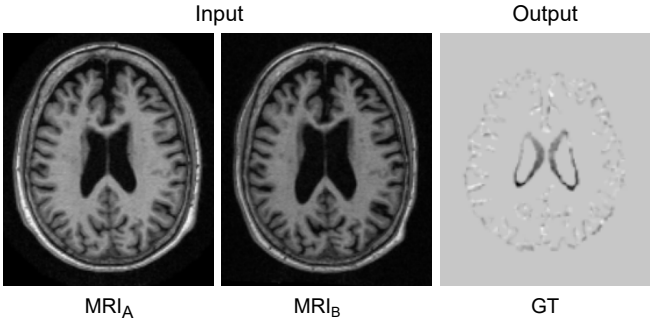

Fig. 2. The first two images show the central slices of the MRIs (baseline and follow-up) used as input to SIENA. The last image displays the output (our ground truth) obtained from SIENA.

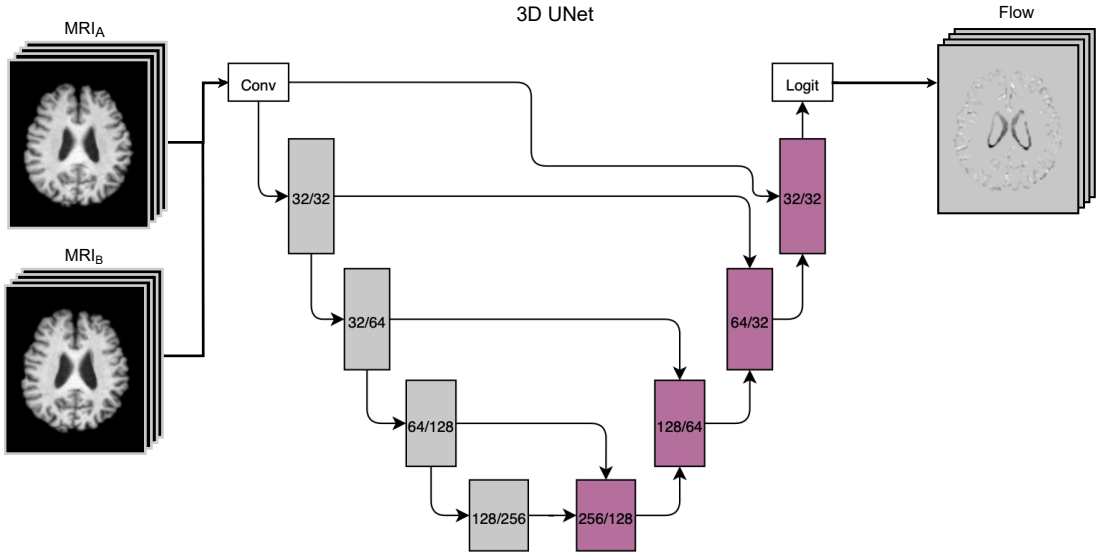

Fig. 3. 3D U-Net architecture employed. A and B represent MRI_A and MRI_B, respectively. The rectangle labeled Conv represents a two-channel convolution, while each gray-colored rectangle represents the layers used in the downsampling phase, with the number of input and output features indicated inside. Pink-colored rectangles represent the part of the network responsible for the upsampling phase, with concatenation to the encoder feature map. Again, the values inside indicate the number of features in input and output. The output of the last upsampling layer is then passed to a Logit layer, which will provide the final prediction.

MONAI framework [34]. Our input-output model can be mathematically summarized as follows:

$$f_\theta : \mathcal{R}^{2\times H\times W\times D} \rightarrow \mathcal{R}^{H\times W\times D} \tag{2}$$

where *H*, *W* and *D* represent height, width, and depth.

For the training phase, various configurations were tested using the AdamW optimizer [35] with different learning rates, batch sizes, and weight decay settings as shown in Table I. In the last row, the square brackets indicate the maximum and minimum learning rate values used with the decay scheduler. The loss function utilized is the following:

$$\mathcal{L} = \max_{i=1,\ldots,\mathcal{N}} (\hat{y}_i - y_i)^2 \tag{3}$$

Where $\mathcal{N}$ corresponds to the batch size, $\hat{y}_i$ represents the model's prediction, and y_i denotes the GT. Three different methods of batch loss calculation were also considered: mean, max, and sum.

IV. Result

This section describes the results obtained from the proposed approach. The first part showcases the quantitative results, comparing the execution times with respect to SIENA method, as well as the evaluation metrics used. The second part is focused on the qualitative results displaying also a visual comparison with the GT. We want to note that SIENA does not support GPU usage; therefore, the SIENA values will be reported without specifying the device (only CPU).

For our experiments, we used an NVIDIA A100 GPU and an Intel(R) Xeon(R) CPU E5-2650 v4 @ 2.20GHz.

The best configuration of our approach (the fifth model in Table I) achieves an average Mean Square Error (MSE) of 0.0008 and an average Structural Similarity Index Measure (SSIM) of 0.88. From this point onward, when we refer to our model, we mean this specific model.

Table II shows the computation time required by different approaches. Given a pair of MRIs (MRI_A, MRI_B), the processing time using our method averages 0.05 seconds with a GPU and 2.98 seconds with a CPU. From the table, it is evident that our solution is more efficient than SIENA, which requires instead 840 seconds.

In Fig. 4, we present an example comparing our model's results to the ground truth (GT) obtained using SIENA. The similarity between them indicates that our approach provides good qualitative performance. Upon closer inspection, the GT images (Fig. 4, first row) contain only some lighter or white spots. This example gives us an indication of a close match with SIENA and therefore the effectiveness of our solution. We

Learning Rate	Batch Size	Loss Reduction	MSE	SSIM
0.001	4	mean	0.001	0.83
0.001	4	sum	0.001	0.84
0.001	8	mean	0.003	0.71
0.01	4	max	0.0012	0.77
0.001	4	max	0.0008	0.88
[0.001, 0.01]	4	max	0.0014	0.83

TABLE I
THIS TABLE PRESENTS THE RESULTS OF THE MODELS TESTED WHILE VARYING THE LEARNING RATE, BATCH SIZE, AND LOSS FUNCTION REDUCTION, ALSO INCLUDING THE AVERAGE MSE AND SSIM VALUES CALCULATED ON THE TEST SET.

Method	Preprocessing time Registration + skull rem.	Processing time
Our Model CPU	190 s ± 3 s	∼2.98 s
Our Model GPU	190 s ± 3 s	∼0.05 s
SIENA	190 s ± 3 s	∼840 s

TABLE II
IN THIS TABLE, THE EXECUTION TIMES OF OUR BEST MODEL AND SIENA ARE REPORTED, WITH EXECUTION USING CPU AND GPU. THE EXECUTION TIMES ARE EXPRESSED IN SECONDS.

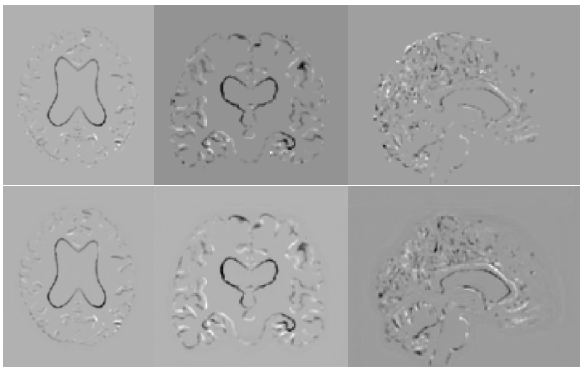

Fig. 4. The first row of these images shows the model's results, while the second row presents the GT results, which are the outputs from SIENA.

can conclude that, by leveraging deep learning, we developed a model that can compute the brain atrophy map in a very short time from two MRI scans. The main disadvantage of SIENA lies in its inability to utilize the GPU for inference, leading to significantly longer execution times. Notably, even with the use of only a CPU, our method is more than 300 times faster.

As further evidence for the quantitative assessment of our model, we compared the PBVC between SIENA (target PBVC) and our approach. Our results show a Pearson's correlation coefficient of 0.92, indicating a relatively high correlation, with a standard deviation of 1.49. The graph in Fig. 5 visually represents these results.

V. CONCLUSIONS AND FUTURE DEVELOPMENTS

In this study, we propose a novel deep-learning model based on a 3D U-Net capable of generating atrophy flow maps. These flow maps are useful for describing atrophy and assessing the progression of neurological diseases. We compare our method with SIENA, which is currently the most used method in the neuroimaging community for flow maps and PBVC estimation.

Our results demonstrate that we achieve outcomes comparable to SIENA while significantly optimizing execution times. Our model processes an MRI scan in approximately 3 minutes, compared to the 17 minutes required by SIENA for each pair of MRIs. Furthermore, our model shows strong similarity to SIENA when analyzed using MSE and SSIM. Additionally, the high Pearson's correlation coefficient further validates our model's accuracy in assessing PBVC.

Given the promising results of the model, we plan to explore its application in real-time clinical settings for monitoring the progression of neurological diseases such as Alzheimer's, multiple sclerosis, and Parkinson's disease. This approach could also be integrated into large-scale studies to provide quick and accurate assessments of brain atrophy across diverse populations.

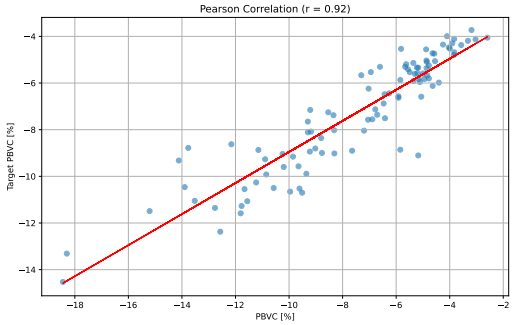

Fig. 5. Pearson's correlation between the PBVC values calculated by SIENA and those obtained by replacing SIENA's intermediate flow with that produced by our model: the value r in brackets indicating the resulting coefficient.

ACKNOWLEDGMENTS

Alessia Rondinella is a PhD candidate enrolled in the National PhD in Artificial Intelligence, XXXVII cycle, course on Health and life sciences, organized by Università Campus Bio-Medico di Roma. Francesco Guarnera is funded by the PNRR MUR project PE0000013-FAIR. Lemuel Puglisi is enrolled in a PhD program at the University of Catania, fully funded by PNRR (DM 118/2023).

REFERENCES

[1] NC Fox, RI Scahill, WR Crum, and MN Rossor. Correlation between rates of brain atrophy and cognitive decline in ad. *Neurology*, 52(8):1687–1687, 1999.
[2] Holger Jahn. Memory loss in alzheimer's disease. *Dialogues in clinical neuroscience*, 15(4):445–454, 2013.
[3] Laura A Rabin, Colette M Smart, and Rebecca E Amariglio. Subjective cognitive decline in preclinical alzheimer's disease. *Annual review of clinical psychology*, 13:369–396, 2017.
[4] Lorraine V Kalia and Anthony E Lang. Parkinson's disease. *The Lancet*, 386(9996):896–912, 2015.
[5] Ranjan Dutta and Bruce D Trapp. Pathogenesis of axonal and neuronal damage in multiple sclerosis. *Neurology*, 68(22_suppl_3):S22–S31, 2007.
[6] Ranjan Dutta and Bruce D Trapp. Mechanisms of neuronal dysfunction and degeneration in multiple sclerosis. *Progress in neurobiology*, 93(1):1–12, 2011.
[7] Giovanni B Frisoni, Nick C Fox, Clifford R Jack Jr, Philip Scheltens, and Paul M Thompson. The clinical use of structural mri in alzheimer disease. *Nature Reviews Neurology*, 6(2):67–77, 2010.
[8] Adam L Boxer, Michael D Geschwind, Nataliya Belfor, Maria Luisa Gorno-Tempini, Guido F Schauer, Bruce L Miller, Michael W Weiner, and Howard J Rosen. Patterns of brain atrophy that differentiate corticobasal degeneration syndrome from progressive supranuclear palsy. *Archives of neurology*, 63(1):81–86, 2006.
[9] Robert A Bermel and Rohit Bakshi. The measurement and clinical relevance of brain atrophy in multiple sclerosis. *The Lancet Neurology*, 5(2):158–170, 2006.
[10] John Ashburner and Karl J Friston. Voxel-based morphometry—the methods. *Neuroimage*, 11(6):805–821, 2000.

[11] Christian Gaser, Igor Nenadic, Bradley R. Buchsbaum, Erin A. Hazlett, and Monte S. Buchsbaum. Deformation-based morphometry and its relation to conventional volumetry of brain lateral ventricles in mri. *NeuroImage*, 13(6):1140–1145, 2001.

[12] P.A. Freeborough and N.C. Fox. The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat mri. *IEEE Transactions on Medical Imaging*, 16(5):623–629, 1997.

[13] Stephen M Smith, Yongyue Zhang, Mark Jenkinson, Jacqueline Chen, Paul M Matthews, Antonio Federico, and Nicola De Stefano. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *Neuroimage*, 17(1):479–489, 2002.

[14] Adrien Payan and Giovanni Montana. Predicting alzheimer's disease: a neuroimaging study with 3d convolutional neural networks. *arXiv preprint arXiv:1502.02506*, 2015.

[15] Mengjin Dong, Long Xie, Sandhitsu R. Das, Jiancong Wang, Laura E.M. Wisse, Robin deFlores, David A. Wolk, and Paul A. Yushkevich. Deepatrophy: Teaching a neural network to detect progressive changes in longitudinal mri of the hippocampal region in alzheimer's disease. *NeuroImage*, 243:118514, 2021.

[16] Maria A Rocca, Marco Battaglini, Ralph HB Benedict, Nicola De Stefano, Jeroen JG Geurts, Roland G Henry, Mark A Horsfield, Mark Jenkinson, Elisabetta Pagani, and Massimo Filippi. Brain mri atrophy quantification in ms: from methods to clinical application. *Neurology*, 88(4):403–413, 2017.

[17] Dallas P Veitch, Michael W Weiner, Paul S Aisen, Laurel A Beckett, Nigel J Cairns, Robert C Green, Danielle Harvey, Clifford R Jack Jr, William Jagust, John C Morris, et al. Understanding disease progression and improving alzheimer's disease clinical trials: Recent highlights from the alzheimer's disease neuroimaging initiative. *Alzheimer's & Dementia*, 15(1):106–152, 2019.

[18] Ruoxuan Cui, Manhua Liu, Alzheimer's Disease Neuroimaging Initiative, et al. Rnn-based longitudinal analysis for diagnosis of alzheimer's disease. *Computerized Medical Imaging and Graphics*, 73:1–10, 2019.

[19] Manhua Liu, Danni Cheng, Weiwu Yan, and Alzheimer's Disease Neuroimaging Initiative. Classification of alzheimer's disease by combination of convolutional and recurrent neural networks using fdg-pet images. *Frontiers in neuroinformatics*, 12:35, 2018.

[20] Igor Koval, Alexandre Bône, Maxime Louis, Thomas Lartigue, Simona Bottani, Arnaud Marcoux, Jorge Samper-Gonzalez, Ninon Burgos, Benjamin Charlier, Anne Bertrand, et al. Ad course map charts alzheimer's disease progression. *Scientific Reports*, 11(1):8020, 2021.

[21] Benoît Sauty and Stanley Durrleman. Riemannian metric learning for progression modeling of longitudinal datasets. In *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*, pages 1–5. IEEE, 2022.

[22] Maxime Louis, Raphaël Couronné, Igor Koval, Benjamin Charlier, and Stanley Durrleman. Riemannian geometry learning for disease progression modelling. In *Information Processing in Medical Imaging: 26th International Conference, IPMI 2019, Hong Kong, China, June 2–7, 2019, Proceedings 26*, pages 542–553. Springer, 2019.

[23] Raphaël Couronné, Paul Vernhet, and Stanley Durrleman. Longitudinal self-supervision to disentangle inter-patient variability from disease progression. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part II 24*, pages 231–241. Springer, 2021.

[24] Qingyu Zhao, Zixuan Liu, Ehsan Adeli, and Kilian M Pohl. Longitudinal self-supervised learning. *Medical image analysis*, 71:102051, 2021.

[25] Francesco Paolo Casale, Adrian Dalca, Luca Saglietti, Jennifer Listgarten, and Nicolo Fusi. Gaussian process prior variational autoencoders. *Advances in neural information processing systems*, 31, 2018.

[26] Matthew Ashman, Jonathan So, Will Tebbutt, Vincent Fortuin, Michael Pearce, and Richard E Turner. Sparse gaussian process variational autoencoders. *arXiv preprint arXiv:2010.10177*, 2020.

[27] Siddharth Ramchandran, Gleb Tikhonov, Kalle Kujanpää, Miika Koskinen, and Harri Lähdesmäki. Longitudinal variational autoencoder. In *International Conference on Artificial Intelligence and Statistics*, pages 3898–3906. PMLR, 2021.

[28] Stefan Bauer, Christian May, Dimitra Dionysiou, Georgios Stamatakos, Philippe Buchler, and Mauricio Reyes. Multiscale modeling for image analysis of brain tumor studies. *IEEE transactions on biomedical engineering*, 59(1):25–29, 2011.

[29] Marc Niethammer, Yang Huang, and François-Xavier Vialard. Geodesic regression for image time-series. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011: 14th International Conference, Toronto, Canada, September 18-22, 2011, Proceedings, Part II 14*, pages 655–662. Springer, 2011.

[30] Alexandre Bône, Maxime Louis, Alexandre Routier, Jorge Samper, Michael Bacci, Benjamin Charlier, Olivier Colliot, Stanley Durrleman, and Alzheimer's Disease Neuroimaging Initiative. Prediction of the progression of subcortical brain structures in alzheimer's disease from baseline. In *Graphs in Biomedical Image Analysis, Computational Anatomy and Imaging Genetics: First International Workshop, GRAIL 2017, 6th International Workshop, MFCA 2017, and Third International Workshop, MICGen 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 10–14, 2017, Proceedings 1*, pages 101–113. Springer, 2017.

[31] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 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*, pages 234–241. Springer, 2015.

[32] Pamela J LaMontagne, Tammie LS Benzinger, John C Morris, Sarah Keefe, Russ Hornbeck, Chengjie Xiong, Elizabeth Grant, Jason Hassenstab, Krista Moulder, Andrei G Vlassenko, et al. Oasis-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and alzheimer disease. *MedRxiv*, pages 2019–12, 2019.

[33] Clifford R Jack Jr, Matt A Bernstein, Nick C Fox, Paul Thompson, Gene Alexander, Danielle Harvey, Bret Borowski, Paula J Britson, Jennifer L. Whitwell, Chadwick Ward, 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.

[34] M Jorge Cardoso, Wenqi Li, Richard Brown, Nic Ma, Eric Kerfoot, Yiheng Wang, Benjamin Murrey, Andriy Myronenko, Can Zhao, Dong Yang, et al. Monai: An open-source framework for deep learning in healthcare. *arXiv preprint arXiv:2211.02701*, 2022.

[35] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.