Abstract

Normative modelling is an emerging method for understanding the underlying heterogeneity within brain disorders like Alzheimer Disease (AD) by quantifying how each patient deviates from the expected normative pattern that has been learned from a healthy control distribution. Existing deep learning based normative models on magnetic resonance imaging (MRI) neuroimaging data use unimodal autoencoders with a single encoder and decoder that may fail to capture the relationship between brain measurements extracted from different MRI modalities. In this work, we propose multi-modal variational autoencoder (mmVAE) based normative modelling framework that can capture the joint distribution between different modalities and apply it for normative modeling. The deviation maps generated by our proposed multimodal model (mmVAE) are more sensitive to disease staging within AD, have a better correlation with patient cognition and higher number of brain regions with statistically significant deviations compared to a unimodal baseline model with all modalities concatenated as a single input.

1 Introduction

Traditional case-control analyses on neurodegenerative disorders like Alzheimer Disease (AD) assume that there is a single pattern that distinguishes the two contrasted groups and focus on 1st order statistics (group means) to estimate it, effectively ignoring the underlying disease heterogeneity [3, 8]. Normative modeling is a technique for parsing heterogeneity in clinical cohorts, while allowing predictions at an individual subject level [3, 8]. Assuming that neurodegenerative disorders like AD manifest as deviations from a normal pattern of functioning, the parameters of a normative model are learned such that they characterise healthy brains from a control population and provide a statistical measure of normality. Thus, applying the normative model to a disease cohort allows for quantification of the deviation of disease patients from the norm. [9, 5, 18, 6, 2, 16]. Regression models such as hierarchical linear models, support vector regression, and gaussian process regression (GPR) have traditionally been used as normative models (for an extensive list, see [8]). However, it is necessary to train a regression model for each individual brain region which does not incorporate the interactions between brain regions.
Given advances in deep learning technology and the growing availability of large scale datasets, there have been a number of deep learning-based normative models proposed in recent years that uses autoencoders to learn complex structures in the data to best capture patterns in healthy brains [7, 13, 14]. Since AD is a multifactorial disease with more than one biological pathways, using multimodal neuroimaging data can provide complementary information about AD. However, the existing autoencoder approaches have a unimodal structure with a single encoder and decoder for multimodal magnetic resonance imaging (MRI) neuroimaging data, and hence might fail to capture correlations between different MRI modalities (e.g., T1-weighted and T2-weighted) [13, 14].

Contributions: In this work, we aim at implementing a multi-modal variational autoencoder (mm-VAE) based normative modeling framework that can capture the joint distribution between different MRI modalities. Further, we use mmVAE as a normative model to learn the brain regional patterns of healthy subjects and subsequently quantify the deviation of AD patients. We hypothesize that deviations generated by our proposed multimodal framework were more sensitive to disease staging within AD, had a better correlation with patient cognition and estimated more regions with statistically significant deviations compared to a unimodal baseline model. To the best of our knowledge, ours is the first work on developing a multimodal normative modeling framework that can capture the joint distribution between T1 and T2 MRI modalities.

Figure 1: Our proposed multimodal normative modeling framework (mmVAE). Cortical and subcortical brain volumes extracted from T1-weighted MRI scans and hippocampal volumes extracted from T2-weighted MRI scans are used as input to two modality specific encoders. The latent space parameters of the individual modalities can be combined by the Product of Experts (POE) approach as followed in [17] to form the shared latent space, which feeds the modality-specific decoders for reconstructions. The deviations in brain volumes of disease patients are calculated, normalized with respect to the healthy controls to form Z-scores. The regions with statistically significant deviations are identified ($p_{FDR} < 0.05$) after correcting false positives by FDR approach.

2 Proposed Methodology

In our proposed multimodal normative modeling framework (mmVAE) shown in Figure 1, cortical and subcortical brain volumes extracted from T1-weighted Magnetic Resonance Imaging (MRI) scans and hippocampal volumes extracted from T2-weighted MRI scans are used as input to two modality specific encoders. The latent space parameters of the individual modalities can be combined by the Product of Experts (POE) approach as followed in [17]. The main idea is to assume that the joint distribution over the 2 modalities factorizes into a product of single-modality data-generating distributions when conditioned on the latent space and use this to derive the structure and factorization of the variational posterior. The shared latent space formed by the POE approach models the joint distribution between the two modalities (for details, see Appendix A). mmVAE is first trained to reconstruct the brain volumes of cognitively normal participants and subsequently applied on patients with different stages of AD. The main idea of the normative approach is that since mmVAE only learns how to reconstruct the brain region volumes of HC subjects, it will be less precise (more reconstruction error) in reconstructing the brain volumes of AD patients.
Calculating the deviations: For each disease patient, the deviations $d_{ij}$ with respect to the healthy controls are calculated as the absolute signed difference between the original and reconstructed brain region volumes. Assuming that the proposed model is not able to perfectly reconstruct the brain region volumes of the healthy subjects, the deviations are normalized with respect to the mean $\mu_{\text{norm}}$ and variance $\sigma_{\text{norm}}$ of deviations $d_{ij}^{\text{norm}}$ of the healthy participants calculated from a separate held-out validation cohort. The final normalized deviation $\text{deviation}_{ij}$ (Z-scores) are calculated by normalizing the deviations of disease patients with respect to the mean $\mu_j$ and variance $\sigma_j$ of the deviations of healthy participants for each brain region $j$ (see equation below)

$$\text{deviation}_{ij} = \frac{d_{ij} - \mu_{\text{norm}}(d_{ij}^{\text{norm}})}{\sigma_{\text{norm}}(d_{ij}^{\text{norm}})}$$

We identified the brain regions of each patient whose deviations (Z-scores) are significantly different from those of the HC subjects ($p < 0.05$). Since the normalized deviations $\text{deviation}_{ij}$ are estimated independently for each brain region for every patient, FDR (False Discovery Rate) correction was applied to control the Type 1 error rate (false positive correction) [1]. Our proposed framework has been summarized in Figure 1.

3 Experiments and Results

Data: For training the mmVAE framework, we selected 9633 healthy controls (HC) from the UKBiobank dataset [15] after excluding all subjects with recent history of anxiety, depression and nerve disorders. 862 disease patients (106 Significant Memory Concerns (SMC), 312 Early Mild Cognitive Impairment (EMCI), 263 Late Mild Cognitive Impairment (LMCI) and 181 AD) were selected from the Alzheimer’s Disease Neuroimaging (ADNI) dataset [12]. For data harmonization between UKB and ADNI participants, the normative model estimated on UKB was re-trained on ADNI by a transfer learning approach. As patients progress from SMC to the AD stage, the severity of impairment increases. For both the datasets, we used the FreeSurfer software (version 5.1) [4] to estimate the volumes of 64 cortical and 35 subcortical brain regions from T1-weighted MRI images as well as the volumes of 16 hippocampal regions from T2-weighted MRI images, respectively. For information about baselines and other implementation details, see Appendix B.

Results

Sensitivity towards disease staging: For both our proposed multimodal and baseline unimodal framework, the patients exhibited higher magnitude of mean deviations (more brain abnormality) across all brain regions with increasing severity of their condition from SMC to AD (Figure 2A). Higher slope across disease categories suggests that our proposed model is more sensitive to the different stages of AD compared to the baseline. The pairwise differences in deviations between the disease categories were statistically significant ($p_{\text{FDR}} < 0.05$) except for SMC and EMCI pair (Figure 2B). Correlation with patient cognition: We analyzed the Pearson correlation between the patient-level deviation maps and cognitive assessment scores, AD Assessment Scale (ADAS13) [11] and Rey Auditory Verbal Learning Test (RAVLT) [10]. ADAS13 and RAVLT are 2 tests to assess the level of cognitive dysfunction due to AD. High scores of ADAS13 and low scores of RAVLT indicate greater loss in memory and cognition. Our proposed framework exhibited higher correlation (high $r$ value) with patient cognition, compared to the unimodal baseline (Figure 2D). Statistically significant brain regions: We also identified the brain regions whose deviations (neuroanatomical alterations in the brain) were statistically significant compared to the healthy controls. Our proposed model had more brain regions for each patient (out of total 115 regions) with statistically significant deviations (Figure 2C) and a higher fraction of significance for each hippocampal region (Figure 2E) respectively compared to the unimodal baseline.

4 Conclusion and Future Work

We propose a multimodal VAE-based normative modelling framework (mmVAE), which models the joint distribution between brain region volumes derived from multiple MRI modalities (T1 and T2 MRI). Our framework quantifies at a subject-level how patients with different stages of AD deviate from the expected pattern learned from the healthy controls. The deviations generated by
our proposed multimodal framework were more sensitive to disease staging within AD, had a better
correlation with patient cognition and more regions with statistically significant deviations compared
to a unimodal baseline model. As part of future work, we plan to perform further validations of
our proposed model to estimate its generalizability on more neuroimaging datasets and utilize the
patient-level deviation maps to further investigate heterogeneity of other neurodegenerative and
neuropsychiatric disorders.

5 Negative Societal Impact

Our work mmVAE takes a step towards addressing heterogeneity in a neurodegenerative disorder like
Alzheimer’s Disease. However, one of the potential limitations of our work is the fact that the ADNI
dataset used only consists of subjects from the United States. With the goal of deploying mmVAE
as part of a computer aided diagnostic system, our model may not generalize successfully when
applied to different populations from other countries. The lack of generalizability to different types
of neuroimaging datasets might lead to false positives, which result in misdiagnosis and inappropriate
treatment.

References

[1] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and
powerful approach to multiple testing. *Journal of the Royal statistical society: series B
(Methodological)*, 57(1):289–300, 1995.
[2] Richard Al Bethlehem, Jakob Seidlitz, Rafael Romero-Garcia, Stavros Trakoshis, Guillaume Dumas, and Michael V Lombardo. A normative modelling approach reveals age-atypical cortical thickness in a subgroup of males with autism spectrum disorder. Communications biology, 3(1):1–10, 2020.

[3] Aoyan Dong, Jon B Toledo, Nicolas Honnorat, Jimit Doshi, Erdem Varol, Aristeidis Sotiras, David Wolk, John Q Trojanowski, Christos Davatzikos, and Alzheimer’s Disease Neuroimaging Initiative. Heterogeneity of neuroanatomical patterns in prodromal alzheimer’s disease: links to cognition, progression and biomarkers. Brain, 140(3):735–747, 2017.

[4] Bruce Fischl, David H Salat, Evelina Busa, Marilyn Albert, Megan Dieterich, Christian Haselgrove, Andre Van Der Kouwe, Ron Killiany, David Kennedy, Shuna Klaveness, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron, 33(3):341–355, 2002.

[5] Wyke Huizinga, Dirk HJ Poot, Meike W Vernooij, Gennady V Roshchupkin, Esther E Bron, Mohammad Arfan Ikram, Daniel Rueckert, Wiro J Niessen, Stefan Klein, Alzheimer’s Disease Neuroimaging Initiative, et al. A spatio-temporal reference model of the aging brain. NeuroImage, 169:11–22, 2018.

[6] Seyed Mostafa Kia and Andre F Marquand. Neural processes mixed-effect models for deep normative modeling of clinical neuroimaging data. In International Conference on Medical Imaging with Deep Learning, pages 297–314. PMLR, 2019.

[7] Sayantan Kumar and Aristeidis Sotiras. Normvae: Normative modeling on neuroimaging data using variational autoencoders. arXiv preprint arXiv:2110.04903, 2021.

[8] Andre F Marquand, Seyed Mostafa Kia, Mariam Zabihi, Thomas Wolfers, Jan K Buitelaar, and Christian F Beckmann. Conceptualizing mental disorders as deviations from normative functioning. Molecular psychiatry, 24(10):1415–1424, 2019.

[9] Andre F Marquand, Iead Rezek, Jan Buitelaar, and Christian F Beckmann. Understanding heterogeneity in clinical cohorts using normative models: beyond case-control studies. Biological psychiatry, 80(7):552–561, 2016.

[10] Sona Matloubi, Ali Mohammadzadeh, Zahra Jafari, and Alireza Akbarzade Baghban. Effect of background music on auditory-verbal memory performance. Audiology, 23(5), 2014.

[11] Richard C Mohs, David Knopman, Ronald C Petersen, Steven H Ferris, Chris Ernesto, Michael Grundman, Mary Sano, Linas Bielaukas, David Geldmacher, Chris Clark, et al. Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the alzheimer’s disease assessment scale that broaden its scope. Alzheimer disease and associated disorders, 1997.

[12] Susanne G Mueller, Michael W Weiner, Leon J Thal, Ronald C Petersen, Clifford R Jack, William Jagust, John Q Trojanowski, Arthur W Toga, and Laurel Beckett. Ways toward an early diagnosis in alzheimer’s disease: the alzheimer’s disease neuroimaging initiative (adni). Alzheimer’s & Dementia, 1(1):55–66, 2005.

[13] Walter HL Pinaya, Andrea Mechelli, and João R Sato. Using deep autoencoders to identify abnormal brain structural patterns in neuropsychiatric disorders: A large-scale multi-sample study. Human brain mapping, 40(3):944–954, 2019.

[14] Walter HL Pinaya, Cristina Scarpazza, Rafael Garcia-Dias, Sandra Vieira, Lea Baecker, Pedro F da Costa, Alberto Redolfi, Giovanni B Frisoni, Michela Pievani, Vince D Calhoun, et al. Using normative modelling to detect disease progression in mild cognitive impairment and alzheimer’s disease in a cross-sectional multi-cohort study. Scientific reports, 11(1):1–13, 2021.

[15] Cathie Sudlow, John Gallacher, Naomi Allen, Valerie Beral, Paul Burton, John Danesh, Paul Downey, Paul Elliott, Jane Green, Martin 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 medicine, 12(3):e1001779, 2015.
[16] Thomas Wolfers, Nhat Trung Doan, Tobias Kaufmann, Dag Alnæs, Torgeir Moberget, Ingrid Agartz, Jan K Buitelaar, Torill Ueland, Ingrid Melle, Barbara Franke, et al. Mapping the heterogeneous phenotype of schizophrenia and bipolar disorder using normative models. *JAMA psychiatry*, 75(11):1146–1155, 2018.

[17] Mike Wu and Noah Goodman. Multimodal generative models for scalable weakly-supervised learning. *Advances in Neural Information Processing Systems*, 31, 2018.

[18] Gabriel Ziegler, Gerard R Ridgway, Robert Dahnke, Christian Gaser, Alzheimer’s Disease Neuroimaging Initiative, et al. Individualized gaussian process-based prediction and detection of local and global gray matter abnormalities in elderly subjects. *NeuroImage*, 97:333–348, 2014.
A Joint distribution between multiple modalities

Our proposed mmVAE has separate modality-specific encoders and decoders for individual modalities. The main idea is to assume the joint distribution over the multiple modalities factorizes into a product of single-modality data-generating distributions when conditioned on the latent space. This assumption is used to derive the structure and factorization of the variational posterior. Without loss of generality, we assume that the N modalities $x_1, \ldots x_N$ are conditionally independent given the common latent variable $z$. So, we assume a generative model of the form $p_\theta(x_1, \ldots x_N, z) = p(z) \prod_{i=1}^{N} p_\theta(x_i | z)$. The conditional independence assumptions in the generative model imply a relation among joint and single-modality posteriors as shown below.

$$p(z | x_1, \ldots, x_N) = \frac{p(x_1, \ldots, x_N | z) * p(z)}{p(x_1, \ldots, x_N)} = \frac{p(z)}{p(x_1, \ldots, x_N)} * \prod_{i=1}^{N} p(x_i | z)$$

$$= \frac{p(z) * \prod_{i=1}^{N} p(z | x_i) p(x_i)}{p(x_1, \ldots, x_N)}$$

$$= \frac{\prod_{i=1}^{N} p(z | x_i) \prod_{i=1}^{N} p(x_i)}{\prod_{i=1}^{N-1} p(x_i)}$$

$$\propto \prod_{i=1}^{N} p(z | x_i)$$

Here we see that the joint posterior is a product of individual posteriors, with an additional quotient by the prior. Alternatively, if we approximate $p(z | x_i) = q(z | x_i) \equiv \tilde{q}(z | x_i)p(z)$ where $\tilde{q}(z | x_i)$ the underlying inference network, we can avoid the quotient term $p(z)$. Now, we can approximate the joint posterior as shown below. In other words, we can use a product of experts (PoE), including a “prior expert” $p(z)$, as the approximating distribution for the joint-posterior

$$p(z | x_1, \ldots, x_N) \propto \prod_{i=1}^{N} p(z | x_i)$$

Data from the individual modalities are fed into the corresponding encoders to form their respective latent space parameters (mean and variance). The product distribution required above are not in general solvable in closed form. However, if we approximate both $p(z)$ and $\tilde{q}(z | x_i)$ as Gaussian, then we utilize the solution shown in [?] that a product of Gaussian experts is itself Gaussian with mean $\mu = \left( \sum_i \mu_i * T_i \right) / \left( \sum_i T_i \right)^{-1}$ and variance $\sigma = \left( \sum_i T_i \right)^{-1}$ where $\mu_i$ and $\sigma_i$ are parameters of the $i$-th Gaussian expert and $T_i = \sigma_i^{-1}$.

B Feature Engineering

As part of the feature pre-processing step, the brain region volumes of each subject were normalized by their Intracranial Volume (ICV). The healthy controls from UK Biobank data was split into 80% for training and 20% for validation, which was used for early stopping to prevent overfitting. The 269 CN participants in ADNI were split into 70% for model training and validation, a 15% held-out validation set for estimating the parameters of the normative population and 15% in the test for estimating the deviations (see next subsection). The volumes of each region were scaled between 0 and 1 using MinMax scaling. The mean and variance calculated in the training set were also used to scale the data in validation and test sets.

C Baselines and Training details

We compared our proposed framework with a baseline VAE having a unimodal VAE (single encoder and decoder) which takes the cortical, subcortical and hippocampal region volumes into a single...
concatenated input. We conditioned both the proposed and baseline VAE networks on the age and sex of patients, represented as one-hot encoding vectors, to ensure that the deviations in regional brain volumes reflect only the disease pathology and not deviations due to effects of covariates. Both the VAE models were trained using the Adam optimizer with model hyperparameters as follows: epochs = 500, learning rate = $10^{-5}$, batch size = 256 and latent dimension = 64. The encoder and decoder networks have 4 dense layers of sizes 512, 256, 128, 64 and 64, 128, 256, 512 respectively.

D Statistically significant brain regions - cortical and subcortical

Figure 3: Fraction of significance: Number of times (%) each cortical (top) and subcortical (bottom) region exhibited statistically significant deviations.