INVESTIGATING CONVERSION FROM MILD COGNITIVE IMPAIRMENT TO ALZHEIMER'S DISEASE USING LATENT SPACE MANIPULATION

Deniz Sezin Ayvaz and İnci M. Baytaş

Department of Computer Engineering, Boğaziçi University
Bebek, Istanbul, Turkey 34342
sezin.ayvaz@boun.edu.tr, inci.baytas@boun.edu.tr

ABSTRACT

Alzheimer's disease, a progressive neurologic disorder, is the most common cause of dementia, affecting millions worldwide. Mild Cognitive Impairment (MCI) is considered an intermediate stage before Alzheimer's. Early prediction of the conversion from MCI to Alzheimer's is crucial to take necessary precautions for decelerating the disease progression and developing suitable treatments. This study proposes a deep learning framework to identify patients whose diagnoses might change from MCI to Alzheimer's in the following stages. In particular, latent space manipulation techniques are applied to the latent space of a variational auto-encoder trained with MCI and Alzheimer's patients. The manipulation step aims to reveal significant attributes triggering the conversion. Secondly, a correlation between the manipulation's magnitude and the conversion time is investigated to introduce a predictive perspective. Experimental results show promising quantitative and qualitative results on one of the literature's most extensive and commonly used Alzheimer's disease neuroimaging datasets.

1. INTRODUCTION

Neurodegenerative diseases are a common concern threatening millions of lives worldwide. Alzheimer's disease (AD) is the most well-known cause of dementia that affect cognitive functions [1]. The transition between normal cognitive functions and dementia is called Mild Cognitive Impairment (MCI) [2]. While 1-2% of the cognitively normal elderly develop dementia each year, approximately 12% of those with MCI convert to dementia [3]. Some patients with MCI symptoms have AD in the later stages. However, some may remain in the MCI stage or progress very slowly [4]. Therefore, predicting the conversion from MCI to AD is of great importance.

AD can be diagnosed with an accuracy close to 95% [5]. However, for the early diagnosis, it is imperative to take the clinical history of a patient and their families, perform a neurological examination and neuropsychological tests suitable to the education level, and evaluate the brain tissue with neuroimaging methods, such as Magnetic Resonance Imaging (MRI), and Positron Emission Tomography (PET). Today, machine learning and deep learning techniques can assist clinicians in simultaneously analyzing multiple modalities and developing predictive models for early AD diagnosis.

Although predictive models successfully solve various healthcare problems, they are not always easily interpretable. However, it is crucial to know the underlying reasons behind the behavior of a model. This study proposes a deep learning framework to discover the underlying factors leading to the conversion from MCI to AD. A latent space manipulation technique is adopted to obtain the significant attributes and decipher their behavior. The contributions of the proposed study are outlined below.

- A variational auto-encoder (VAE) network is trained to learn the latent representations of the MCI and AD patient data without any labels.
- Supervised and unsupervised latent space manipulation techniques are applied to the latent representations of the MCI patients. The VAE decoder outputs new patient data from the manipulated latent representations. To the best of our knowledge, the proposed study is the first attempt to apply latent space manipulation to tabular patient data.
- The proposed approach facilitates discovering relationships between the directions in the latent space and the AD diagnosis.
- We show that the generated AD patients via latent space manipulation and the actual MCI patients diagnosed with AD in the future stages share similar characteristics. This finding may imply that our approach may correspond to a clinically realistic transition between MCI and AD. Thus, the proposed framework may reveal more information about the conversion than a binary classifier.

We conduct experiments with the dataset publicly shared by The Alzheimer's Disease Prediction of Longitudinal Evolution, TADPOLE challenge [6]. The credibility of the experimental results from a clinical perspective is confirmed by a neurologist.

2. BACKGROUND

Clinical studies show that not all MCI patients convert to AD [7]. Therefore, detecting MCI patients likely to develop AD is crucial to improving the quality of life. This section reviews the studies for detecting MCI to AD conversion and the latent space manipulation providing a basis for our study.

2.1 Studies on the Conversion Between MCI and AD

Biological and behavioral markers from the lab and neuropsychological tests are often used to detect AD [7]. There has been a quest for distinctive noninvasive identifiers and fusion

of multiple identifiers to diagnose a patient with dementia. For instance, Lu *et al.* proposed to fuse AD biomarkers with an event-based probabilistic framework [8]. The authors state that screening abnormal amyloid and tau proteins has been essential for accurate diagnosis, but it is not always possible [8]. Therefore, it is necessary to enable accurate and early AD diagnosis considering the biomarkers obtained by noninvasive screening. It is also essential to analyze multiple factors simultaneously. For instance, Sun *et al.* considered a hierarchical Bayesian model to simultaneously extract underlying factors from atrophy patterns and cognitive impairments [9].

Mofrad *et al.* also focused on brain atrophy due to aging [10]. The authors applied an ensemble of predictive models on MRI images to predict the conversion from cognitively normal patients to MCI and from MCI to AD [10]. For the same task, Bae *et al.* used a 3-dimensional Convolutional Neural Network (CNN) to extract features from structural MRI images [11]. Sufficient data is required when complex models are utilized. However, it is often not possible to collect large-scale patient data. Therefore, the authors resorted to fine-tuning a pre-trained CNN [11]. Instead of directly utilizing the MRI images, some studies aim to discover new biomarkers from MRI. Kung *et al.* proposed a biomarker, the ratio of principal curvatures, that may be an identifier to detect the transition between MCI and AD [12]. Kuang *et al.*, on the other hand, considered a drastically different dataset collected via a cognitive questionnaire [13]. The authors employed logistic regression and Multi-layer Perceptron (MLP). Although less complex and more interpretable models are used, relying only on the questionnaire dataset is a limitation [13].

In addition to CNN and MLP, some studies consider recurrent neural network architectures, such as Long Short-Term Memory (LSTM), to learn the temporal patterns in the disease progression. Li *et al.* proposed an LSTM-based framework to predict the risk of progression from MCI to AD using cognitive measures and features extracted from the MRI of the hippocampus [14]. Wegmayr *et al.* took the temporal analysis of the progression to another level by training a recursive generator architecture to synthesize the following stages of a patient's brain image [15].

2.2 Latent Space Manipulation

Latent space manipulation is usually applied to the latent spaces of Generative Adversarial Networks (GANs). Controlling the latent space enables reshaping the generated output in specific ways. For example, one of the studies in computer vision employs latent space manipulation to generate face images that look more feminine, older, or with more makeup [16]. Another study from the healthcare domain utilizes latent space manipulation for high-resolution medical image synthesis using GANs [17].

The latent space manipulation can be learned in supervised and unsupervised manners. In the supervised techniques, an attribute classifier is trained to discover relevant directions within the latent space [16, 18]. Principal Component Analysis (PCA) can also be applied to data points sampled from a latent space to find relevant directions [19].

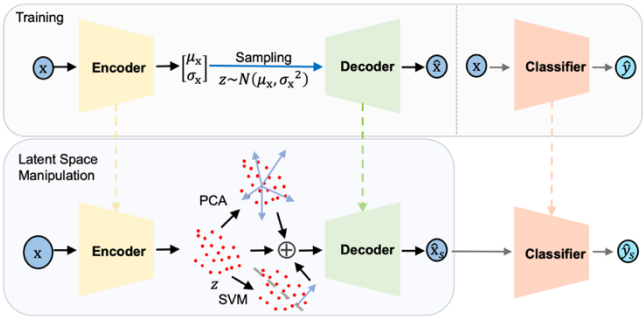

Figure 1. Proposed framework.

On the other hand, DeVries *et al.* proposed using simple transformations, such as extrapolation, interpolation, and Gaussian noise, to the latent representations learned via an auto-encoder [20].

The studies presented in Section 2.1 mainly focus on training predictive models by posing the conversion from MCI to AD as a classification problem. Despite the high performance of deep learning models, it is challenging to unravel biomarkers and modalities contributing to predictive performance. This study employs supervised and unsupervised latent space manipulation techniques to investigate the underlying factors of the conversion between MCI and AD. We aim to reveal more information about MCI patients and their conversion to AD.

3. METHODS

In this section, we present the proposed framework shown in Figure 1. Details of the proposed framework are presented in the following sections.

3.1 Problem Definition

Predicting the conversion from MCI to AD may be posed as a supervised learning problem. However, the primary goal of this study is to investigate the reasons behind this progression. Therefore, we pose this problem as an exploration of the meaningful directions in the latent space where the patients are represented. In this study, we try to find an answer to the following question: How can the latent directions help us understand the conversion from MCI to AD?

3.2 Model Architecture

The proposed framework in Figure 1 comprises a binary classifier and a VAE trained on MCI and AD patients. We did not consider a control group in this study since we focused on the conversion between MCI and AD rather than on the expected MCI.

3.2.1 Binary Classifier

A binary classifier is trained to distinguish patient visits with MCI and AD diagnoses. We do not consider temporal modeling at this stage. The number of time steps in publicly available datasets is often insufficient to capture a temporal pattern. Therefore, an MLP network is trained with the patient visits resulting in MCI and AD diagnoses. The binary cross-entropy

loss, $\mathcal{L}_{CE} = -\frac{1}{N}\sum_{i=1}^{N} y_i \log \hat{y}_i + (1 - y_i)\log(1 - \hat{y}_i)$, with ground-truth y_i and predicted $\hat{y}_i$ is used to train the classifier.

The binary classifier is used to evaluate the output of the manipulation procedure. Otherwise, any feedback from the binary classifier is not used during the training of the VAE. The probabilistic outcome of the classifier also provides insight into the manipulated patient and the amount of manipulation towards the direction of AD. For instance, if an MCI patient is classified as AD with high confidence after the manipulation, this may mean that the MCI patient is pushed beyond the decision boundary and shows the characteristics of an AD diagnosis.

3.2.2 Variational Autoencoder

The VAE aims to learn a latent space for the input samples in an unsupervised way. Unlike the traditional autoencoders, the encoder outputs a distribution rather than projecting the input to a fixed-sized latent vector. Thus, the VAE prevents overfitting specific patients in the training set but captures the distributions of MCI and AD patients. The decoder gains a generative property by reconstructing the input from a sample drawn from a normal distribution. Therefore, VAE's latent space is more versatile for latent space manipulation than the traditional autoencoder. This study uses fully-connected encoder and decoder architectures with ReLU activation functions and is trained to minimize the following loss function.

$$\mathcal{L}_{\text{VAE}} = \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|_2^2 + \gamma \text{KL}[N(\boldsymbol{\mu}_x, \boldsymbol{\sigma}_x^2), N(0,1)] \qquad (1)$$

where $\mathbf{x}_i \in \mathbb{R}^d$ is the input, $\hat{\mathbf{x}}_i \in \mathbb{R}^d$ is the reconstructed input, $\boldsymbol{\mu}_x, \boldsymbol{\sigma}_x^2$ denote mean and variance of the distribution learned by the encoder, KL is the Kullback-Leibler divergence. Gamma controls the effect of KL divergence that measures the distance between the two distributions.

3.3 Latent Space Manipulation

Inspired by the unsupervised latent space manipulation technique [19], we investigate the latent direction in the latent space of the VAE by applying PCA to the latent representations of the patients. Thus, the principal components can capture the directions containing the most variation in the latent space. Since only MCI and AD patients are considered in this study, we expect that the first principal component should be associated with the direction of the disease progression.

We also conducted experiments with a supervised manipulation technique [16] to be able to control the manipulation direction directly. An SVM classifier with a linear kernel is trained on the latent representations of MCI and AD patients. The opposite direction of the parameter vector denoting the direction of the conversion to AD is used to manipulate MCI patients. Our analysis shows that both supervised and unsupervised manipulation techniques lead in similar directions for the dataset used in this study. For this reason, we continue with unsupervised latent manipulation in our experiments.

The latent representation of a patient is manipulated by adding either a principal component vector or the unit direction vector of the SVM classifier, multiplied by a coefficient, as shown below.

$$\mathbf{z}^* = \mathbf{z} + \alpha\mathbf{u} \qquad (2)$$

where $\mathbf{z}$ is the latent representation learned by the VAE, $\mathbf{u}$ is a principal component or the direction obtained by the SVM classifier, and α denotes a scalar controlling the manipulation amount. Once the manipulated latent representation of the patient, $\mathbf{z}^*$, is obtained, the decoder of the VAE is used to synthesize a new patient from the manipulated latent representation; bold equals $= \text{decoder}(\mathbf{z}^*)$. Next, the diagnosis of the manipulated patient $\hat{\mathbf{x}}_s$ is predicted using the binary classifier. The following steps summarize the procedure.

1) The reconstructed MCI patient vectors by the decoder of the VAE are fed to the binary classifier.
2) The MCI patients, whose reconstructed samples are correctly identified as MCI by the binary classifier, are selected for the latent space manipulation step. This step aims to choose the correctly reconstructed data points so that the effect of manipulation on label shift can be observed.
3) PCA is applied to the latent space, and principal components (PCs) are obtained.
4) The selected MCI patients are manipulated, and the binary classifier obtains the predicted labels of the synthesized manipulated patients.

The key findings of the proposed approach are presented in the next section.

4. RESULTS

Experiments are conducted with the TADPOLE dataset [6], which is retrieved from Alzheimer's Disease Neuroimaging Initiative (ADNI) [21]. Since similar results from the supervised and unsupervised latent space manipulation techniques are obtained, only unsupervised latent space manipulation results are reported in this section.

4.1 Data Preprocessing

TADPOLE dataset comprises multi-modal data such as MRI, PET, CSF, diffusion tensor imaging (DTI), cognitive tests, and some genetic and demographic information from the ADNI database [22]. The visits of 1737 subjects diagnosed with MCI, AD, and control patients are recorded in the dataset. On average, each subject has 6.68 different visits up to 6 years

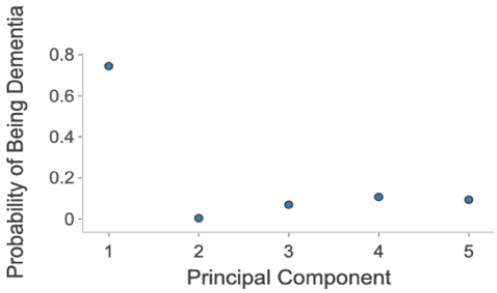

Figure 2. Manipulation effects of principal components for fixed α.

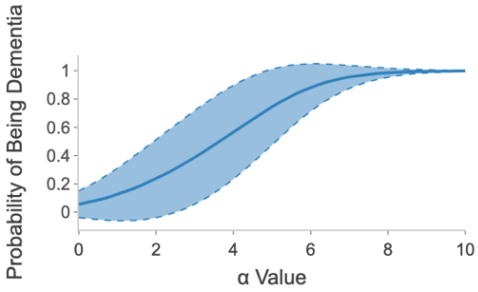

Figure 3. Manipulation effect of the first PC for α. y-axis is the average dementia probability of the patients after manipulation and the shaded region denotes the standard deviation.

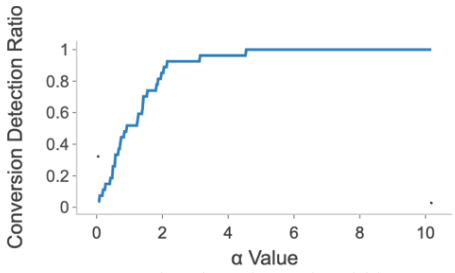

Figure 4. Conversion detection ratio within 2-year follow-up with different $\boldsymbol{\alpha}$ values

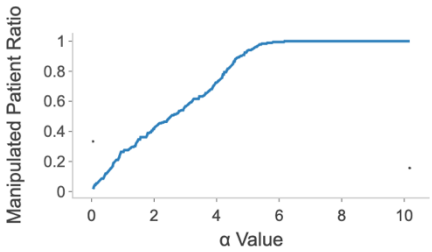

Figure 5. The ratio of number of manipulated patients with different $\boldsymbol{\alpha}$ values

from their first examination. The TADPOLE challenge recommends 23 variables, including diagnosis, neuropsychological test scores, and anatomical features derived from T1 MRI, PET, and CSF markers [6]. However, we selected a subset of size 16 with a missing value rate of less than 50%. Missing data is imputed with the forward filling by replacing the most recent measurement of the patient before the missing. If the forward filling cannot be used, an average value calculated from the visits with the same label is used.

Measurements that are not clinically plausible are removed from the dataset. PET and MRI measurements, namely hippocampus, whole brain, middle temporal cortical, entorhinal, and fusiform volumes, are normalized by the intracranial volume of the particular patient, as suggested by Voevodskaya *et al.* [23]. In our experiments, we use the patient visits with only MCI and AD labels. Thus, we have 1343 data points with AD and 4021 with MCI labels. During the evaluation, we focus only on correctly classified test MCI data points, which contain 185 patients, and 30 of them convert to dementia in the future. We use the standard scaling method for normalizing the numerical columns.

4.2 Training Scheme

Hyperparameter optimization is employed with the 5-fold cross-validation. The encoder and decoder are designed as three-layer fully connected networks, and the binary classifier is a two-layer fully connected network. The hyperparameter γ in Eq. 1 is set to 0.1. VAE and the binary classifier are trained with the training set. PCA is applied to latent space representations of the training dataset. Manipulation experiments are conducted on latent space representations of the test dataset after the manipulation direction vectors are obtained. Google Colab[1] and Tensorflow[2] are used to implement the proposed approach.

4.3 Choice of Principal Components

We experiment with the top 5 PCs. Each PC is added to the latent representations of the MCI patients as in Eq. 2. The manipulation effect of PCs is evaluated by the number of MCI patients classified as dementia after the manipulation. In Figure 2, for a fixed value of $\alpha = 5$, we plot the average probability of the classifier for dementia computed using manipulated MCI patients against the five different principal directions. This empirical evidence indicates that the first PC moves the latent representation of an MCI patient towards dementia.

4.4 Investigating the Effect of α

Next, we investigate the effect of varying alpha values. Figure 3 demonstrates the average softmax probabilities of being dementia against different α values for the first principal component. Different α values have a non-linear behavior for converting the MCI patients to dementia. We observe similar behavior when the number of manipulated MCI patients labeled as dementia is investigated for each α value.

5. DISCUSSION

5.1 Clinical Interpretation

[1] https://colab.research.google.com/

[2] https://www.tensorflow.org/

To investigate the individual changes in variables after manipulation, we manipulated the MCI patients diagnosed with dementia in future time steps, using the first PC with a fixed alpha value. We compare the average changes in a subset of variables between the manipulated and original time steps labeled as dementia in Table 1.

Table 1. Average changes as % between the manipulated and original time steps labeled as dementia for a fixed α. CDRSB: Clinical Dementia Rating Scale Sum of Boxes, ADAS: Alzheimer's Disease Assessment Scale, RAVLT: Rey Auditory Verbal Learning Test

Variable	Manipulated vs. Original	Future Visit vs. Original
CDRSB	27.85	46.29
RAVLT immediate	-8.01	-7.31
ADAS-Cog11	15.43	15.49
FAQ	28.38	47.88
MMSE	-4.4	-6.49
Ventricles	5.79	6.87
Hippocampus	-6.67	-4.93
Whole brain vol.	-2.25	-2.39
Entorhinal cortical vol.	-6.84	-4.98
Fusiform cortical vol.	-3.77	-2.68
Middle temporal vol.	-3.76	-3.04

Future visit information after the manipulated data points of the corresponding patients shows that the patients undergo similar changes to be diagnosed with dementia. For instance, clinically plausible results include an increase in the CDRSB score and a decrease in the MMSE score as the disease progresses. The reported results in this section imply that we can discover a direction that contains a significant amount of information about AD in the latent space of a deep network.

Although the manipulation effect falls behind on changing CDRSB and FAQ variables accordingly, the changes in other variables enable the visit data to be labeled as dementia by the classifier. This implies that the variation in MRI values and the other cognitive tests can better explain the shift in diagnosis.

In Figure 4, we investigate the ratio of detected test patients who convert to AD within a 2-year follow-up by the manipulation method. This figure demonstrates that the detected converted patients are more realistic as the manipulation amount increases. The first visit of each MCI patient among the manipulated visit data for each $\boldsymbol{\alpha}$ is selected. The time interval between the first visit of an MCI patient at which the conversion is detected by our method and the patient's actual conversion time to AD is extracted from the dataset. If the time interval is less than or equal to 24 months, the patient is included in Figure 4.

In Figure 5, we demonstrate the ratio of patients labeled as AD by the classifier after manipulating the test set. Visits

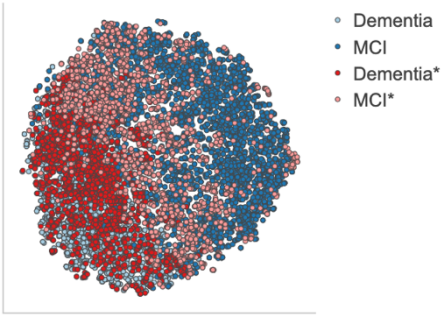

Figure 6. t-SNE of the manipulated data points with their original versions in the classifier's latent space with for a fixed $\boldsymbol{\alpha = 4}$.

belonging to different time points of the same MCI patient can be manipulated simultaneously by a fixed α value. Therefore, we take a unique number of patients for this figure. One might consider that if the α value is too high, the number of manipulated patients covers almost all patients in the test set, which is not desirable to detect the patients who will convert to AD in real life. There needs to be a balance point between these two figures, where the method enables detecting a maximum number of progressive MCI and AD-to-be patients while focusing on the minimum portion of the cohort. For example, when $\alpha = 2$, the ratio of detected progressive MCI patients is ~86%, while the percentage of manipulated patients is ~42%. These two figures imply that using a reasonable α value would enable clinicians to detect most of the progressive MCI cases that will convert to dementia by focusing only on a small portion of the cohort. This result shows the potential to associate α with the risk of conversion. In future work, we plan to formalize this relationship for predictive evaluation.

5.2 Visualizing the Classifier's Latent Space

Figure 6 displays manipulated data points with their original counterparts in the latent space of the binary classifier. Since we apply manipulation only on the MCI labeled patients, Dementia* and MCI* represent manipulated versions of MCI data points. With this visualization, we can observe the diversity in generated data points. We infer that the proposed approach may be suitable for generating realistic synthetic dementia patients under the low data regime.

ACKNOWLEDGMENT

This work is supported by Boğaziçi University Research Fund under grant number 17004 and by the Scientific and Technological Research Council of Turkey ARDEB 3501 under grant number 121E594. The authors thank neurologist Dr. Nazlı Gamze Bülbül for her assistance in interpreting the results from a clinical perspective.

REFERENCES

[1] Y. Hou, X. Dan, M. Babbar, Y. Wei, S. G. Hasselbalch, D. L. Croteau and V. A. Bohr, "Ageing as a risk factor

for neurodegenerative disease," *Nature Reviews Neurology,* vol. 15, no. 10, pp. 565-581, 2019.

[2] M. J. Aminoff and R. B. Daroff, "Mild Cognitive Impairment," in *Encyclopedia of the Neurological Sciences (Second Edition)*, Academic Press, 2014, pp. 72-75.

[3] B. L. Plassman, K. M. Langa, G. G. Fisher, S. G. Heeringa, D. R. Weir, M. B. Ofstedal, J. R. Burke, M. D. Hurd, G. G. Potter, W. L. Rodgers and others, "Prevalence of cognitive impairment without dementia in the United States," *Annals of internal medicine,* vol. 148, no. 6, pp. 427-434, 2008.

[4] S. Minhas, A. Khanum, F. Riaz, S. A. Khan and A. Alvi, "Predicting progression from mild cognitive impairment to Alzheimer's disease using autoregressive modelling of longitudinal and multimodal biomarkers," *IEEE Journal of Biomedical and Health Informatics,* vol. 22, no. 3, pp. 818-825, 2017.

[5] L. Mucke, "Alzheimer's disease," *Nature,* vol. 461, no. 7266, pp. 895-897, 2009.

[6] R. V. Marinescu and others, "The Alzheimer's Disease Prediction Of Longitudinal Evolution (TADPOLE) Challenge: Results after 1 Year Follow-up," *arXiv preprint,* vol. arXiv:2002.03419, 2020.

[7] R. M. Chapman, M. Mapstone, J. W. McCrary, M. N. Gardner, A. Porsteinsson, T. C. Sandoval, M. D. Guilliy, E. DeGrush and L. A. Reilly, "Predicting conversion from mild cognitive impairment to Alzheimer's disease using neuropsychological tests and multivariate methods," *Journal of Clinical and Experimental Neuropsychology,* vol. 33, no. 2, pp. 187-199, 2011.

[8] X. Lu, J. Chen, H. Shu, Z. Wang, Y.-m. Shi, Y.-g. Yuan, C.-m. Xie, W.-x. Liao, F. Su, Y.-c. Shi and others, "Predicting conversion to Alzheimer's disease among individual high-risk patients using the Characterizing AD Risk Events index model," *CNS neuroscience & therapeutics,* vol. 26, no. 7, pp. 720-729, 2020.

[9] N. Sun, E. C. Mormino, J. Chen, M. R. Sabuncu and B. T. T. Yeo, "Multi-modal latent factor exploration of atrophy, cognitive and tau heterogeneity in Alzheimer's disease," *NeuroImage,* vol. 201, no. 116043, 2019.

[10] S. A. Mofrad, A. Lundervold and A. S. Lundervold, "A predictive framework based on brain volume trajectories enabling early detection of Alzheimer's disease," *Computerized Medical Imaging and Graphics,* vol. 90, no. 101910, 2021.

[11] J. Bae, J. Stocks, A. Heywood, Y. Jung, L. Jenkins, V. Hill, A. Katsaggelos, K. Popuri, H. Rosen, M. F. Beg and others, "Transfer learning for predicting conversion from mild cognitive impairment to dementia of Alzheimer's type based on a three-dimensional convolutional neural network," *Neurobiology of Aging,* vol. 99, pp. 53-64, 2021.

[12] T.-H. Kung, T.-C. Chao, Y.-R. Xie, M.-C. Pai, Y.-M. Kuo, G. Lee and C. Giun, "Neuroimage biomarker identification of the conversion of mild cognitive impairment to Alzheimer's Disease," *Frontiers in Neuroscience,* vol. 15, 2021.

[13] J. Kuang, P. Zhang, T. Cai, Z. Zou, L. Li, N. Wang and L. Wu, "Prediction of transition from mild cognitive impairement to Alzheimer's disease based on a logistic regression-artificial neural network-decision tree model," *Geriatrics & Gerontology International,* vol. 21, no. 1, pp. 43-47, 2021.

[14] H. Li and Y. Fan "Early Prediction of Alzheimer's Disease Dementia Based on Baseline Hippocampal MRI and 1-Yeear Follow-Up Cognitive Measures usign Deep Receurrent Neural Networks," in *IEEE 16th International Symposium on Biomedical Imaging,* 2019.

[15] V. Wegmayr, M. Hörold and J. M. Buhmann, "Generative Aging of Brain MR-Images and Prediction of Alzheimer Progression," in *German Conference on Pattern Recognition*, 2019.

[16] Y. Shen, G. Jinjin, T. Xiaoou and Z. Bolei, "Interpreting the Latent Space of GANs for Semantic Face Editing," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020.

[17] L. Fetty, M. Bylund, P. Kuess, G. Heilemann, T. Nyholm, D. Georg and T. Löfstedt, "Latent space manipulation for high-resolution medical image synthesis via the StyleGAN," *Zeitschrift für Medizinische Physik,* vol. 30, no. 4, pp. 305-314, 2020.

[18] L. Goetschalckx, A. Andonian, A. Oliva and P. Isola, "GANalyze: Toward Visual Definitions of Cognitive Image Properties," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019.

[19] E. Härkönen, A. Hertzmann, J. Lehtinen and S. Paris, "Ganspace: Discovering interpretable gan controls.," in *Advances in Neural Information Processing Systems* , 2020.

[20] T. DeVries and G. W. Taylor, "Dataset Augmentation in Feature Space," *preprint arXiv:1702.05538,* 2017.

[21] R. C. Petersen and others, "Alzheimer's disease neuroimaging initiative (ADNI): clinical characterization," *Neurology,* vol. 74, no. 3, pp. 201-209, 2010.

[22] J. R. Clifford and others, "Magnetic resonance imaging in Alzheimer's Disease Neuroimaging Initiative 2," *Alzheimer's & Dementia,* vol. 11, no. 7, pp. 740-756, 2015.

[23] O. Voevodskaya, A. Simmons and others, "The effects of intracranial volume adjustment approaches on multiple regional MRI volumes in healthy aging and Alzheimer's disease," *Frontiers in Aging Neuroscience,* vol. 6, p. 264, 2014.