1 Introduction

Alzheimer’s is a progressive, neurodegenerative disease, that causes irreversible damage to the brain tissue. It impairs the ability to form and retrieve memory, and eventually disrupts the natural flow of life, by affecting the ability to carry out even day to day activities. The disease is typically diagnosed from the symptoms (Mini Mental State Examination, [8]), such as decline in cognitive abilities, visual and/or speech impairment, loss of memory, rather than the structural changes in the brain (biomarker) that causes it. But the pathological changes in the brain start decades before the manifestation of the symptoms [7]. Magnetic Resonance Imaging (MRI) is capable of capturing the complex changes in the brain, even if it is difficult for humans to extract those features from the low contrast, multi-dimensional MRIs [1]. There is a considerable amount of work on analyzing Alzheimer’s disease. However, the vast majority intends to predict the state of the disease at the current time step.

In this work, we aim to learn early indicators of the disease and focus on predicting the chances of a person getting affected by Alzheimer’s in the near future (6 months), from their current T1 structural MRI (sMRI). To that end, we train end-to-end deterministic baselines. However, since there are various ways in which a disease might evolve from a patient’s current physiological state, it is even more useful to model the distribution over the possible future states of the disease. We address this problem by learning a distribution over possible disease progressions. Results on the Alzheimer’s Disease Neuroimaging Initiative (ADNI) dataset [19] suggest that sMRI contains visual information that characterizes Alzheimer’s dementia, and highlights the potential of distribution-based models to identify boundary cases, which may require different treatment or follow up.

2 Data Description

Finding open source medical datasets is difficult due to data access and patient privacy issues. Collecting data being expensive, finding a dataset large enough to train deep learning models is even
more challenging. As more data often beats sophisticated algorithms, scarcity of data is a bottleneck in training high performing models. ADNI [19] is an excellent initiative to remove the obstacle of data availability from research, but in terms of size, the dataset is nowhere close to the benchmark datasets that state of the art deep learning models are trained on (see e.g. [5, 17]). For the experiments described in this abstract, we used images of the ADNI dataset, denoised, intensity normalized and registered as in [2, 10, 3, 11]. Labels in the ADNI dataset are not extracted from the MRIs, but from the MMSE scores. MMSE score is not a reliable indicator of the presence of Alzheimer’s or cognitive impairment since other factors such as educational qualification can heavily affect the score [4]. Unless the patients are at an advanced stage of the disease, their performance at the test might vary. These factors make the labels noisy, as for the same MRI, the label might be different on different days. Due to this reason, the dataset also has evidence of retraction of the disease, whereas Alzheimer’s is a strictly progressive disease.

Brain MRI can be visualized as a stack of 2D images with each slice capturing brain anatomy at a different depth. Intrinsically brain MRIs are high dimensional. Images used in this research are of dimension 233 x 197 x 189 voxels. In coronal view, 233 is the dimension of the depth, the others being the shape of the 2D slices. We have chosen coronal view over axial or sagittal view in order to have maximum visibility of the hippocampus, which is known to be one the first regions affected [9]. We have not used 10 slices from far end of either side of the depth dimension as they are mostly background. Even after that, the overall dimension of the images remains high, hindering the application of recent architectures containing vast amounts of parameters. Additionally, Alzheimer’s being an active area of research in the medical domain, even radiologists are not fully confident on the biomarkers of the disease. Hence, we choose to deal with full brain MRIs as opposed to relying on 2D information exclusively [12, 18], pre-selecting a region of interest or downscaling the volume resolution [20], even at the cost of higher dimensionality.

Finally, we considered 4046 MRIs from 1092 patients, for whom at least 2 images were available throughout the course of study. To avoid the presence of MRIs of the same patient in both the training and validation/test set, the data is split into training, validation and test set with 8:1:1 ratio at the patient level. Training, validation and test set consist of 3257, 396 and 393 MRIs from 873, 109 and 110 patients, respectively.

3 Models

All the models described in this section are trained with the MRIs at one time stamp and disease labels from MMSE score at the next time stamp. We applied a volume-wise Gaussian normalization to each MRI and used them as input to our models, with the goal to predict the probability of a patient being normal or diseased (MCI or AD) in 6 months.

As deterministic baselines, we trained a 3D Convolutional Neural Network (CNN) with crossentropy loss. Figure 1a depicts the model, with interleaved convolution and pooling layers, followed by fully connected layers and a classifier. To improve generalization, we augmented the 3D CNN with an upsampling path trained to perform reconstruction. This model is referred to as auto-encoder (AE) and also operates on 3D volumes (see Figure 1b). However, the mapping between current MRI and the next state of the disease is not necessarily one to one. The progression of any disease depends on various salient factors and hence, a patient might evolve in various ways at different pace from their current physiological state. This uncertainty in the disease progression can be captured either in the categorical distribution of the output (classifier) or the latent variable, such as in Variational Autoencoder (VAE) [15]. To that end, we introduce a third model based on 3D convolutional VAE and endow it with a classifier, which we connect to the latent variable (see Figure 1c). The model is trained to optimize for both negative evidence lower bound and crossentropy.

4 Empirical Studies

All models were trained with Adam optimizer [14], with a learning rate of 0.0001, which was reduced a factor of 10 after 15, 25 and 35 epochs. We applied dropout [21] with 20% probability to all of the convolutional layers. Mini-batch size varied between 2 and 4, and was set to maximize GPU memory utilization. The reconstruction hyper-parameter $\lambda$ was set to 2.5 for AE and to 1 for VAE.
Figure 1: **Model Architectures.** (a) CNN: All Convolutions are 3D with kernel size 3 and 11 channels. Each convolution is followed by a 3D maxpool and a 3D batch normalization layer [13]. All of the hidden layers are passed through the ReLU non-linearity. The fully connected layers have 4096 neurons each. (b) AE: Composed of an encoder (same architecture as CNN) and a decoder, mirroring the encoder. (c) VAE: same encoder-decoder architecture as AE. The latent code dimension is 1024.

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
<th>F1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN</td>
<td>76.77%</td>
<td>0.65</td>
</tr>
<tr>
<td>AE</td>
<td>80.05%</td>
<td>0.71</td>
</tr>
<tr>
<td>VAE</td>
<td>75.76%</td>
<td>0.61</td>
</tr>
</tbody>
</table>

Table 1: **Performance.** AE performed best, followed by the CNN and VAE. Highest true positive rate of AE makes it the best classifier. Note that VAE learns the distribution over the future labels, and might assign different labels on different runs.

Table 2: **Predicting distribution over future disease states.** For each MRI, the trained VAE predicts the future label 20 times. Patient 1, 2 and 3 are not always predicted normal; they might require different treatment/follow up than patient 4.

Classification performance is reported in Table 1 in terms of accuracy and F1-score. The trained CNN, regularized with AE achieved 81.93% test accuracy when predicting normal and diseased patients of ADNI dataset. Note that random accuracy on test set is 60.55%. Moreover, we used the trained VAE to predict disease states of 4 randomly selected Normal (NL) patients multiple times (see Table 2). As shown in the table, by sampling multiple times from the trained VAE, the classifier makes different predictions for the same patient, potentially identifying boundary cases, which might benefit from specific treatment or follow up. Finally, Figure 2 depicts the t-SNE visualization of the representations extracted from the final fully connected layers of the models. It is worth mentioning that, for VAE, mis-classification errors are mostly driven by those uncertain cases.
The results discussed in this abstract suggest that sMRI contains visual information that characterizes Alzheimer's dementia, paving the road towards addressing increasingly challenging questions. A natural next step would be to perform a thorough analysis on the trained model, to better understand how classification decisions are made. This could lead to confirming the presence of known patterns in the sMRI scans of the patients affected by Alzheimer, and eventually discover new relevant features that characterize the dementia. Finally, a particularly interesting research direction would be to learn a predictive model of the disease conditioned on a sequence of past image acquisitions. We envision that such approach could be reminiscent to the recent work on video generation presented in [6]. Moreover, we could condition the posterior distribution over latent codes on both the input volume and the label, following the work on segmentation of ambiguous medical images introduced in [16].

References


