석 사 학 위 논 문
Master's Thesis

작은 데이터 셋에서 딥러닝을 활용한 알츠하이머 치매 진단

Alzheimer's disease classification using deep learning: A novel small-data approach.

2021

하시브 라자 (Haseeb, Raja)

한 국 과 학 기 술 원

Korea Advanced Institute of Science and Technology

석 사 학 위 논 문

작은 데이터 셋에서 딥러닝을 활용한 알츠하이머 치매 진단

2021

하시브 라자

한 국 과 학 기 술 원

전기및전자공학부

작은 데이터 셋에서 딥러닝을 활용한 알츠하이머 치매 진단

하시브 라자

위 논문은 한국과학기술원 석사학위논문으로
학위논문 심사위원회의 심사를 통과하였음

2021년 5월 18일

심사위원장   김 종 환   (인)

심 사 위 원   명 현   (인)

심 사 위 원   하 동 수   (인)

Alzheimer's disease classification using deep learning: A novel small-data approach.

Raja Haseeb

Advisor: Jong-Hwan Kim

A dissertation submitted to the faculty of
Korea Advanced Institute of Science and Technology in
partial fulfillment of the requirements for the degree of
Master of Science in Electrical Engineering

Daejeon, Korea
May 18, 2021

Approved by

_____

Jong-Hwan Kim
Professor of Electrical Engineering

The study was conducted in accordance with Code of Research Ethics[1].

Abstract

Alzheimer's disease (AD) is a progressive neurodegenerative disease that causes cognitive impairment among elder people. It is the most common cause of dementia. Early and timely diagnosis of AD can help with the overall treatment process of the patients. Classifying AD patients from mild cognitive impairment (MCI) and cognitive normal (CN) patients is very important for AD diagnosis. Recently use of machine learning techniques and deep learning-based algorithms have become a popular choice for the AD classification task. However, most of the existing approaches depend upon large data sets or suffer from data leakage. In this work, we present a novel framework for AD classification in a small data regime. Our approach is based on three main steps: 1) PGGAN based medical image generation to deal with data scarcity; 2) Model pretraining using SimCLR framework; 3) Train the pretrained model for the final classification task. We used ResNet-18 in combination with CBAM (Convolutional Block Attention Module). The CBAM module enhances the useful features in the images, which helps in better training of the model. We performed a clinical evaluation of our model on novel test data set. We achieved an accuracy of 83% for the AD vs. CN classification task using only a few slices for the training process. We also compared our model with the previous approaches and found our results to be comparable even with such small data. Results indicate the effectiveness of our proposed framework and can help in the early diagnosis of AD patients.

Keywords Alzheimer's disease classification, Alzheimer's disease diagnosis, convolutional neural network (CNN), machine learning, deep learning, contrastive learning, CBAM, generative adversarial network (GAN), self-supervised learning.

Contents

List of Tables

List of Figures

Chapter 1. Introduction

Alzheimer's disease (AD), also referred to simply as Alzheimer's, is a progressive neurodegenerative disease. It is the most common cause of dementia in the world, affecting 1 out of 9 people over the age of 65 years [18]. In 2018, 5.7 million Americans were reported to have AD. This number can rise to 14 million by the year 2050 [18]. AD causes a loss of connection between nerve cells in elders. AD results in brain shrinkage, cause enlargement of brain ventricles and decrease the size of the hippocampus. Alzheimer's disease progressively causes cognitive impairment, commonly associated with early memory loss, affecting a person's ability to think and perform common day-to-day tasks. The available treatments for AD at the present can only slow the progression at the best, and there is no validated cure for the patients that already have AD. Therefore, early diagnosis of AD is crucial for the timely treatment of patients and can help to abate the progression of the disease. Based on this, the concept of Mild Cognitive Impairment (MCI) was introduced as the prodromal form of AD [23]. It describes people having mild symptoms of brain impairment. The MCI patients are still able to perform daily activities up to some extent. However, their ability to do so declines with time as the disease progresses and the patients in this phase have high chances of progressing into dementia [18][8][87][200]. As a result, patients are classified into three main categories i.e. AD class representing the patients with the disease, Cognitive Normal (CN) people which have normal brain function, and the MCI patients which represents the stage in between the CN and AD patients.

Consequently, physicians have been trying to apply various clinical methodologies to classify and diagnose AD patients. Cerebrospinal fluid (CSF) concentration in the brain is reported to indicate the presence of AD. As the disease progresses, the level of norepinephrine increase in the CSF. A ventricular puncture is used for the collection of CSF; the physician collects CSF directly from one of the brain ventricles by making a hole in the skull [45]. This process can be arduous and can cause bleeding in the brain. A lot of focus has been put on the development of medical imaging techniques in recent years. Neuroimaging techniques like Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), Functional MRI, and Computerized Tomography (CT) are widely used to diagnose functional and structural changes in the brain. MRI modality is commonly used for the diagnosis of AD, as it is easily accessible and provides good results. MRI sequences are usually T1 or T2-weighted scans. It is evident from research that the progression from CN to MCI to AD results in certain changes in the brain structure. A healthy brain can be distinguished from a diseased brain by using morphological changes in the brain structure, volume, and texture via medical imaging [228][134]. Changes in hippocampus size can be an early indicator of AD and are useful for AD classification [2]. The area of the brain affected by AD can be estimated by observing changes in the white matter (WM) [138]. Gray matter (GM) can help in analyzing AD [107]. However, all these methods for diagnosis require specialists to analyze the neuroimaging sequences and make a decision. This process can be costly, laborious, time-consuming, and sometimes even prone to human errors. Therefore, there is a need for an automated way for the classification of AD patients which takes less time and effort, is reliable, less costly. Such a process can help the practitioners speed up the diagnosis process and treat the patients at an early stage.

This is where the exciting field of machine learning comes in. This is the age of machine learning and artificial intelligence. These are the most interesting areas for many engineers, researchers, and practitioners these days. It is the simulation of human perception and intelligence by computer systems.

These days everyone is utilizing modern machine learning tools somehow. The motivation behind this is pretty obvious; computers and machines can be utilized to perform repetitive tasks that can be fatiguing and tiresome for human beings. Recently, machines have shown tremendous success in the implementation and learning of very complex tasks, and have shown promising results in the decision-making process. In all this, the most exciting part is the finding that in some cases, machines and computers can identify patterns that are even beyond the human perception [46].

The paradigm of machine learning mainly involves learning from experience. Given a set of training examples, the algorithm can learn the underlying features of data and can make a decision based on those features or can even predict future data. In recent decades, various algorithms have been designed for the implementation of machine learning. These approaches have been for various tasks like classification, regression, clustering, and so on. In the early days, traditional machine learning approaches were used which required a human expert to select appropriate features from the data. This process is known as feature engineering. These features are then passed into the model. Typical examples of such algorithms include logistic regression, support vector machines (SVM), decision trees, random forest, and so on.

Deep learning [209][183][117][61][131] is the latest paradigm in machine learning which involves high-level abstraction of features. In traditional machine learning, a domain expert is required to abstract useful features from the data to reduce data complexity and make the underlying pattern more visible to the algorithm. Deep learning is a more sophisticated evolution of traditional ML, which eliminates the need for a domain expert and extracts features from the data in an automated way. In the past few years, deep learning techniques completely transformed the way researchers and data scientists thought and achieved state-of-the-art results in tasks across multiple domains. It is a type of neural network, instead of linear mapping utilizes convolution filters and many hidden layers along with various activation. Numerous deep learning architectures have been developed in recent years including convolutional neural networks (CNNs) [118][110], recurrent neural networks (RNNs) [125], recursive neural networks [198], auto-encoders [147] and Generative Adversarial Networks (GANs) [62]. The details of these deep learning models and architectures are out of the scope here. However, some important papers are referenced above.

Deep learning can be used for various tasks. These tasks include classification [110][194][206][207][76], object detection [172][175], object tracking [243], depth estimation [112][126], semantic and instance segmentation [33][132][148][70], speech recognition [13][66], natural language processing (NLP) [38][199][60] and so on. Deep learning models have a wide range of applications across various domains like autonomous driving, robotics, fault detection, visual recognition, natural language processing, fraud detection, anomaly detection, image captioning, pixel restoration, surveillance, healthcare and so on.

Such discoveries aroused the interest of data scientists and researchers in the field of machine learning and also adapting it to the medical domain. Deep learning caught the eye of many healthcare practitioners, and it did not take long for researchers to realize yet another opportunity for the application of deep learning. The healthcare sector is not small anymore. Due to the increase in the number of patients and medical devices, the amount of medical data is quite large. Therefore, it makes medical diagnosis and analysis of the patient result more challenging. Such rapid growth of medical data requires tedious efforts by doctors and medical experts and are usually susceptible to human error, and can have various interpretations among different experts. There is a need for efficient and effective tools and methods to help speed up the diagnosis process by proper management and processing of this large data while maintaining high levels of accuracy and performance [170].

Soon enough deep neural networks found their application in the medical and healthcare sector for diagnosis of various diseases [149]. This lead to the use of deep learning methods for the recognition
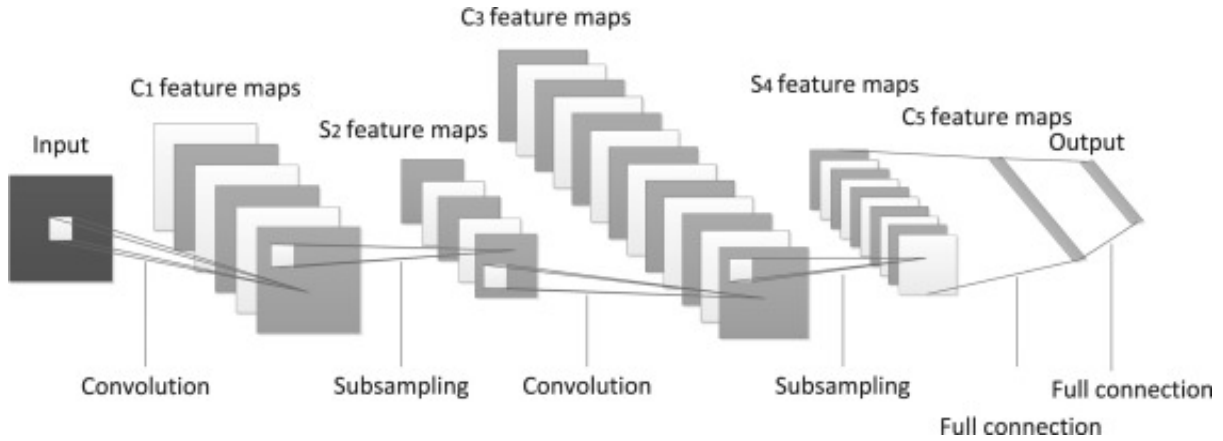
Figure 1.1: A schematic diagram of standard CNN [131].

of Alzheimer's Disease (AD). Many scientists and researchers started applying deep learning models to medical biomarkers and imaging data [129][130][202]. As expected, deep learning showed high performance in AD diagnosis like every other field. Modern deep learning algorithms like stacked auto-encoders (SAE), residual networks, and Deep Boltzmann Machine (DBM) outperformed previously used traditional machine learning algorithms like SVMs and random forest.

However, despite this tremendous success of deep learning techniques across multiple domains, one important issue still needs to be addressed, and researchers and data scientists are making continuous efforts to address that issue i.e. the need for a large training data set for deep learning models. Although deep learning models show promising results with a very high level of accuracy, this requires a very careful training process along with a large number of training examples. It is a common belief among all the researchers that the more data you feed into your network, the better the performance will be. So you need a sufficiently large volume of data to train a successful machine learning algorithm [117][43]. Data with high-class imbalance or insufficient data can lead to the poor performance of the algorithm [191]. In the case of common real-world images, this is not a problem as we have enough training data available from various resources. However, this issue becomes more prominent when we talk about domains like medical and healthcare.

Despite the rapid growth in medical records and data, there are many cases when training data is not sufficient. One main reason is that patient data is well protected by the patient data laws. Also, it is expensive and time-consuming to obtain more medical data especially annotated ones since you need a medical expert for that purpose. The imaging and data retrieval standards also differ from country to country and even from one hospital to another, which makes the process even more complicated. In recent years, some big hospitals and authorities made some medical data available to the public in an anonymous way to further drive deep learning research. However, still, this data is not big enough to properly train the current deep learning models. Also, the associated annotation varies considerably due to differences in data acquisition systems and different annotators. Therefore, in clinical settings, usually small data sets are obtained. Creating large data sets is usually time-consuming and might not be feasible especially in the case of some rare diseases or new imaging modalities [145]. For example, in the case of the recent COVID-19 infection, the training data is still an issue for many researchers [154] and practitioners are in continuous efforts to obtain more data through clinical assessments, which is going to take quite some time because the infection is novel and trials are still in early stages.

This problem is more significant, especially when employing supervised learning algorithms that

require large labeled training examples. As mentioned earlier, the collection of large medical data is a complex, expensive, and time-consuming process which requires the collaboration of radiologists and researchers [68]. In recent years, many researchers have tried to overcome the issue of small data sets by using various techniques. It has been and still a very important research area for many researchers and data scientists across the world. There have been many attempts to resolve this issue in different domains by using various networks and techniques like data augmentation and synthetic data using GANs [62] [208], Meta-learning [217], using surrogate data, n-shot (zero-shot, one-shot and few-shot) learning [56][237][197][233] and Siamese networks [111][146]. For the case of limited labeled data (since the annotations are done by experts which are costly), there are other learning methods like unsupervised learning [142][21], semi-supervised learning [244] and more recently self-supervised learning [91], which do not require much-labeled data for training purposes and can still achieve higher accuracies.

Such methods also got the attention of researchers and experts working with medical data and the healthcare sector. Recently there have been several attempts to work with small medical data. Researchers have employed methods used in other domains in the healthcare sector. Such work can help mitigate the need for huge medical data required for training the algorithm and can help reduce the time and cost of the annotation process. This research presents a novel approach to deal with medical data scarcity for the case of Alzheimer's Disease (AD) diagnosis and classification. A Progressive GAN (PGGAN) [96] based data augmentation approach is combined with a recent simple framework for contrastive learning of visual representations (SimCLR) [34] and Convolutional block attention module (Cbam) [231] to achieve better results on AD vs. CN vs. MCI classification using small data set. The details of the algorithm and methodology are explained in the later sections.

Chapter 2. Related Works and Contributions of this Dissertation

2.1 Alzheimer's Disease Classification

In the recent few years, computer-aided diagnosis (CAD) has been utilized to assist physicians. Many researchers and data scientists have been putting efforts to develop CAD systems to diagnose AD. Most of the practitioners rely on Mini-Mental State Examination and other physical tests to verify the stage of AD [228][106]. Clinically, AD classification is done by developing different biomarkers and by collecting different parameters to test the AD stage [1]. Machine learning has been deployed in recent CAD systems to examine and analyze the pattern in the medical data. Machine learning has shown promising results with better accuracies based on the features extracted from single and multi-modal brain images [188].

There have been various attempts to classify CN, AD, and MCI patients using image data, cognitive features, and other medical biomarkers. A random forest-based approach is used for AD classification using CSF measurement, genetic information, shape, voxel intensity and brain volume as feature values as features to the network [67]. The dimensionality of these features is reduced using PCA and classification is performed using particle swarm optimization (PSO) and support vector machine (SVM) [236]. Brain tissues like WM, GM, and hippocampus are affected as AD progresses. Texture changes in GM and WM are also a good indicator to distinguish AD, MCI, and CN patients. These changes are used for patient classification using SVM [173] and KNN [155]. Hybrid features generated by the combination of volume and texture information such as GM volume along with texture features are used to perform AD classification using SVM-Random Fourier Expression (SVM-FRE) [234]. A combination of features extracted from clinical data and segmented brain images is used for multi-class classification of AD from CN and MCI [11]. A novel automatic classification and tissue segmentation method based on Independent Component Analysis (ICA) and SVM classifier was proposed in [94] where the input to the network is a combination of T1, T2, and proton density (PD) scans. Connectivity information from fMRI data was used in [190][80]. Vemuri *et al.* [219] observed improved classification for CN vs. AD by combining sMRI scan features with genetic and demographic information of the patients. Various researchers also focused on identifying and extracting the most discriminative regions and features for AD classification task [186][139][57][171][31][164][186][158][156][184][165]. Tangaro *et al.* [210] used a fuzzy-based classes approach for hippocampus volume to classify MCIc vs. MCIn and AD vs. CN. Several other researchers also utilized SVM-based (linear, polynomial and RBF kernel) approach for AD classification [100][101][144][64][107][113][239][114]. Some works [239][30] also used wavelet-based features. Kernel ensembles are also adopted by some researchers. Alam *et al.* [7] used multiple kernel SVM for AD vs. MCI vs. CN classification. A combination of RBF, linear and polynomial kernel is used by Kamath *et al.* [94]. A multi-modal (features from MRI and genetic data) and multiple SVM kernel [162] approach is presented for the classification of subjects. SPECT images of Alzheimer's and normal subjects are used as input to a contiguous SVM (CSVM) classifier for more spatial information [55]. A twin support vector machine (TWSVM)-based approach is used by Zhang *et al.* for AD vs. CN classification [239]. Improved prediction accuracy is reported in various research for AD classification using ensemble-based SVM [12][84][178]
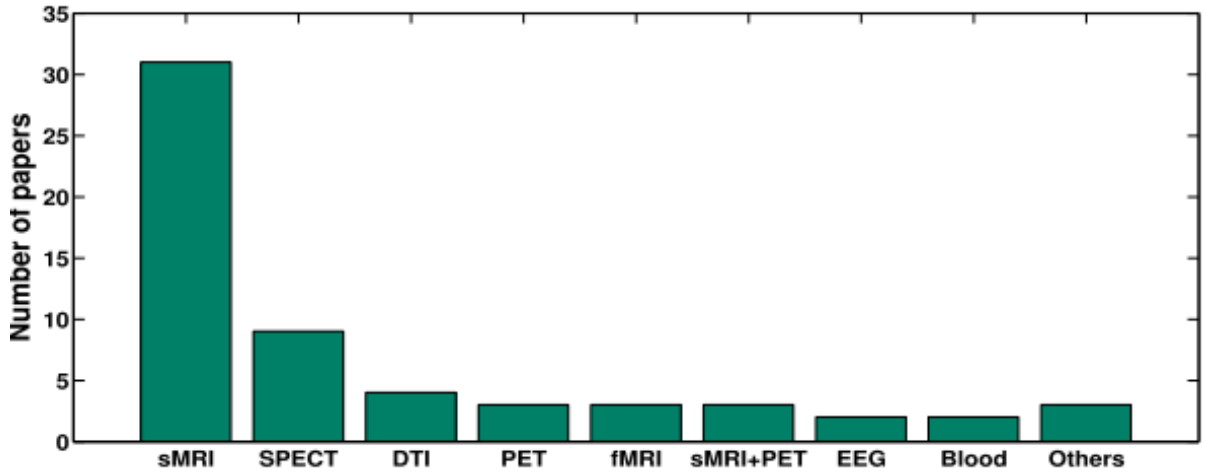
Figure 2.1: Plot showing data and image modalities used with SVM for Alzheimer's in various research [211].

Artificial neural networks (ANNs) can model complex patterns of non-linear data and are highly used for machine learning (ML) applications. Many researchers adopted ANNs for AD classification. A standard ANN is shown in Figure 3. EEG signals are used to classify Ad vs. CN with ANN [41][179]. A comparison between different variants of ANN by Savio *et al.* [182] and they found that learning vector quantization neural network (LVQNN) shows best results. Long et al. [133] utilized an unsupervised learning approach for AD vs. CN classification. A comparison is made between various algorithms like random forest (RF), decision tree, bagging, multi-layer perceptron (MLP) and ANN for AD classification [93]. Mahmood *et al.* used ANN in combination with PCA to reduce the complexity of high dimensional data [140]. ANN is used for AD vs. CN classification [6]. A bidirectional gated recurrent neural network (BGRU) is cascaded with an MLP for AD classification [40]. AD subjects are classified based on sequential auditory data by using an LSTM (long short-term memory) network [54]. To discriminate AD subjects from CN, Kar *et al.* [95] used an ANN with a fuzzy-based approach on features from DTI images. Recently, deep learning (DL) has shown good results and success in AD classification [189]. Suk *et al.* reported improved accuracy for AD vs. MCI classification by applying a stacked auto-encoder (SAE) [203]. A deep Boltzmann machine (DBM) is used for AD vs. MCI vs. CN classification using patches from PET and MRI data [202]. To perform AD classification, an entropy-based convolutional neural network is used to extract radiomic features from 3D MRI images [27]. A multi-scale deep neural network is proposed based on multi-modal data from FDG-PET and MRI images by Lu *et al.* [135]. Segmented gray matter is divided into patches which are then used for feature extraction. Diagnosis of AD and its stages is performed using a deep CNN on MRI scans [47]. 3D-CNN is reported to perform better than SAE on AD classification task [161]. To classify CN, AD, MCI patients, an SAE-based classifier is used on top of DL-based latent features [201]. A multi-modal deep learning architecture based on SAE is presented for the AD diagnosis [129]. To classify AD, MCI, and CN subjects, Hosseini *et al.* [82] presented a deeply supervised adaptive 3D-CNN (DSA-3DCNN) using MRI data. The network was pretrained by a 3D convolutional auto-encoder (3D-CAE). High classification accuracy is obtained by training a DBN-SVM classifier on extracted 3D patches [157]. Multiple GRU architectures are used to extract features from multi-model patient data like MRI, cognitive performance, CSF volumes, and demographic information [119]. Segmented GM voxels from T2-weighted MRI are enhanced via hybrid
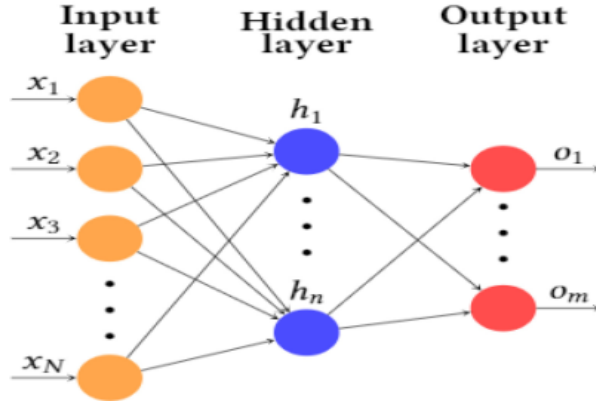
Figure 2.2: Standard Neural Network Architecture [211].

enhanced independent component analysis [23]. A CNN is then used on these voxels to classify AD, MCI, and CN patients. A deep learning model was trained on fMRI data for the first time by Sarraf *et al.* [181]. Data from multiple modalities i.e. PET and MRI are used for feature extraction [65]. An LSTM-based approach is presented for AD vs. CN classification [81]. Basaia *et al.* utilized T1-weighted 3D sMRI data to build a deep network that was able to differentiate AD, CN and MCI patients [22]. FDG-PET data is used to train a multi-scale deep neural network (MDNN) [136]. A VGG16 [195] network pretrained on Imagenet is used in a transfer learning approach to classify AD patients with sMRI slices [88]. Khan *et al.* presented another transfer learning approach for AD prediction based on VGG19 architecture [98]. To handle class imbalance, a data augmentation approach is presented [4]. To enhance performance for AD vs MCI diagnosis on sMRI images and reduce overfitting, Whang et al. [222] proposed an ensemble of 3D-DenseNets.

2.2 Medical Data Scarcity

Data is the driving force for any machine learning model. More training data means better training hence better performance. However, as mentioned in previous sections, large data sets are not always available, especially in medical diagnosis. This section mentions some of the recent works and methodologies to deal with the issue of having small training data in medical cases.

2.2.1 N-shot learning methods

One way to deal with small data sets is to use n-shot learning techniques [48][51]. The aim of few-shot learning (FSL) is to learn good representations given a very small amount of training examples [196][220][174][73]. Contrary to the classic method of supervised learning in which the model is fed with a large number of training samples, here we only feed the model with a limited number of training samples per category [226]. One-shot learning (OSL) and zero-shot learning (ZSL) are special cases of n-shot learning in which the model is fed with only one/few training samples or with zero training samples respectively. So zero-shot learning deals with situations where at inference time, the model observes a data sample that was not present in the training data set and still needs to predict the correct category to which the new example belongs to [225].

Kim *et al.* leveraged few-shot learning for glaucoma diagnosis using high-quality FUNDUS images [103]. Inspiration for this came from the earlier work which combined deep learning with few-shot

learning and demonstrated high potential [180]. They employed Matching Networks (MNs) introduced in [220] along with few-shot learning. Another such approach makes use of a deep Siamese neural network [108]. The intuition behind the approach is that the Siamese network will learn the distance and feature representation from the textures present in the source domain, and then apply this learned knowledge to classify the tumoral and healthy tissues in a target by using few-shot learning. A three-headed siamese network is presented [145]. During training, embeddings for images that belong to similar classes are forced closer and those that belong to different classes have their embedding vectors pulled apart. This embedding vector is the set to train a classifier like Support Vector Machine (SVM). One important medical task in drug discovery, which unfortunately also lacks the amount of data available for predictions. to resolve this issue, various research adopted one-shot learning [10][48][51].

Segmentation of the medical results is also an important application that can help radiologists in identifying various diseases. With enough labeled data, supervised deep learning-based segmentation methods have shown state-of-the-art results. However, manual segmentation labeling of medical images is resource and time-intensive. Also, the use of different image acquisition procedures can cause wide variations in tissue appearance, image noise and resolution [120]. One approach to mitigate the issue of less labeled data is to synthesize realistic and diverse labeled examples by leveraging unlabeled examples [241], which results in an automated data augmentation method for synthesizing labeled medical images. A novel few-shot learning framework has been presented by Feyjie *et al.* [50]. Training is done in an episodic manner where unlabeled data is also presented during each episode. A semi-supervised few-shot learning approach is presented for medical image segmentation [50]. By making use of information from two different modalities i.e. CT and x-ray, generalized zero-shot learning can prove valuable for the diagnosis of chest radiographs [160].

Recently, there has been an outbreak of novel coronavirus (COVID-19) which is affecting the whole world. More than 132 million cases of COVID-19 have been confirmed as of April 2021, with a death toll of around 3 million the world, according to World Health Organization [230]. Researchers are trying to develop approaches for the diagnosis of COVID-19 using deep learning. However, due to the rare nature of this virus, large data sets are still unavailable. To deal with this situation, a few-shot-based approach has been presented to make an accurate prediction by using a small number of chest CT images [35]. The authors employ the prototypical network [196] to apply a few-shot classification. Experimental results indicate that the proposed methods achieve better performance than ResNet-50 in limited data cases.

2.2.2 GAN-based methods

Generative Adversarial Networks or GANs have been used extensively for data synthesis and augmentation in recent years [62][208]. They became quite popular due to their immense success across various tasks. Their most significant breakthrough has been in the domain of computer vision where they helped in dealing with challenges among various tasks like image generation, facial attribute manipulation, style transfer, and so on [227]. Due to their immense popularity, GANs have also been used in the medical domain nowadays. Research has been conducted in the health sector to make use of these networks for data augmentation and synthesis to help deal with the issue of small and less annotated data.

In their research, Costa *et al.* presented a way to generate new eye FUNDUS images from the existing data using adversarial learning [39]. In another work, the authors attempted to generate CT images given the MR images by using GANs [151]. GANs can are used to synthesize high-quality and diverse medical images for liver lesion classification task [53]. Along with bone and lungs, the liver is a common site for

metastatic cancer. In 2012 alone, 8.2 a million deaths were caused by cancer worldwide, out of which 745,000 are due to liver cancer [49]. Work has been done by researchers to automate the classification process using CT images and help radiologists in the diagnosis procedure [59][29]. The authors of [53] adopted DCGAN presented in [167] for generating synthetic liver lesion in combination with standard image perturbation techniques for the improvement of liver lesion classification. An innovative GAN-based method has been proposed in [71] for image augmentation. This is done by adding the bounding box conditions in the original PGGANs [96].

Besides providing a diverse and additional form of data augmentation, GANs can also serve as an effective model for the anonymization of medical data [192]. In recent years, researchers have been utilizing noise-to image GANs [58] or image-to image GANs [182] to improve the classification results with augmented images. An innovative method has been presented by Han et al. [72]. A two-step GAN-based data augmentation technique is presented. First, PGGANs [96] generates 256×256 diverse and realistic images. In the second step, a Multimodal Unsupervised image-to-image Translation (MUNIT) [83] model which is a combination of GANs/Variational AutoEncoders (VAEs) [104] or SimGAN [193], is used to refine the shape and texture of the images generated in the first step which furthers fits them into the distribution of real images. Results indicate that the presented combination of noise-to-image and image-to-noise GANs for augmentation of medical images improves the accuracy of CNN-based classification for tumor detection, especially when combined with classic DA techniques. Two publicly available data sets of brain MRI have been used by Shin *et al.* [192] including Alzheimer's Disease Neuroimaging Initiative (ADNI) data set and Multi-modal Brain Tumor Image Segmentation Benchmark (BRATS) data set. To translate MRI to labels and vice versa, image-to-image translation conditional GAN (pix2pix) model introduced in [86] has been utilized.

2.2.3 Meta-learning methods

In recent years, meta-learning has gained a lot of popularity and more meta-learning techniques are being developed. The term meta-learning or learning to learn refers to techniques that leverage prior experience gained on other tasks [217]. We can leverage more types of meta-data depending on the similarity index between the new task and the prior task. There are different types of meta-learning like metric-based, model-based, and optimization-based.

Like GAN-based synthesis, meta-learning also found its way into the medical regime. Maicas *et al.* citemaicas2018training learned good initialization on a series of tasks by utilizing a meta-learning-based method. A novel difficulty-aware meta-learning (DAML) method has been proposed to deal with this issue especially for rare disease classification [124]. First, a meta-classifier is trained on a series of related tasks like the classification of common diseases. Then these learned internal representations are transferred to the new task of rare disease classification. For the better optimization of the meta-classifier, the authors formulated a dynamic modulating function over the learning tasks. The well-learned takes are automatically down-weighted by the function which puts more focus on the hard tasks. The proposed methods also outperform standard fine-tuning methods and widely used few-shot learning approaches like Relation Net [204], MAML [52] and Task sample [141] for the task of rare disease classification.

Domain generalization is one of the most crucial tasks in deep learning i.e learning a model from multi-domain data and generalizing it to a new and unseen target domain. Various attempts have been made to address the domain generalization problem [20][44][121]. Such tasks become even more complex in the healthcare sector due to the complexity of medical images. To tackle this issue, recently a novel shape-aware meta-learning (SAML) scheme is presented by Liu *et al.* [128] for the improved model

generalization in prostate MRI segmentation.

2.2.4 Surrogate methods

Some other methods utilize some surrogate data or surrogate supervision to handle small data sets in the medical sector. Shaikhina *et al.* [187] employ surrogate data analysis and method of multiple runs for the application of artificial neural network to regression problems involving limited medical data sets. Such surrogate data set resembles the original data set in terms of the statistical properties. However, the surrogates do not keep the complex interrelationships among various components of the original data set. Various methods [185][215][122] can be used for the generation of surrogate data and can be validated via NN techniques.

Motivated by the work done of Shaikhina *et al.* [187], Sabay *et al.* [177] also utilize the surrogate data generation method to overcome the problem of the limited data set in the prediction of heart disease. However, the authors claim that the number of neurons and computational resources required in the prior work is large. Hence, an alternate solution for the generation of surrogate data is presented. There various tools available for that purpose. One such library is Synthpop, which is an R language library and provides functionality for data comparison and synthesis [153].

2.2.5 Self-supervised learning methods

Self-supervised learning enables us to utilize a variety of labels that are already present in the data. As we already know that generating a set of annotated examples is both time and resource-consuming. However, a huge amount of unlabeled is available most of the time. Therefore, an efficient solution would be to use this unlabelled data somehow to generate labels for the training process. Self-supervised learning enables us to do so by setting learning objectives in a way such that the data itself provides supervision.

Self-supervised learning has emerged as the new paradigm in recent years due to its effectiveness [89][91] especially for the case of small annotated data sets. It is clear from our previous discussions, that one domain the suffers most from the curse of small datasets and annotations is the medical sector. Therefore, in recent years, researchers have made attempts to utilize self-supervised learning methods to mitigate the need for large annotated datasets. One such approach has been presented by Tajbakhsh *et al.* [?]. The underlying idea presented in the paper is to utilize the large available unlabeled medical data and assign labels to it. This is done by a process knows as surrogate supervision. This is mainly employed at the pre-training stage, in which a deep learning model is pretrained using surrogate supervision including tasks like rotation, reconstruction, and colorization. The findings of the research indicated that pre-training with surrogate supervision is effective and can outperform the models trained from scratch or transfer learning from an unrelated domain, especially in the case of medical data.

Some of the prior research proposed various surrogate tasks to apply self-supervised learning with surrogate supervision in the medical domain. However, the literature is still very limited, since it is still a new trend and requires more attention. Jamaluddin *et al.* [90] proposed a surrogate task based on the longitudinal relationships between medical images. Alex *et al.* [9] used small image patches for noise removal as the surrogate task. Ross *et al.* [176] employed image colorization as the surrogate task. Rotation prediction [58] is used as the surrogate task in cases when the geometry is consistent and data has adequate landmarks.

GAN-based surrogate supervision is used where rotation might not be the appropriate choice. Such
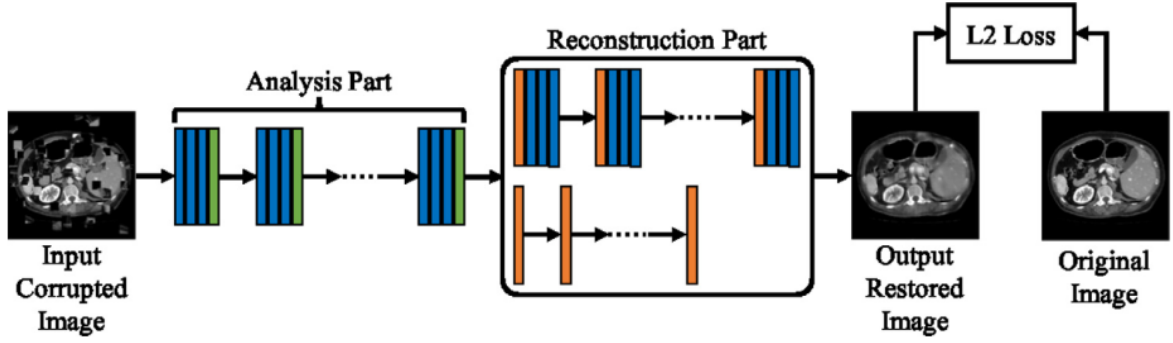
Figure 2.3: General CNN architecture self-supervised learning based on context restoration pretext task [32].

schemes include image colorization and patch reconstruction tasks by using conditional and Wasserstein GAN respectively [115][15]. Such surrogate tasks enable the model to learn the relative location of various structures in medical data, distinguish between them, and learn the overall geometry of the data. An early example of the adaptation of self-supervised learning to medical diagnosis was provided by Ammara *et al.* [143]. The authors proposed a classification model in which a deep belief net is used as the architecture for classification and another SASVM (self-advised support vector machine) alongside the first classifier.

A novel pretext task for self-supervised learning has been presented by Chen *et al.* [32] for medical data, which they termed as context restoration. Input and output images are paired together for training, and then the mapping is learned between them. Two steps are involved in the context restoration approach. Two small patches are randomly selected and swapped in the given input image. Repeating this process multiple times results in the formation of a new altered image. We can model this as a corruption function that changes the context of the original image and results in a new version of the image. Afterward, CNN is used to approximate an inverse function that restores the altered image to the original version. The proposed network consists of two main parts i.e. analysis part and the reconstruction part. The proposed network is shown in Figure 2.3.

According to Bai *et al.* [19], an efficient and effective pretext task for medical data is to predict the anatomical positions defined by some bounding boxes. The intuition behind this approach is that in this way, the network not only learns features related to particular parts but also starts to recognize their neighboring regions. Then these learned features can be passed to the actual task like performing segmentation of different anatomical structures. Solving Jigsaw Puzzles is also another valuable proxy task for self-supervised learning [152]. This approach has been extended to multi-modal medical data [152] with the intuition that by mixing data from various modalities, we can learn more informative features and representations. The authors [152] also make use of Pix2Pix GAN-based image-to-image translation network [86] to generate more data for the modality with less amount of data, by utilizing data from another larger modality.

2.2.6 Unsupervised learning methods

Unsupervised machine learning models mainly deal with the methods of utilizing unlabeled data for training. The main goal is to identify and learn hidden structures, similarities, and relationships in the provided data samples [235]. Different clusters are formed based on the distance or similarity measures,

hence samples that are alike are clustered together. Besides clustering algorithms, unsupervised learning algorithms can also use some kind of association rule which is used to find similarities and relationships between samples. Some common unsupervised algorithms include K-means clustering, K-nearest neighbors (KNN), hierarchical clustering, anomaly detections, Principal Component Analysis (PCA), singular value decomposition, and neural networks.

Unsupervised learning has been adopted in various domains in recent years and also in the past. However, the amount of work done in medical diagnosis via unsupervised algorithms are still very limited. Goswami *et al.* [63] utilized a neural network based on unsupervised learning for the detection of brain tumor using MRI images. ICA is used to extract the most useful features and then an unsupervised learning-based classification technique "Self-organizing map" (SOM) is used to classify the MRI brain images along with the K-means clustering algorithm.

Unsupervised domain adaptation [77][159] transfers knowledge learned on a source domain with a large number of labeled examples to a target domain with unlabeled data. This is very useful in medical diagnosis when there is limited annotated data. A novel Collaborative Unsupervised Domain Adaptation method (CoUDA) is proposed in [240]. Extensive experiments on real-world data show the generalization and effectiveness of the proposed network.

2.2.7   Semi-supervised learning methods

Semi-supervised learning [244] is another approach for machine learning which uses a combination of labeled and unlabeled i.e. it utilizes a small number of labeled examples with a large number of unlabeled data. Due to their success in various domains, semi-supervised algorithms have been used in the medical domain as well in recent years.

A semi-supervised learning-based algorithm for dimensionality reduction can be used to extract useful information from unlabeled data which helps in improving the classification accuracy on medical images [24]. Zhao *et al.* [242] proposed a compact graph-based semisupervised learning approach for the medical diagnosis specifically for Alzheimer's disease. Authors claim that the newly proposed model can represent the data manifold in a more compact form and also predict whether a patient is suffering from dementia or not. To their knowledge, this is the first work that applied a semi-supervised graph-based method for medical diagnosis and shows better accuracy and performance compared to other graph-based semi-supervised learning methods. An immune-inspired semisupervised algorithm has been presented for breast cancer diagnosis [163].

A semi-supervised learning-based GAN model is proposed for automated diagnosis of cardiac disease [137]. The loss function is comprised of three different loss functions namely supervised, unsupervised, and generated loss. Training the model involves one pass through the generator and three passes through the discriminator during each iteration. Results indicate that model can learn with a very limited amount of labeled data i.e. only 4data is kept unlabeled and achieves an accuracy of up to 80for classification.

2.2.8   Some miscellaneous methods

This section highlights some of the very recent works which are unique based on their method and lack much prior and follow-up work. One such method is presented by Chung *et al.* [37]. To learn the representations of medical data with less supervision and training data, the authors propose a Deep Siamese CNN architecture. Deep SCNN architecture is a variant of neural networks that can find correlations and similarities between the members or input data. Two identical CNN sharing the same

weights have been used for the construction of deep SCNN.

To prevent the overfitting of the classifier on small medical data and to improve generalization, a novel reinforced classifier is proposed [5]. Instead of trying to hard-fit its output to the training example, the proposed model tries to directly improve the generalization performance on the unseen data based on the policy gradient methods in reinforcement learning (RL) [16]. Deep learning (DL) approaches by utilizing chest X-ray has been actively explored for the classification of COVID-19 [223][150][14].

Inspired by these early works, the authors of [154] further explore the feasibility of deep learning methods for coronavirus detection. The intuition behind the presented approach is that the patch-wise intensity distribution has statistically significant differences in COVID-19 CXR (Chest X-ray). These findings led to the formation of a novel patch-based deep neural network architecture with patch cropping in a random manner. The presented approach reduces the network complexity and in each image, multiple patches can be used to augment the training data set. Hence, the neural network can be trained efficiently even with such small data without overfitting. The authors used a simple ResNet- 18 architecture for classification task to avoid overfitting, and it also allows to use of the pretrained ImageNet weights for the small training set and makes the overall training process stable.

2.3 Contrastive Representation Learning

Recently, contrastive learning has received great interest from many researchers due to its success in self-supervised representation learning. *Representation learning* involves the process of learning a parametric mapping between input data domain and feature vector. This can help in extracting more abstract features that can help in improving the performance of the downstream tasks [116]. Deep network architectures can achieve abstract and distributed representations to some extent. However, disentangled and invariant representations are not easy to learn. Contrastive learning provides a simple approach to learn and encode these representations and properties. Due to its success in natural language processing (NLP) and computer vision (CV) tasks, contrastive learning recently became popular. However, the original ideas originated in 1990s [25][26].

Several recent studies present promising results on unsupervised and self-supervised visual representation learning using approaches related to contrastive loss. Wu *et al.* [232] presented an unsupervised feature learning approach that focuses on an extreme case of instance discrimination, instead of class discrimination. To extract useful representation from high-dimensional data, a universal unsupervised learning approach, called Contrastive Predicting Coding (CPC) is presented by Oord *at al.* [216]. Powerful auto-regressive models are used to predict the future samples in the latent space. To capture useful information in the latent space, a predictive contrastive loss is utilized. Strong performance is achieved in text, images, speech, and reinforcement learning in 3D environments. To obtain locally consistent representations, a novel method based on unsupervised learning is presented [79] to maximize mutual information.

Zhuang *et al.* [245] presented a method that trains an embedding function based on a local aggression metric. Instances that are similar to each other are pushed apart in the embedding space, while similar instances are pulled closer. CPC also enables data-efficient image recognition [78]. Given small amounts of labeled data, CPC greatly improves the performance of classifiers and detectors. Powerful representations are learned by maximizing mutual information in an unsupervised manner based on multiple views of a data set [213].

All these methods are similar to building dynamic dictionaries. Given this idea, it is desirable

to build dictionaries that are large and stay consistent as they evolve [74]. To build consistent large dictionaries with unsupervised learning and contrastive loss, a Momentum Contrast (MoCo) approach is presented by He *et al.* [74]. To effectively utilize label information, Khosla *et al.* [102] extended the idea of self-supervised contrastive learning to fully supervised settings. Clusters of points belonging to the different classes are pushed apart in the embedding space. At the same time, clusters belonging to the same class are pulled closer. A comprehensive analysis is presented by Tian et al. [214] on the view selected for the multiple views contrastive representation learning. A simple framework (SimCLR) is presented for contrastive learning of visual representations [34]. Unlike previous approaches, the proposed method does not require a memory bank or a specialized architecture.

The application of contrastive representation learning to the medical domain is still very limited. However, some researchers are making attempts to utilize its effectiveness for medical data cases. To alleviate the need for large annotated training sets for segmentation of volumetric medical images with limited labeled data, Chaitanya et al. [28] utilized the contrastive learning framework in a semi-supervised setting. Extensive experimentation shows improved performance on MRI data sets in case of limited labeled data. Self-supervised-based contrastive learning is applied to discriminate physiological signals of the patients [105]. Spatial and temporal representations. that correspond to instances from the same patient are encouraged to be similar to each other by minimizing the euclidean distance. Medical visual representations are learned in an unsupervised manner from naturally occurring pairing of text and images [238]. A bidirectional objective is used to contrast text data with the learned image representations.

2.3.1 Limitations of past work

No focus on small data regime. Despite extensive research on applying machine learning methods to the medical domain and specifically for AD classification, there remain some limitations in past works. First and foremost, all the previous work for AD diagnosis utilized a huge amount of labeled data. The resulting performance of the model is good due to the big data set. However, there are many cases when the amount of labeled data is very small.

Transfer Learning. Another limitation is the use of using transfer learning like pre-training on ImageNet data. Such methods can restrict the designer to architectures that are needlessly complex for the medical data set, and can cause performance degradation during training and inference. Also, transfer learning does not work most of the time in the case of 3D medical analysis applications, since the 3D and 2D kernel shapes are not compatible. Besides, medical data is different from real-world data, and pre-training on real-world images like ImageNet is not a very useful idea for medical diagnosis. Hence, transfer learning has its limitations in the case of medical data.

Biased Evaluation. One more shortcoming of the previous works is no proper testing procedure. Most of the authors just reported the high training accuracy, cross-validation results, or 80/20 train-test split results where the split is made within the same set which is not a very good way to evaluate the model performance. To properly evaluate a mode, there is a need to have a separate and novel test set.

Difference in data and metrics. Also, it is not possible to make a direct comparison between various past works. These approaches differ in the participants involved, image processing procedures, cross-validation methods, and the evaluation metrics.

Data Selection. There has been some limitation in data selection methods as well. Some work utilized all the slices from a patient's data. Usually not all the slices are good for AD classification, and some bad slices can also cause performance degradation. Khan *et al.* [99] presented a way to select most

useful slices based on entropy calculation. However, during our research, we found that this method also has its limitations and is not a good way to select the slices.

Data Leakage. To assess the potential of classifiers, the evaluation must be completely unbiased, especially for the case of medical data. Data Leakage, which refers to the presence of test data in any part of the training process, is a major source of bias during the evaluation. Since DL approaches are flexible and complex, data leakage can be very hard to detect. Wen *et al.* [229] presented a detailed analysis of data leakage among various approaches. Following are the three main causes of data leakage.

- **Wrong data split.** Training, validation, and test set should be separated at the subject-level and not the data-level. If not, then data from the same patient can appear in several sets, resulting in biased evaluation of the model.

- **Late split.** Procedures like feature selection, data augmentation, or pre-training must never use test data. These steps should be performed on the training set after separating the data into training, validation, and test set.

- **Absence of independent test set.** To correctly evaluate the performance of the classifier, the test set should be separate and should only be used in the final stage to assess the classifier. It should not be used during model training or parameter tuning. A validation set should be prepared for hyper-parameter optimization. It is also suggested that for meaningful hyper-parameter selection, the validation set should also be separate i.e. instead of splitting training set into training and validation set, prepare a separate validation set on subject-level, just like the test set.

2.3.2 Contribution of this research

In this work, we present a novel approach to classify AD patients and to deal with medical data scarcity. The key contributions of this work are listed below.

- **A small-data approach.** Dealing with medical data scarcity specifically for the case of AD diagnosis. We use a small data set approach and achieve good results using only a few MRI slices.

- **Dealing with class imbalance.** Also removed imbalance among classes using data augmentation.

- **Novel medical image synthesis.** We present a novel PGGAN-based approach for medical image synthesis in combination with traditional data augmentation.

- **Applciation of SimCLR to AD classification.** We present application of SimCLR [34] and CBAM [231] to the case of AD classification for the first time.

- **Unbiased Evaluation.** To correctly evaluate our model without data leakage, we prepared separate training, validation, and test sets on subject-level. This results in effective hyper-parameter optimization and evaluation of our approach.

- **An analysis of various approaches.** We also present an analysis of some previous approaches and different architectures. Results indicate the presence of data leakage and overfitting in these different approaches.

This research aims to design an automated model with improved performance for AD diagnosis specifically for the small data set settings. Although there have been some approaches to deal with

medical data scarcity. However, none of them was focused on the AD classification task, specifically using GAN-based image synthesis. Contrastive representation learning is still novel in the medical sector. There hasn't been any work based on contrastive learning for the classification of Alzheimer's Disease (AD).

Chapter 3.   Proposed Method

3.1   Data Preparation

The first setup was to select and prepare the data set. We collected T2-weighted MRI scans from the ADNI (Alzheimer's Disease Neuroimaging Initiative) website. Data for 246 patients (82 AD, 82 MCI, and 82 CN) is used for training. Additional data from 54 patients were kept separate for testing purposes. This test data is further divided into two parts i.e. validation and test data. To clear evaluate our model, validation data is selected from test data instead of training data. Validation data consisted of 115 slices (51 for AD, 64 for CN) and test data consisted of slices from 10 patients for each category. Demographic information of the training data is shown in Table 3.1.

In our case, we only used axial view slices and the dimensions of each slice are 256x256. Initially, each patient had around 48 slices. Manual data cleaning was performed and only good slices with visible brain regions were selected. The slices with no or less information were removed from the data.

3.2   Data Augmentation

In this research, we presented a small-data approach. One common problem faced when training with limited data is the overfitting. To deal with this, we use a novel GAN architecture for image synthesis. However, training a GAN also requires a large amount of data, which brings us to traditional data augmentation techniques. We performed traditional data augmentation on our data including small rotations, shear, zoom, width, and height shift. The resulting number of slices is 2435, 2429, and 2389 for AD, CN, and MCI respectively.

3.3   GAN-based image generation

After the data augmentation, we generated whole-brain MRI images using GAN. There are various GAN architectures available. We decided to implement three architectures i.e. DCGAN [168], RaLSGAN [92] and PGGAN [97]. DCGAN is hard-coded for 64x64 images. Since we wanted to keep our image size 256x256, we tried to train DCGAN on 256x256 images. However, the training of the network is not stable for higher resolutions.

A comparison was made between RaLSGAN and PGGAN for medical image synthesis based on the FID (Fréchet Inception Distance) score. The FID score is a measure based on the similarity between real and generated images and is commonly used to evaluate GAN models. The lower the FID score, the better the model. In our case, PGGAN outperformed RaLSGAN for image synthesis. Moreover, Progressive Growing of GANs (PGGANs) with the Wasserstein loss [69] using gradient penalty generated very realistic and high-resolution images. The FID score for PGGAN-WP for our case is around 40 while for RaLSGAN it is around 140. A comparison of generated images by the two models is shown in Figure 3.1. The final training data set consisted of 4870 slices for each category. Table 3.2 shows the number of training samples after PGGAN-based image generation.

Table 3.1: Demographic representation of training MRI images

Data source	Research group	Total subjects	Sex		Age (years)	Number of MRI volumes	Image slices	Imaging protocol
			M	F				
ADNI	AD	82	46	44	55-93	82	406	Axial,
	CN	82	46	46	70-96	82	347	2D,
	MCI	82	46	46	58-96	82	279	1.5-Tesla

Table 3.2: Total training data after PGGAN-based image generation

Data source	Research group	MRI volumes		
		Before Augmentation	After Augmentation	After PGGAN
ADNI	AD	406	2435	4870
	CN	347	2429	4870
	MCI	279	2389	4870

3.4 Alzheimer's Disease Classification

3.4.1 Comparison of various architectures

As a first step, a detailed analysis was made on the AD vs. CN classification accuracy using different architectures. These architectures include Custom CNN architecture, VGG16 [195], inception model [205], residual attention network [221], CBAM [231] architectures and Multi-scale CNN. We also evaluate pretrained models on ImageNet data. Since the validation data was unseen and novel, all the models were evaluated based on the validation accuracy. Table 3.2 shows a comparison of various architectures. Testing results are reported based on the majority voting decision for the patients.

All these architectures showed higher accuracy during training. However, the validation accuracy is not very high. Models were also used to predict the patients using the test data. The testing accuracy of all the models was nearly the same as the validation accuracy. Majority voting was used to make and categorize the patients. If a patient has more slices with AD prediction, then the patient was classified as AD. Otherwise, the patient belongs to CN. The best performance was achieved in some cases with 2-3 misclassifications per category.

Based on these results, we can conclude that these models are not very promising for small medical data classification. As mentioned earlier, transfer learning from real-world data is also not very useful in the case of medical data. Most of the models seem to overfit due to the small size of training data. Raghu *et al.* evaluated transfer learning for small medical data regime. It has been reported that for medical image tasks, transfer learning does not significantly affect the model performance [169]. Models trained from scratch showed comparable performance to the standard ImageNet architectures. The reason can be that these large ImageNet models might be over-parametrized for very small data sets. Therefore, there is a need for a better model and approach that can achieve robust performance for small medical data.

Alzheimer's disease (AD)
with PGGAN

Cognitive Normal (CN)
with PGGAN

Alzheimer's disease (AD)
with RaLsGAN

Cognitive Normal (CN)
with RaLsGAN
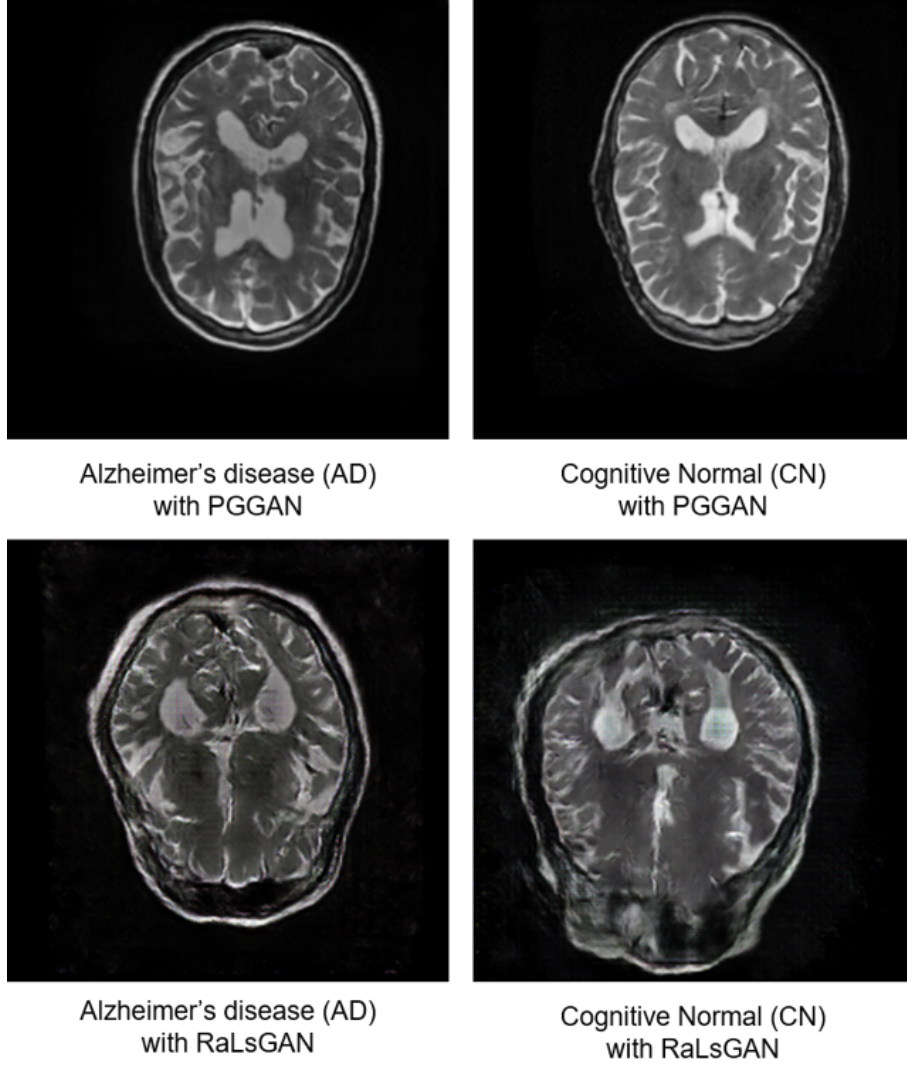
Figure 3.1: Comparison of MRI images generated by PGGAN-WP and RaLSGAN

3.4.2 Proposed Architecture

In this research, we propose a novel method for AD classification in small data regime. Proposed methodology consists of three main parts.

- PGGAN-based image generation

- Model pre-training with SimCLR

- Training the pre-trainined model for classification task

We used ResNet-18 [75] with CBAM [231] as our base model. After the PGGAN-based image generation, the base model is first pretrained via SimCLR [34] to learn discriminative representations. The new train this final model again for the classification task. Details and working of the above architectures are mentioned in the next sections.

Table 3.3: Comparison of various architectures for AD vs. CN classification

Accuracy	Custom CNN	ResNet-18	Inception pretrained	Residual Attention Network	CBAM Inception V3	Multi-scale CNN	DenseNet
Train (%)	98.04%	99%	99.8%	85%	99%	100%	97.4%
Validation(%)	72.91%	75%	70%	76.67%	70.8%	70.83%	72.32%
Testing (#)	15/20	16/20	13/20	12/20	15/20	12/20	15/20

Figure 3.2: Backpropagation in generator training [42].

PGGAN

The Progressive-Growing GAN architecture [97] release from NVIDIA has shown impressive performance in GAN-based image synthesis. Classical GAN architectures struggle to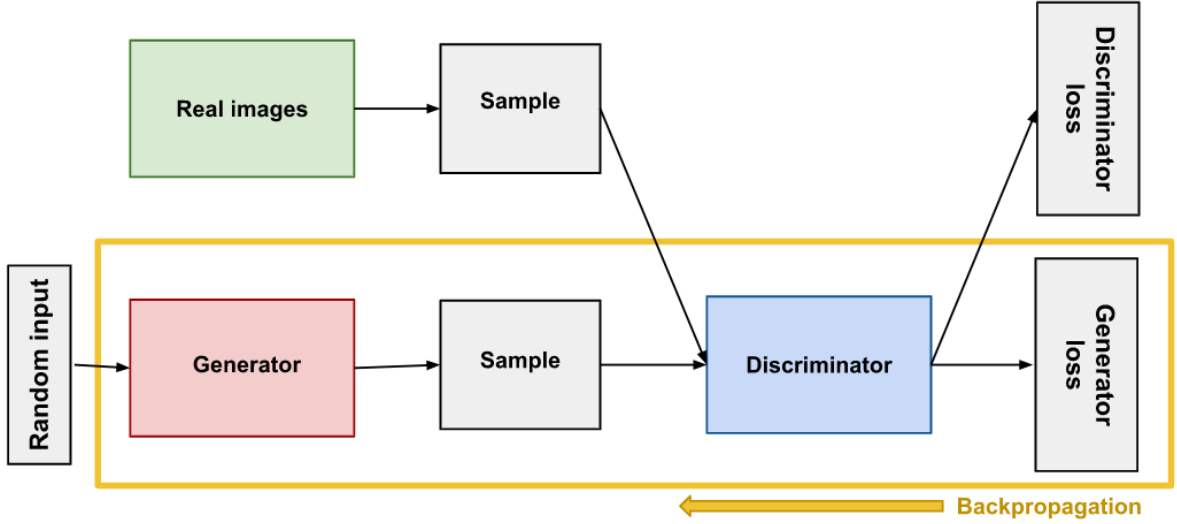 produce good images even for low resolution like 32x32 or 128x128. However, PGGAN can output good quality high-resolution images of up to 1024x1024.

Typically, a GAN consists of two parts i.e. a generator and a discriminator (aka critic). The generator takes random noise as input and produces an image sample. Ideally, this image sample should be indistinguishable from the training distribution. The discriminator network acts as a critic that assesses the generated images. The goal of the generator is to fool the discriminator network by generating as real images as possible, while the discriminator penalizes the generator if it fails to do so. Loss is calculated for the discriminator classification, which is then backpropagated through the network for gradient calculation. These gradients are then used to update the generator's parameters. A typical example of generator training is shown in Figure 3.2.

In PGGAN, the main difference is in the training methodology. It utilizes a multi-scale architecture where the training starts at low-resolution and then the resolution is increased by adding new layers to the network in a progressive manner as shown in Figure. The original data set with real training sampled is first classified into different resolutions like 4^2, 8^2 up to 1024^2. The generator first produces 4^2 images
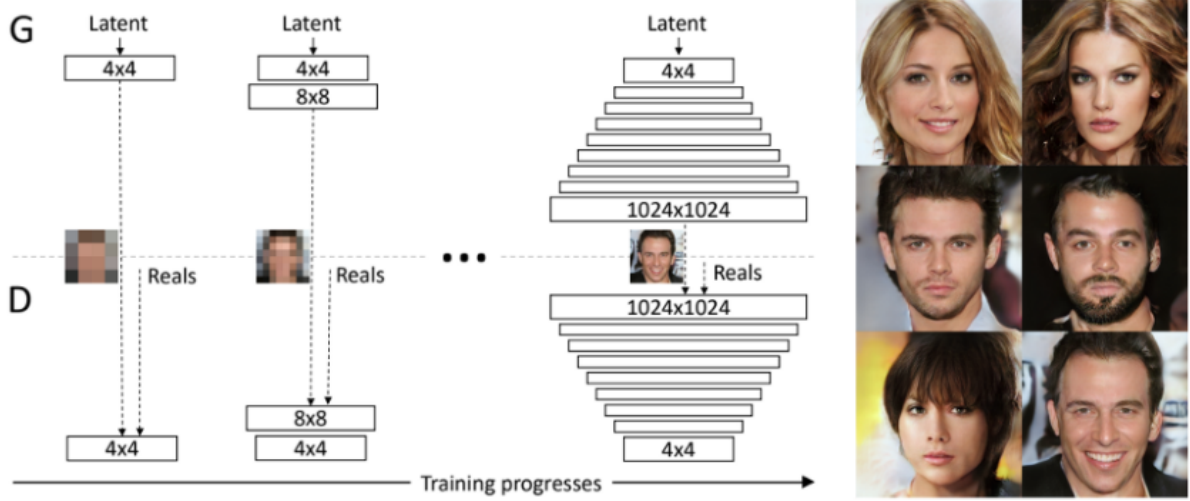
Figure 3.3: Diagram of the Multi-Scale Architecture used in Progressively-Growing GANs as proposed in the original paper [97].

until it reaches some equilibrium point. Then the training proceeds with 8^2 images up to 1024^2. This strategy improves the training stability. The main reason is the when we go from a latent space z to high-resolution like 1024^2, there is an enormous amount of variance present in the space. As has been the trend in previous GAN research, generating 128^2 RGB ImageNet images is much harder than low-resolution images such as 28^2 MNIST images. Figure 3.3 shows the proposed multi-scale architecture.

This study adopts the PGGAN architecture to generate 256x256 brain MRI images using an ADNI data set. In our case, we limited the resolution up to 256^2. We train and generate AD, CN, and MCI images separately. The details of the model used are shown in Figure 3.4.

Convolutional Block Attention Module

Attention modules are used in computer vision to make the model learn and focus more on the important information, rather than learning background information. The idea of attention was first introduced in the domain of Natural Language Processing (NLP), introduced by Google Brain [218] in NeurIPS 2017. Recently, there have been many works that adopted the idea of the attention to computer vision tasks.

A typical attention module generates a mask of the input feature map using a simple 2D-convolutional layer, multi-layer perceptron (MLP), and a sigmoid function at the end. Convolutional Block Attention Module or CBAM [231] is a recently proposed attention module for CNNs. Provided an input feature map, it computes the attention maps along two dimensions i.e. channel and spatial. These inferred attention maps are then multiplied with the input feature map to further refine the features. The intuition behind this idea is that blind attachment of an attention module can result in a 3D attention map which can be computationally expensive. This work [231] proposes the idea to compute channel attention and spatial attention separately using two sequential sub-modules called the Channel Attention Module (CAM) and the Spatial Attention Module (SAM). Results indicate that the proposed method achieves a similar effect with much fewer parameters. Convolutional Block Attention Module has shown consistent improvements in classification and detection tasks over the previous state-of-the-art models. The general structure of the CBAM module is shown in Figure 3.5.

Generator	Activation	Output Shape		
Latent vector	–	512 ×	1 ×	1
Conv 4 × 4	LReLU	512 ×	4 ×	4
Conv 3 × 3	LReLU	512 ×	4 ×	4
Upsample	–	512 ×	8 ×	8
Conv 3 × 3	LReLU	512 ×	8 ×	8
Conv 3 × 3	LReLU	512 ×	8 ×	8
Upsample	–	512 ×	16 ×	16
Conv 3 × 3	LReLU	256 ×	16 ×	16
Conv 3 × 3	LReLU	256 ×	16 ×	16
Upsample	–	256 ×	32 ×	32
Conv 3 × 3	LReLU	128 ×	32 ×	32
Conv 3 × 3	LReLU	128 ×	32 ×	32
Upsample	–	128 ×	64 ×	64
Conv 3 × 3	LReLU	64 ×	64 ×	64
Conv 3 × 3	LReLU	64 ×	64 ×	64
Upsample	–	64 ×	128 ×	128
Conv 3 × 3	LReLU	32 ×	128 ×	128
Conv 3 × 3	LReLU	32 ×	128 ×	128
Upsample	–	32 ×	256 ×	256
Conv 3 × 3	LReLU	16 ×	256 ×	256
Conv 3 × 3	LReLU	16 ×	256 ×	256
Conv 1 × 1	Linear	1 ×	256 ×	256

Discriminator	Activation	Output Shape		
Input image	–	1 ×	256 ×	256
Conv 1 × 1	LReLU	16 ×	256 ×	256
Conv 3 × 3	LReLU	16 ×	256 ×	256
Conv 3 × 3	LReLU	32 ×	256 ×	256
Downsample	–	32 ×	128 ×	128
Conv 3 × 3	LReLU	32 ×	128 ×	128
Conv 3 × 3	LReLU	64 ×	128 ×	128
Downsample	–	64 ×	64 ×	64
Conv 3 × 3	LReLU	64 ×	64 ×	64
Conv 3 × 3	LReLU	128 ×	64 ×	64
Downsample	–	128 ×	32 ×	32
Conv 3 × 3	LReLU	128 ×	32 ×	32
Conv 3 × 3	LReLU	256 ×	32 ×	32
Downsample	–	256 ×	16 ×	16
Conv 3 × 3	LReLU	256 ×	16 ×	16
Conv 3 × 3	LReLU	512 ×	16 ×	16
Downsample	–	512 ×	8 ×	8
Conv 3 × 3	LReLU	512 ×	8 ×	8
Conv 3 × 3	LReLU	512 ×	8 ×	8
Downsample	–	512 ×	4 ×	4
Minibatch stddev	–	513 ×	4 ×	4
Conv 3 × 3	LReLU	512 ×	4 ×	4
Conv 4 × 4	LReLU	512 ×	1 ×	1
Fully-connected	Linear	1 ×	1 ×	1

Figure 3.4: PGGAN architecture details for generator/discriminator used for the MRI synthesis [97].

Two pooling methods i.e. average and max pooling, are used at the same time to compute channel-wise attention for the given input feature map. This results in two Cx1x1 vectors, one produced by max-pooling and the other by average pooling. These are then passed through a simple bottleneck dense layer and then combined with a summation. The sigmoid function is applied at the end to obtain a Cx1x1 vector which shows the importance of each channel in the original feature map. Figure 3.6 shows the channel attention sub-module. This channel attention vector is applied to the input feature in a pointwise manner, which creates a new vector $\mathbf{F}'$ which is shaped the same as the original input feature map $\mathbf{F}$.

After channel dimension, the next step is to process features in width and height dimensions. From the channel attention module, we obtain a CxHxW map. In the spatial attention module, average and max pooling are applied pointwise which results in two 1xHxW features. A 7x7 convolution is applied after concatenating the two features. In the end, the sigmoid function is applied to get a 1xHxW shaped feature, which is called a spatial attention map. This spatial attention map is applied to $\mathbf{F}'$ pointwise, resulting in a CxHxW vector and we get the final output of CBAM. Figure 3.7 shows the channel attention sub-module.

SimCLR

Numerous self-supervised methods have been presented in recent years, each better than the previous ones. However, their performance was still lower than the supervised learning methods. Cheng *et al.* changed this by presenting their research paper "SimCLR: a simple framework for contrastive learning of visual representations" [34]. The SimCLR paper not only improves upon the previous state-of-the-art self-supervised learning methods but also shows better performance for ImageNet classification than the supervised learning methods when scaling up the architecture.
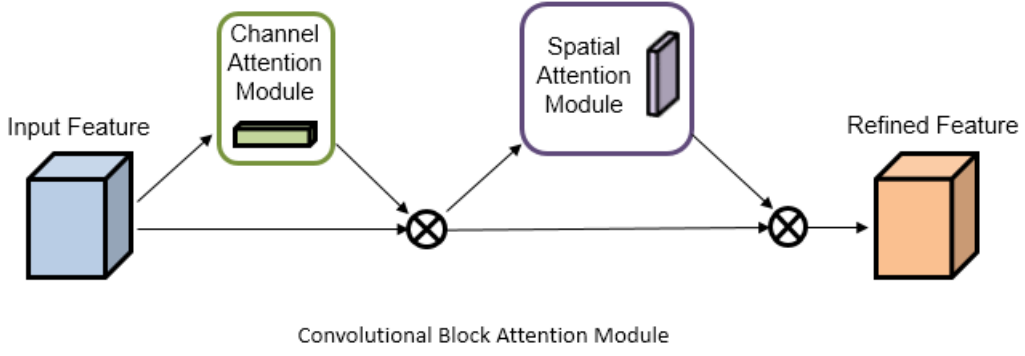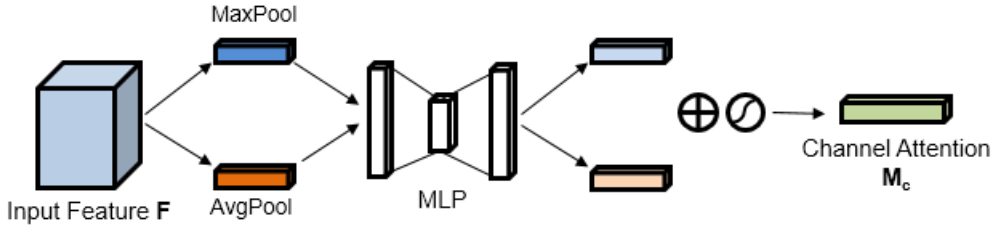
Figure 3.5: The overview of CBAM [231].

Figure 3.6: Channel attention sub-module [231].

The basic intuition behind contrastive learning is to teach a machine how to distinguish between similar and dissimilar things. SimCLR blends contrastive learning with some novel ideas to learn the visual representation without human supervision. The idea of SimCLR framework is rather very simple. Given an input image, random transformations are applied to get two augmented versions of the image x_i and x_j. Representation is then obtained for these augmented images by passing them through an encoder network. These encoded vector are represented as h_i and h_j. Then a non-linear fully connected layer is applied to get representations z. The objective is to maximize the similarity between these two representations z_i and z_j for augmented versions of the same image. Figure 3.8 shows working of SimCLR framework.

In their research [34], the authors used a combination of random transformations including crop, flip, color jitter, and grayscale. A random transformation function (T) is applied to get a pair of two images for each input image in the batch. Hence for the two-class case and batch size of 2 (one from each class), we have 4 images after augmentation. To get the encodings h_i and h_j, the authors used ResNet-50 [76] architecture. These encoded representations are then passed through a series of **Dense -> ReLU -> Dense** non-linear transformation layers to obtain representations z_i and z_j. In the paper, the non-linear projection part is referred to as the projection head and is denoted by $g(.)$. Cosine similarity shown in equation 3.1, is calculated between the representations z_i and z_j. The similarity of the augmented images belonging to the same class will be higher compared to the similarity between images from different classes.

$$s_{i,j} = \frac{z_i^T z_j}{||z_i||||z_j||} \tag{3.1}$$

SimCLR uses contrastive loss called NT-Xent loss. The augmented pair in the batch are taken one by one and the probability of the two images being similar is calculated by applying the softmax
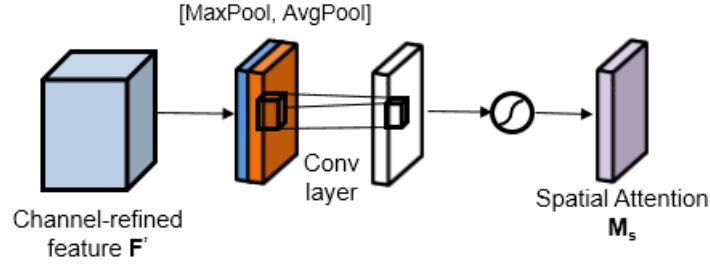
Figure 3.7: Spatial attention sub-module [231].
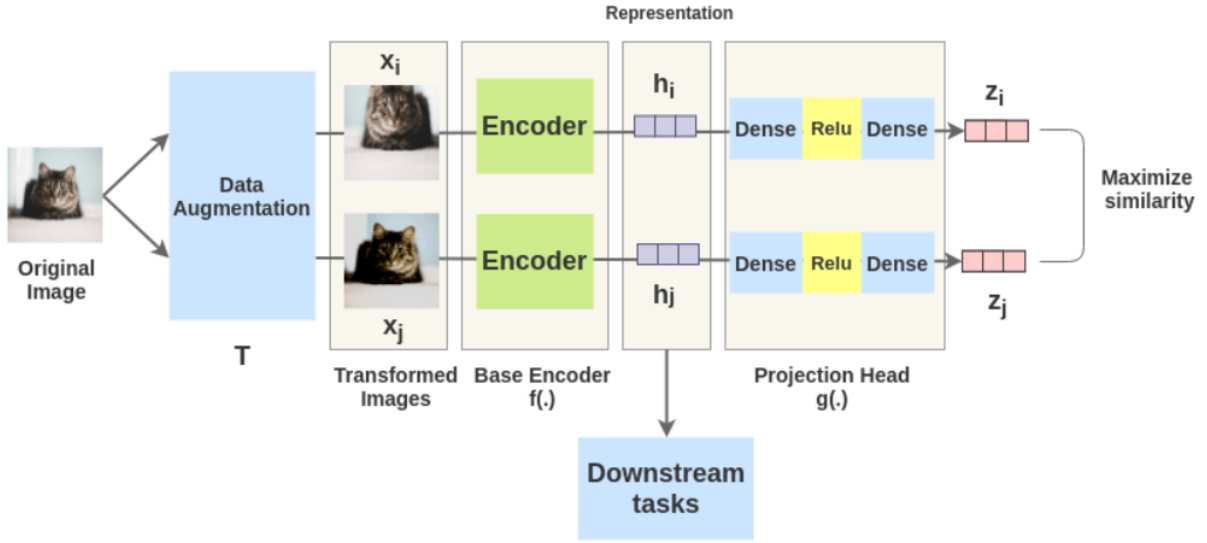
Figure 3.8: Working of SimCLR framework [34].

function. The loss is calculated by taking the negative log of the above calculation (equation 3.2). Loss is computed for the same pair a second time, by interchanging the positions of the images in the pair. Finally, we compute loss over all the pairs in the batch of size N=2 and take an average.

$$L = \frac{1}{2N} \sum_{k=1}^{2N} [l(2k-1, 2k) + l(2k, 2k-1)] \qquad (3.2)$$

The encoder and projected representation are improved over time based on the above loss values placing similar images closer to each other and dissimilar images apart. Once the training is finished on the contrastive learning task, this pretrained model can be used in a transfer learning fashion to perform a downstream task. For this, the representations from the encoder are used instead of representations obtained from the projection head.

In our case, we used random cropped versions of the input images as transformations, and the downstream task is the AD vs. CN vs. MCI classification. The ResNet-18 model with CBAM is pretrained using the SimCLR framework and is then applied to the classification task.

Chapter 4. Experiments

We then performed experiments using our approach. Training and testing are done for two-category cases (AD vs. CN) and three-category cases (AD vs. CN vs. MCI). Our training details and results are mentioned in the next sections.

4.1 Training Details

4.1.1 System and environment

We used PyTorch modules to design and train our proposed model. All the training is performed using NVIDIA RTX 2080 Ti GPU.

4.1.2 Data Preprocessing

The original images obtained from the ADNI repository are 256x256 DICOM images. These images are converted into PNG format using a python library named pydicom [166]. The resulting images are then center-cropped to 224x224 and image pixels are normalized between 0 and 1. As mentioned earlier, images with no or little information are removed manually and only informative slices are used.

4.1.3 SimCLR

Next we pretrain our model using SimCLR [34] framework. Model is trained for 1000 epochs with a batch size of 128. SGD optimizer is used with a learning rate of 0.05, momentum value of 0.9, and weight decay regularization of 0.0004. We used 80%:20% split to divide our data into training and validation sets.

4.1.4 Classification task

Next, we load our pre-trained model for the classification task. Model is trained for 100 epochs with a batch size of 128. SGD optimizer is used with a learning rate of 0.001 which is decayed after a certain number of epochs. Momentum value of 0.9 and weight decay regularization of 0.0004. As mentioned earlier, separate data sets are prepared and used for training, validation, and testing the model.

4.2 Results

We trained our ResNet-18 [76] model for comparison. We also compare results with and without CBAM and SimCLR architectures to highlight their effect. We also compared our model to a custom-designed CNN (five convolutional layers and five fully connected layers). Classification results are shown in the table. Here we only mention validation and test accuracies. Training accuracy can easily be very high and does not convey much meaningful information in terms of the actual model performance. Table 4.1 and 4.2 show results of the proposed framework for AD vs. CN and AD vs. CN vs. MCI classification task respectively.

Table 4.1: AD vs. CN classification results of proposed framework

Accuracy	Custom CNN	ResNet-18 (scratch)	ResNet-18 (SimCLR)	ResNet-18 + CBAM (SimCLR)
Validation	75%	70%	75%	81%
Test(per slices)	77%	69%	75%	83%
Test(per patient)	9/10, 7/10	6/10, 8/10	8/10, 8/10	9/10, 9/10

Table 4.2: AD vs. CN vs. MCI classification results of proposed framework

Accuracy	ResNet-18 + CBAM (SimCLR)
Validation	66%
Test(per slices)	65%
Test(per patient)	7/10, 7/10, 7/10

As seen in the results shown in the table 4.1 and 4.2, ResNet-18 architecture combined with CBAM and SimCLR provides better results. When training from scratch, the model seems to overfit and the training process is unstable. However, after pretraining, the model training on classification tasks is more stable and the resulting accuracy also goes up to 75%. The model performance is further refined when combined with CBAM architecture since the model learns to focus and refine important features for the classification task. This shows that both SimCLR and CBAM help the model in achieving considerable performance gains on medical data.

In further experiments, the model was trained for the three-way classification task. However, the model validation and testing accuracies ranged between 55-60%. Here model shows lower performance. The main reason is that MCI class is very hard to distinguish from AD and CN classes. Also, the amount of data is small which can result in poor training of the model. We also investigated the effect of PGGAN-based medical image generation. The model shows 3-4 misclassifications during the testing phase when trained on data with only traditional data augmentation. Adding PGGAN data increases the amount and also a variety of data which can help the model learn more generalized features.

4.2.1 Comparison with existing methods

Finally, we compare our results with some other deep learning-based approaches in recent years for the AD classification task. However, the comparison is not completely fair since most of these past approaches use large data sets or suffer from data leakage and also lack proper evaluation methods.

Aderghal *et al.* achieved an accuracy of 84% percent for the AD vs. CN classification task using the ROI-based approach. Accuracy of 85% is achieved using a 3D subject-level approach [36]. Cheng *et al.* reported an accuracy of 87% for AD vs. CN using a 3D patch-level approach. Korolev *et al.* [109] and Li *et al.* [123] obtained accuracies of 80% and 88% using 3D subject-level approaches. Valliani and Soni [] reported accuracy of 81% for AD vs. CN classification and 57% for three-way classification using 2D slices. Senanayake *et al.* achieved 76% accuracy using 3D subject-level approach for AD vs. CN classification. Compared with our approach, some of these methods have higher accuracy. These

methods are free from data leakage. However, the amount of training data was large.

An accuracy of 91% is achieved using ROI-based approach by Aderghal *et al.* [3]. However, their approach suffers from late-split and the absence of an independent test set. Similar limitations are observed in [22][85][212][224] where reported accuracies are very high but there was no proper test set and evaluation was biased. Liu *et al.* [127] achieved 85% accuracy for AD vs. CN task using 3D patches. However, their approach suffers from biased transfer learning (see Section 2.3.1). Farooq *et al.* achieved a very high accuracy of 99% for AD vs. CN vs. MCI classification using 2D slices. However, their data split method was wrong and resulted in a biased evaluation of the model. Vu *et al.* reported 86% and 80% accuracy for AD vs. CN and AD vs. CN vs. MCI classification respectively using 3D data. However, this approach also suffers from a wrong data split. Accuracy of 95% is reported by Wu *et al.* for three-way classification using 2D slices but there was no independent test set or proper data split. Approaches by Hosseini *et al.* [17] and Wang *et al.* [222] also suffer from these limitations. Our model also shows an accuracy of up to 65% for AD vs. CN vs. MCI classification, which is higher than the previous approaches without data leakage, despite the small data set.

In our approach, we avoided all kinds of data leakage and provided an unbiased evaluation of our model which is very important for clinical applications. Our model achieves comparable results to the past approaches even with a very small data set. Table 4.2 and 4.3 show comparison of proposed framework with previous approaches.

Table 4.3: Comparison of AD vs. CN classification

Study	Total Subjects	Performance	Approach	Data Leakage
Aderghal et al., 2017	815 (T1 MRI)	ACC=0.84	ROI-based	None
Cheng and Liu, 2017	193 (T1 MRI + PET)	ACC=0.85	3D subject-level	None
Korolev et al., 2017	231 (T1 MRI)	ACC=0.80	3D subject-level	None
Valliani and Soni, 2017	417 (T1 MRI)	ACC=0.81	2D slice-level	None
Senanayake et al., 2018	515 (T1 MRI)	ACC=0.76	3D subject-level	None
Li et al., 2017	427 (T1 MRI)	ACC=0.88	3D patch-level	None
Basaia et al., 2019	646 (T1 MRI)	ACC=0.99	3D subject-level	Unclear (b)
Hon and Khan, 2017	416 (T1 MRI)	ACC=0.96	2D slice-level	Unclear (a, c)
Hosseini et al., 2018	120 (T1 MRI)	ACC=0.99	3D subject-level	Unclear (a)
Lin et al., 2018	417 (T1 MRI)	ACC=0.89	ROI-based	Unclear (b)
Taqi eet al., 2018	400 (T2 MRI)	ACC=1.00	2D slice-level	Unclear (b)
Vu et al., 2017	317 (T1 MRI)	ACC=0.85	3D subject-level	Unclear (a)
Vu et al., 2018	400 (T1 MRI)	ACC=0.86	3D subject-level	Clear (a, c)
Wang et al., 2019	400 (T1 MRI)	ACC=0.99	3D subject-level	Clear (b)
Basheera et al., 2019	242 (T1 MRI)	ACC=1.00	2D slice-level	Clear (b)
Proposed	**164 (T2 MRI)**	**ACC=0.83**	**2D slice-level**	**None**

Table 4.4: Comparison of AD vs. CN vs. MCI classification

Study	Total Subjects	Performance	Approach	Data Leakage
Valliani and Soni, 2017	660 (T1 MRI)	ACC=0.57	2D slice-level	None
Hosseini et al., 2018	210 (T1 MRI)	ACC=0.95	3D subject-level	Unclear (a)
Farooq et al., 2017	355 (T1 MRI)	ACC=0.99	2D slice-level	Clear (a, c)
Vu et al., 2018	615 (T1 MRI)	ACC=0.80	3D subject-level	Clear (a, c)
Wang et al., 2019	624 (T1 MRI)	ACC=0.97	3D subject-level	Clear (b)
Basheera et al., 2019	349 (T2 MRI)	ACC=0.86	2D slice-level	Clear (b)
Proposed	**246 (T2 MRI)**	**ACC=0.65**	**2D slice-level**	**None**

*Types of data leakage: a: wrong data set split; b: absence of independent test set; c: late split

Chapter 5.   Concluding Remarks

We have proposed a novel framework for the AD classification task. The proposed architecture comprises a baseline ResNet-18 model in combination with CBAM which helps the model to refine the important discriminative features and put less focus on the background features. The whole model is pretrained using SimCLR [34] that enables the model to learn useful representations in a self-supervised manner. To deal with the issue of data scarcity, we presented a novel PGGAN-WGP based medical image synthesis scheme which generates very realistic MRI images.

We compare our results with the previous approaches. We evaluated our model on separate validation and test data without any kind of data leakage as seen in many previous approaches [229]. Our proposed model achieves an accuracy of 83% on the test data set with one misclassification per category for the task of AD vs. CN classification. These results are comparable to previous approaches despite the small data set. In the case of three-way classification, model achieves an accuracy of up to 60% on the test data. The proposed model shows promising results and can help in early diagnosis which can result in timely treatment of AD patients.

5.1   Future direction

There is still room for improvement in the model performance, especially for the case of AD vs. CN vs. MCI classification. As part of future work, we will test our model with more and more large data sets and compare the results. This can result in better training and make the model more generalizable. We also plan to look into better data preprocessing methods like segmentation, background removal, registration and better slice selection approaches. Improving the data can help the model performance. Taking help from medical experts can help in this process. Different pretraining tasks can also be used as pretext task instead of patch-based approach, and results can be compared.

Another future direction is to investigate the performance of 3D CNN architecture instead of 2D CNN with our approach. This framework can also be extended for the diagnosis of other diseases. We can also test this framework for other modalities or multiple modalities and compare the results. Another possible application can be explainable AI in which we can interpret how our model is making decisions during the evaluation process. Therefore, there are a lot of possible paths for future work and further work in this direction can help with the early diagnosis of various diseases, especially in small data regimes.

Bibliography

[1] K. Adamczuk, J. Schaeverbeke, H. M. Vanderstichele, J. Lilja, N. Nelissen, K. Van Laere, P. Dupont, K. Hilven, K. Poesen, and R. Vandenberghe. Diagnostic value of cerebrospinal fluid aβ ratios in preclinical alzheimer's disease. *Alzheimer's research & therapy*, 7(1):1–11, 2015.

[2] S. Adaszewski, J. Dukart, F. Kherif, R. Frackowiak, B. Draganski, A. D. N. Initiative, et al. How early can we predict alzheimer's disease using computational anatomy? *Neurobiology of aging*, 34(12):2815–2826, 2013.

[3] K. Aderghal, J. Benois-Pineau, K. Afdel, and C. Gwenaëlle. Fuseme: Classification of smri images by fusion of deep cnns in 2d+ ε projections. In *Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing*, pages 1–7, 2017.

[4] S. Afzal, M. Maqsood, F. Nazir, U. Khan, F. Aadil, K. M. Awan, I. Mehmood, and O.-Y. Song. A data augmentation-based framework to handle class imbalance problem for alzheimer's stage detection. *IEEE Access*, 7:115528–115539, 2019.

[5] W. A. Al and I. D. Yun. Reinforcing medical image classifier to improve generalization on small datasets. *arXiv preprint arXiv:1909.05630*, 2019.

[6] B. Al-Naami, N. Gharaibeh, and A. A. Kheshman. Automated detection of alzheimer disease using region growing technique and artificial neural network. *World Acad. Sci. Eng. Technol. Int. J. Biomed. Biol. Eng*, 7(5), 2013.

[7] S. Alam, G.-R. Kwon, and A. D. N. Initiative. Alzheimer disease classification using kpca, lda, and multi-kernel learning svm. *International Journal of Imaging Systems and Technology*, 27(2):133–143, 2017.

[8] M. S. Albert, S. T. DeKosky, D. Dickson, B. Dubois, H. H. Feldman, N. C. Fox, A. Gamst, D. M. Holtzman, W. J. Jagust, R. C. Petersen, et al. The diagnosis of mild cognitive impairment due to alzheimer's disease: recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. *Alzheimer's & dementia*, 7(3):270–279, 2011.

[9] V. Alex, K. Vaidhya, S. Thirunavukkarasu, C. Kesavadas, and G. Krishnamurthi. Semisupervised learning using denoising autoencoders for brain lesion detection and segmentation. *Journal of Medical Imaging*, 4(4):041311, 2017.

[10] H. Altae-Tran, B. Ramsundar, A. S. Pappu, and V. Pande. Low data drug discovery with one-shot learning. *ACS central science*, 3(4):283–293, 2017.

[11] T. Altaf, S. M. Anwar, N. Gul, M. N. Majeed, and M. Majid. Multi-class alzheimer's disease classification using image and clinical features. *Biomedical Signal Processing and Control*, 43:64–74, 2018.

[12] I. Álvarez, M. López, J. M. Górriz, J. Ramírez, D. Salas-Gonzalez, C. G. Puntonet, and F. Segovia. Automatic classification system for the diagnosis of alzheimer disease using component-based

svm aggregations. In *International Conference on Neural Information Processing*, pages 402–409. Springer, 2008.

[13] D. Amodei, S. Ananthanarayanan, R. Anubhai, J. Bai, E. Battenberg, C. Case, J. Casper, B. Catanzaro, Q. Cheng, G. Chen, et al. Deep speech 2: End-to-end speech recognition in english and mandarin. In *International conference on machine learning*, pages 173–182. PMLR, 2016.

[14] I. D. Apostolopoulos and T. A. Mpesiana. Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Physical and Engineering Sciences in Medicine*, 43(2):635–640, 2020.

[15] M. Arjovsky, S. Chintala, and L. Bottou. Wasserstein gan, 2017.

[16] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath. Deep reinforcement learning: A brief survey. *IEEE Signal Processing Magazine*, 34(6):26–38, 2017.

[17] E. H. Asl, M. Ghazal, A. Mahmoud, A. Aslantas, A. Shalaby, M. Casanova, G. Barnes, G. Gimel'farb, R. Keynton, and A. El Baz. Alzheimer's disease diagnostics by a 3d deeply supervised adaptable convolutional network. *FRONTIERS IN BIOSCIENCE-LANDMARK*, 23:584–596, 2018.

[18] A. Association. 2015 alzheimer's disease facts and figures. *Alzheimer's & Dementia*, 11(3):332–384, 2015.

[19] W. Bai, C. Chen, G. Tarroni, J. Duan, F. Guitton, S. E. Petersen, Y. Guo, P. M. Matthews, and D. Rueckert. Self-supervised learning for cardiac mr image segmentation by anatomical position prediction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 541–549. Springer, 2019.

[20] Y. Balaji, S. Sankaranarayanan, and R. Chellappa. Metareg: Towards domain generalization using meta-regularization. *Advances in Neural Information Processing Systems*, 31:998–1008, 2018.

[21] P. Baldi. Autoencoders, unsupervised learning, and deep architectures. In *Proceedings of ICML workshop on unsupervised and transfer learning*, pages 37–49. JMLR Workshop and Conference Proceedings, 2012.

[22] S. Basaia, F. Agosta, L. Wagner, E. Canu, G. Magnani, R. Santangelo, M. Filippi, A. D. N. Initiative, et al. Automated classification of alzheimer's disease and mild cognitive impairment using a single mri and deep neural networks. *NeuroImage: Clinical*, 21:101645, 2019.

[23] S. Basheera and M. S. S. Ram. Convolution neural network–based alzheimer's disease classification using hybrid enhanced independent component analysis based segmented gray matter of t2 weighted magnetic resonance imaging with clinical valuation. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 5:974–986, 2019.

[24] K. N. Batmanghelich, H. Y. Dong, K. M. Pohl, B. Taskar, C. Davatzikos, et al. Disease classification and prediction via semi-supervised dimensionality reduction. In *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 1086–1090. IEEE, 2011.

[25] S. Becker and G. E. Hinton. Self-organizing neural network that discovers surfaces in random-dot stereograms. *Nature*, 355(6356):161–163, 1992.

[26] J. Bromley, I. Guyon, Y. LeCun, E. Säckinger, and R. Shah. Signature verification using a" siamese" time delay neural network. *Advances in neural information processing systems*, 6:737–744, 1993.

[27] A. Chaddad, C. Desrosiers, and T. Niazi. Deep radiomic analysis of mri related to alzheimer's disease. *IEEE Access*, 6:58213–58221, 2018.

[28] K. Chaitanya, E. Erdil, N. Karani, and E. Konukoglu. Contrastive learning of global and local features for medical image segmentation with limited annotations, 2020.

[29] C.-C. Chang, H.-H. Chen, Y.-C. Chang, M.-Y. Yang, C.-M. Lo, W.-C. Ko, Y.-F. Lee, K.-L. Liu, and R.-F. Chang. Computer-aided diagnosis of liver tumors on computed tomography images. *Computer methods and programs in biomedicine*, 145:45–51, 2017.

[30] S. Chaplot, L. M. Patnaik, and N. Jagannathan. Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. *Biomedical signal processing and control*, 1(1):86–92, 2006.

[31] R. Chaves, J. Ramírez, J. Górriz, M. López, D. Salas-Gonzalez, I. Alvarez, and F. Segovia. Svm-based computer-aided diagnosis of the alzheimer's disease using t-test nmse feature selection with feature correlation weighting. *Neuroscience letters*, 461(3):293–297, 2009.

[32] L. Chen, P. Bentley, K. Mori, K. Misawa, M. Fujiwara, and D. Rueckert. Self-supervised learning for medical image analysis using image context restoration. *Medical image analysis*, 58:101539, 2019.

[33] L.-C. Chen, A. Hermans, G. Papandreou, F. Schroff, P. Wang, and H. Adam. Masklab: Instance segmentation by refining object detection with semantic and direction features. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4013–4022, 2018.

[34] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pages 1597–1607. PMLR, 2020.

[35] X. Chen, L. Yao, T. Zhou, J. Dong, and Y. Zhang. Momentum contrastive learning for few-shot covid-19 diagnosis from chest ct images. *Pattern Recognition*, 113:107826, 2021.

[36] D. Cheng and M. Liu. Cnns based multi-modality classification for ad diagnosis. In *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pages 1–5. IEEE, 2017.

[37] Y.-A. Chung and W.-H. Weng. Learning deep representations of medical images using siamese cnns with application to content-based image retrieval. *arXiv preprint arXiv:1711.08490*, 2017.

[38] R. Collobert and J. Weston. A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th international conference on Machine learning*, pages 160–167, 2008.

[39] P. Costa, A. Galdran, M. I. Meyer, M. D. Abramoff, M. Niemeijer, A. M. Mendonça, and A. Campilho. Towards adversarial retinal image synthesis. *arXiv preprint arXiv:1701.08974*, 2017.

[40] R. Cui, M. Liu, and G. Li. Longitudinal analysis for alzheimer's disease diagnosis using rnn. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 1398–1401. IEEE, 2018.

[41] H. F. da Silva Lopes, J. M. Abe, and R. Anghinah. Application of paraconsistent artificial neural networks as a method of aid in the diagnosis of alzheimer disease. *Journal of medical systems*, 34(6):1073–1081, 2010.

[42] G. Developers. GAN Anatomy. https://developers.google.com/machine-learning/gan/generator#:~:text=The%20generator%20part%20of%20a,discriminator%20than%20discriminator%20training%20requires., 2019. Accessed: 2021-04-14.

[43] P. Domingos. A few useful things to know about machine learning. *Communications of the ACM*, 55(10):78–87, 2012.

[44] Q. Dou, D. C. Castro, K. Kamnitsas, and B. Glocker. Domain generalization via model-agnostic learning of semantic features. *arXiv preprint arXiv:1910.13580*, 2019.

[45] R. Elrod, E. R. Peskind, L. DiGiacomo, K. I. Brodkin, R. C. Veith, and M. A. Raskind. Effects of alzheimer's disease severity on cerebrospinal fluid norepinephrine concentration. *The American journal of psychiatry*, 154(1):25–30, 1997.

[46] B. J. Erickson, P. Korfiatis, Z. Akkus, and T. L. Kline. Machine learning for medical imaging. *Radiographics*, 37(2):505–515, 2017.

[47] A. Farooq, S. Anwar, M. Awais, and S. Rehman. A deep cnn based multi-class classification of alzheimer's disease using mri. In *2017 IEEE International Conference on Imaging systems and techniques (IST)*, pages 1–6. IEEE, 2017.

[48] L. Fei-Fei, R. Fergus, and P. Perona. One-shot learning of object categories. *IEEE transactions on pattern analysis and machine intelligence*, 28(4):594–611, 2006.

[49] J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. M. Parkin, D. Forman, and F. Bray. Cancer incidence and mortality worldwide: sources, methods and major patterns in globocan 2012. *International journal of cancer*, 136(5):E359–E386, 2015.

[50] A. R. Feyjie, R. Azad, M. Pedersoli, C. Kauffman, I. B. Ayed, and J. Dolz. Semi-supervised few-shot learning for medical image segmentation. *arXiv preprint arXiv:2003.08462*, 2020.

[51] M. Fink. Object classification from a single example utilizing class relevance metrics. *Advances in neural information processing systems*, 17:449–456, 2005.

[52] C. Finn, P. Abbeel, and S. Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *International Conference on Machine Learning*, pages 1126–1135. PMLR, 2017.

[53] M. Frid-Adar, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan. Synthetic data augmentation using gan for improved liver lesion classification. In *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*, pages 289–293. IEEE, 2018.

[54] J. Fritsch, S. Wankerl, and E. Nöth. Automatic diagnosis of alzheimer's disease using neural network language models. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5841–5845. IEEE, 2019.

[55] G. Fung and J. Stoeckel. Svm feature selection for classification of spect images of alzheimer's disease using spatial information. *Knowledge and Information Systems*, 11(2):243–258, 2007.

[56] V. Garcia and J. Bruna. Few-shot learning with graph neural networks. *arXiv preprint arXiv:1711.04043*, 2017.

[57] E. Gerardin, G. Chételat, M. Chupin, R. Cuingnet, B. Desgranges, H.-S. Kim, M. Niethammer, B. Dubois, S. Lehéricy, L. Garnero, et al. Multidimensional classification of hippocampal shape features discriminates alzheimer's disease and mild cognitive impairment from normal aging. *Neuroimage*, 47(4):1476–1486, 2009.

[58] S. Gidaris, P. Singh, and N. Komodakis. Unsupervised representation learning by predicting image rotations. *arXiv preprint arXiv:1803.07728*, 2018.

[59] M. Gletsos, S. G. Mougiakakou, G. K. Matsopoulos, K. S. Nikita, A. S. Nikita, and D. Kelekis. A computer-aided diagnostic system to characterize ct focal liver lesions: design and optimization of a neural network classifier. *IEEE transactions on information technology in biomedicine*, 7(3):153–162, 2003.

[60] Y. Goldberg. A primer on neural network models for natural language processing. *Journal of Artificial Intelligence Research*, 57:345–420, 2016.

[61] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.

[62] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial networks. *arXiv preprint arXiv:1406.2661*, 2014.

[63] S. Goswami and L. K. P. Bhaiya. Brain tumour detection using unsupervised learning based neural network. In *2013 International Conference on Communication Systems and Network Technologies*, pages 573–577. IEEE, 2013.

[64] G. Gosztolya, V. Vincze, L. Tóth, M. Pákáski, J. Kálmán, and I. Hoffmann. Identifying mild cognitive impairment and mild alzheimer's disease based on spontaneous speech using asr and linguistic features. *Computer Speech & Language*, 53:181–197, 2019.

[65] T. Goto, C. Wang, Y. Li, and Y. Tsuboshita. Multi-modal deep learning for predicting progression of alzheimer's disease using bi-linear shake fusion. In *Medical Imaging 2020: Computer-Aided Diagnosis*, volume 11314, page 113141X. International Society for Optics and Photonics, 2020.

[66] A. Graves and N. Jaitly. Towards end-to-end speech recognition with recurrent neural networks. In *International conference on machine learning*, pages 1764–1772. PMLR, 2014.

[67] K. R. Gray, P. Aljabar, R. A. Heckemann, A. Hammers, D. Rueckert, A. D. N. Initiative, et al. Random forest-based similarity measures for multi-modal classification of alzheimer's disease. *NeuroImage*, 65:167–175, 2013.

[68] H. Greenspan, B. Van Ginneken, and R. M. Summers. Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Transactions on Medical Imaging*, 35(5):1153–1159, 2016.

[69] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville. Improved training of wasserstein gans. *arXiv preprint arXiv:1704.00028*, 2017.

[70] A. M. Hafiz and G. M. Bhat. A survey on instance segmentation: state of the art. *International Journal of Multimedia Information Retrieval*, pages 1–19, 2020.

[71] C. Han, K. Murao, T. Noguchi, Y. Kawata, F. Uchiyama, L. Rundo, H. Nakayama, and S. Satoh. Learning more with less: conditional pggan-based data augmentation for brain metastases detection using highly-rough annotation on mr images. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pages 119–127, 2019.

[72] C. Han, L. Rundo, R. Araki, Y. Furukawa, G. Mauri, H. Nakayama, and H. Hayashi. Infinite brain mr images: Pggan-based data augmentation for tumor detection. In *Neural approaches to dynamics of signal exchanges*, pages 291–303. Springer, 2020.

[73] B. Hariharan and R. Girshick. Low-shot visual recognition by shrinking and hallucinating features. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3018–3027, 2017.

[74] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick. Momentum contrast for unsupervised visual representation learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[75] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition, 2015.

[76] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[77] T. Heimann, P. Mountney, M. John, and R. Ionasec. Learning without labeling: Domain adaptation for ultrasound transducer localization. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 49–56. Springer, 2013.

[78] O. Henaff. Data-efficient image recognition with contrastive predictive coding. In *International Conference on Machine Learning*, pages 4182–4192. PMLR, 2020.

[79] R. D. Hjelm, A. Fedorov, S. Lavoie-Marchildon, K. Grewal, P. Bachman, A. Trischler, and Y. Bengio. Learning deep representations by mutual information estimation and maximization, 2019.

[80] S. H. Hojjati, A. Ebrahimzadeh, A. Khazaee, A. Babajani-Feremi, A. D. N. Initiative, et al. Predicting conversion from mci to ad using resting-state fmri, graph theoretical approach and svm. *Journal of neuroscience methods*, 282:69–80, 2017.

[81] X. Hong, R. Lin, C. Yang, N. Zeng, C. Cai, J. Gou, and J. Yang. Predicting alzheimer's disease using lstm. *IEEE Access*, 7:80893–80901, 2019.

[82] E. Hosseini-Asl, G. Gimel'farb, and A. El-Baz. Alzheimer's disease diagnostics by a deeply supervised adaptable 3d convolutional network. *arXiv preprint arXiv:1607.00556*, 2016.

[83] X. Huang, M.-Y. Liu, S. Belongie, and J. Kautz. Multimodal unsupervised image-to-image translation. In *Proceedings of the European conference on computer vision (ECCV)*, pages 172–189, 2018.

[84] I. Illán, J. Górriz, M. López, J. Ramírez, D. Salas-Gonzalez, F. Segovia, R. Chaves, and C. G. Puntonet. Computer aided diagnosis of alzheimer's disease using component based svm. *Applied Soft Computing*, 11(2):2376–2382, 2011.

[85] J. Islam and Y. Zhang. A novel deep learning based multi-class classification method for alzheimer's disease detection using brain mri data. In *International conference on brain informatics*, pages 213–222. Springer, 2017.

[86] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.

[87] C. R. Jack Jr, M. S. Albert, D. S. Knopman, G. M. McKhann, R. A. Sperling, M. C. Carrillo, B. Thies, and C. H. Phelps. Introduction to the recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. *Alzheimer's & dementia*, 7(3):257–262, 2011.

[88] R. Jain, N. Jain, A. Aggarwal, and D. J. Hemanth. Convolutional neural network based alzheimer's disease classification from magnetic resonance brain images. *Cognitive Systems Research*, 57:147–159, 2019.

[89] A. Jaiswal, A. R. Babu, M. Z. Zadeh, D. Banerjee, and F. Makedon. A survey on contrastive self-supervised learning. *Technologies*, 9(1):2, 2021.

[90] A. Jamaludin, T. Kadir, and A. Zisserman. Self-supervised learning for spinal mris. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 294–302. Springer, 2017.

[91] L. Jing and Y. Tian. Self-supervised visual feature learning with deep neural networks: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.

[92] A. Jolicoeur-Martineau. The relativistic discriminator: a key element missing from standard gan, 2018.

[93] S. Joshi, D. Shenoy, G. V. Simha, P. Rrashmi, K. Venugopal, and L. Patnaik. Classification of alzheimer's disease and parkinson's disease by using machine learning and neural network methods. In *2010 Second International Conference on Machine Learning and Computing*, pages 218–222. IEEE, 2010.

[94] R. S. Kamathe and K. R. Joshi. A novel method based on independent component analysis for brain mr image tissue classification into csf, wm and gm for atrophy detection in alzheimer's disease. *Biomedical Signal Processing and Control*, 40:41–48, 2018.

[95] S. Kar and D. D. Majumder. A novel approach of diffusion tensor visualization based neuro fuzzy classification system for early detection of alzheimer's disease. *Journal of Alzheimer's disease reports*, 3(1):1–18, 2019.

[96] T. Karras, T. Aila, S. Laine, and J. Lehtinen. Progressive growing of gans for improved quality, stability, and variation. *arXiv preprint arXiv:1710.10196*, 2017.

[97] T. Karras, T. Aila, S. Laine, and J. Lehtinen. Progressive growing of gans for improved quality, stability, and variation, 2018.

[98] N. M. Khan, N. Abraham, and M. Hon. Transfer learning with intelligent training data selection for prediction of alzheimer's disease. *IEEE Access*, 7:72726–72735, 2019.

[99] N. M. Khan, N. Abraham, and M. Hon. Transfer learning with intelligent training data selection for prediction of alzheimer's disease. *IEEE Access*, 7:72726–72735, 2019.

[100] L. Khedher, I. A. Illán, J. M. Górriz, J. Ramírez, A. Brahim, and A. Meyer-Baese. Independent component analysis-support vector machine-based computer-aided diagnosis system for alzheimer's with visual support. *International journal of neural systems*, 27(03):1650050, 2017.

[101] L. Khedher, J. Ramírez, J. M. Górriz, A. Brahim, and I. Illán. Independent component analysis-based classification of alzheimer's disease from segmented mri data. In *International Work-Conference on the Interplay between Natural and Artificial Computation*, pages 78–87. Springer, 2015.

[102] P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, A. Maschinot, C. Liu, and D. Krishnan. Supervised contrastive learning, 2021.

[103] M. Kim, J. Zuallaert, and W. De Neve. Few-shot learning using a small-sized dataset of high-resolution fundus images for glaucoma diagnosis. In *Proceedings of the 2nd International Workshop on Multimedia for Personal Health and Health Care*, pages 89–92, 2017.

[104] D. P. Kingma and M. Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

[105] D. Kiyasseh, T. Zhu, and D. A. Clifton. Clocs: Contrastive learning of cardiac signals across space, time, and patients, 2020.

[106] S. Z. Klekociuk, J. J. Summers, J. C. Vickers, and M. J. Summers. Reducing false positive diagnoses in mild cognitive impairment: the importance of comprehensive neuropsychological assessment. *European journal of neurology*, 21(10):1330–e83, 2014.

[107] S. Klöppel, C. M. Stonnington, C. Chu, B. Draganski, R. I. Scahill, J. D. Rohrer, N. C. Fox, C. R. Jack Jr, J. Ashburner, and R. S. Frackowiak. Automatic classification of mr scans in alzheimer's disease. *Brain*, 131(3):681–689, 2008.

[108] G. Koch, R. Zemel, and R. Salakhutdinov. Siamese neural networks for one-shot image recognition. In *ICML deep learning workshop*, volume 2. Lille, 2015.

[109] S. Korolev, A. Safiullin, M. Belyaev, and Y. Dodonova. Residual and plain convolutional neural networks for 3d brain mri classification. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pages 835–838. IEEE, 2017.

[110] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105, 2012.

[111] V. Kumar BG, G. Carneiro, and I. Reid. Learning local image descriptors with deep siamese and triplet convolutional networks by minimising global loss functions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5385–5394, 2016.

[112] H. Laga, L. V. Jospin, F. Boussaid, and M. Bennamoun. A survey on deep learning techniques for stereo-based depth estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.

[113] S. Lahmiri and M. Boukadoum. New approach for automatic classification of alzheimer's disease, mild cognitive impairment and healthy brain magnetic resonance images. *Healthcare technology letters*, 1(1):32–36, 2014.

[114] S. Lahmiri and A. Shmuel. Performance of machine learning methods applied to structural mri and adas cognitive scores in diagnosing alzheimer's disease. *Biomedical Signal Processing and Control*, 52:414–419, 2019.

[115] G. Larsson, M. Maire, and G. Shakhnarovich. Colorization as a proxy task for visual understanding. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6874–6883, 2017.

[116] P. H. Le-Khac, G. Healy, and A. F. Smeaton. Contrastive representation learning: A framework and review. *IEEE Access*, 2020.

[117] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.

[118] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[119] G. Lee, K. Nho, B. Kang, K.-A. Sohn, and D. Kim. Predicting alzheimer's disease progression using multi-modal deep learning approach. *Scientific reports*, 9(1):1–12, 2019.

[120] K. K. Leung, M. J. Clarkson, J. W. Bartlett, S. Clegg, C. R. Jack Jr, M. W. Weiner, N. C. Fox, S. Ourselin, A. D. N. Initiative, et al. Robust atrophy rate measurement in alzheimer's disease using multi-site serial mri: tissue-specific intensity normalization and parameter selection. *Neuroimage*, 50(2):516–523, 2010.

[121] D. Li, Y. Yang, Y.-Z. Song, and T. Hospedales. Learning to generalize: Meta-learning for domain generalization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

[122] D.-C. Li, C.-S. Wu, T.-I. Tsai, and Y.-S. Lina. Using mega-trend-diffusion and artificial samples in small data set learning for early flexible manufacturing system scheduling knowledge. *Computers & Operations Research*, 34(4):966–982, 2007.

[123] F. Li, D. Cheng, and M. Liu. Alzheimer's disease classification based on combination of multi-model convolutional networks. In *2017 IEEE International Conference on Imaging Systems and Techniques (IST)*, pages 1–5. IEEE, 2017.

[124] X. Li, L. Yu, Y. Jin, C.-W. Fu, L. Xing, and P.-A. Heng. Difficulty-aware meta-learning for rare disease diagnosis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 357–366. Springer, 2020.

[125] Z. C. Lipton, J. Berkowitz, and C. Elkan. A critical review of recurrent neural networks for sequence learning. *arXiv preprint arXiv:1506.00019*, 2015.

[126] F. Liu, C. Shen, and G. Lin. Deep convolutional neural fields for depth estimation from a single image. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5162–5170, 2015.

[127] M. Liu, J. Zhang, E. Adeli, and D. Shen. Landmark-based deep multi-instance learning for brain disease diagnosis. *Medical image analysis*, 43:157–168, 2018.

[128] Q. Liu, Q. Dou, and P.-A. Heng. Shape-aware meta-learning for generalizing prostate mri segmentation to unseen domains. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 475–485. Springer, 2020.

[129] S. Liu, S. Liu, W. Cai, H. Che, S. Pujol, R. Kikinis, D. Feng, M. J. Fulham, et al. Multimodal neuroimaging feature learning for multiclass diagnosis of alzheimer's disease. *IEEE Transactions on Biomedical Engineering*, 62(4):1132–1140, 2014.

[130] S. Liu, S. Liu, W. Cai, S. Pujol, R. Kikinis, and D. Feng. Early diagnosis of alzheimer's disease with deep learning. In *2014 IEEE 11th international symposium on biomedical imaging (ISBI)*, pages 1015–1018. IEEE, 2014.

[131] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi. A survey of deep neural network architectures and their applications. *Neurocomputing*, 234:11–26, 2017.

[132] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.

[133] X. Long and C. Wyatt. An automatic unsupervised classification of mr images in alzheimer's disease. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2910–2917. IEEE, 2010.

[134] S. Lorio, F. Kherif, A. Ruef, L. Melie-Garcia, R. Frackowiak, J. Ashburner, G. Helms, A. Lutti, and B. Draganski. Neurobiological origin of spurious brain morphological changes: A quantitative mri study. *Human brain mapping*, 37(5):1801–1815, 2016.

[135] D. Lu, K. Popuri, G. W. Ding, R. Balachandar, and M. F. Beg. Multimodal and multiscale deep neural networks for the early diagnosis of alzheimer's disease using structural mr and fdg-pet images. *Scientific reports*, 8(1):1–13, 2018.

[136] D. Lu, K. Popuri, G. W. Ding, R. Balachandar, M. F. Beg, A. D. N. Initiative, et al. Multiscale deep neural network based analysis of fdg-pet images for the early diagnosis of alzheimer's disease. *Medical image analysis*, 46:26–34, 2018.

[137] A. Madani, J. R. Ong, A. Tibrewal, and M. R. Mofrad. Deep echocardiography: data-efficient supervised and semi-supervised deep learning towards automated diagnosis of cardiac disease. *NPJ digital medicine*, 1(1):1–11, 2018.

[138] T. Maggipinto, R. Bellotti, N. Amoroso, D. Diacono, G. Donvito, E. Lella, A. Monaco, M. A. Scelsi, S. Tangaro, A. D. N. Initiative, et al. Dti measurements for alzheimer's classification. *Physics in Medicine & Biology*, 62(6):2361, 2017.

[139] B. Magnin, L. Mesrob, S. Kinkingnéhun, M. Pélégrini-Issac, O. Colliot, M. Sarazin, B. Dubois, S. Lehéricy, and H. Benali. Support vector machine-based classification of alzheimer's disease from whole-brain anatomical mri. *Neuroradiology*, 51(2):73–83, 2009.

[140] R. Mahmood and B. Ghimire. Automatic detection and classification of alzheimer's disease from mri scans using principal component analysis and artificial neural networks. In *2013 20th International Conference on Systems, Signals and Image Processing (IWSSIP)*, pages 133–137. IEEE, 2013.

[141] G. Maicas, A. P. Bradley, J. C. Nascimento, I. Reid, and G. Carneiro. Training medical image analysis systems like radiologists. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 546–554. Springer, 2018.

[142] S. Marsland. *Machine learning: an algorithmic perspective*. CRC press, 2015.

[143] A. Masood, A. Al-Jumaily, and K. Anam. Self-supervised learning model for skin cancer diagnosis. In *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 1012–1015. IEEE, 2015.

[144] A. Mazaheri, K. Segaert, J. Olichney, J.-C. Yang, Y.-Q. Niu, K. Shapiro, and H. Bowman. Eeg oscillations during word processing predict mci conversion to alzheimer's disease. *NeuroImage: Clinical*, 17:188–197, 2018.

[145] A. Medela, A. Picon, C. L. Saratxaga, O. Belar, V. Cabezón, R. Cicchi, R. Bilbao, and B. Glover. Few shot learning in histopathological images: reducing the need of labeled data on biological datasets. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pages 1860–1864. IEEE, 2019.

[146] I. Melekhov, J. Kannala, and E. Rahtu. Siamese network features for image matching. In *2016 23rd International Conference on Pattern Recognition (ICPR)*, pages 378–383. IEEE, 2016.

[147] T. Mikolov, M. Karafiát, L. Burget, J. Černocký, and S. Khudanpur. Recurrent neural network based language model. In *Eleventh annual conference of the international speech communication association*, 2010.

[148] S. Minaee, Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, and D. Terzopoulos. Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

[149] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley. Deep learning for healthcare: review, opportunities and challenges. *Briefings in bioinformatics*, 19(6):1236–1246, 2018.

[150] A. Narin, C. Kaya, and Z. Pamuk. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. *arXiv preprint arXiv:2003.10849*, 2020.

[151] D. Nie, R. Trullo, J. Lian, C. Petitjean, S. Ruan, Q. Wang, and D. Shen. Medical image synthesis with context-aware generative adversarial networks. In *International conference on medical image computing and computer-assisted intervention*, pages 417–425. Springer, 2017.

[152] M. Noroozi and P. Favaro. Unsupervised learning of visual representations by solving jigsaw puzzles. In *European conference on computer vision*, pages 69–84. Springer, 2016.

[153] B. Nowok, G. M. Raab, C. Dibben, et al. synthpop: Bespoke creation of synthetic data in r. *Journal of statistical software*, 74(11):1–26, 2016.

[154] Y. Oh, S. Park, and J. C. Ye. Deep learning covid-19 features on cxr using limited training data sets. *IEEE Transactions on Medical Imaging*, 39(8):2688–2700, 2020.

[155] W. K. Oleiwi. Alzheimer disease diagnosis using the k-means, glcm and k_nn. *Journal of University of Babylon for Pure and Applied Sciences*, 26(2):57–65, 2018.

[156] A. Ortiz, J. Munilla, I. Álvarez-Illán, J. M. Górriz, J. Ramírez, A. D. N. Initiative, et al. Exploratory graphical models of functional and structural connectivity patterns for alzheimer's disease diagnosis. *Frontiers in computational neuroscience*, 9:132, 2015.

[157] A. Ortiz, J. Munilla, J. M. Gorriz, and J. Ramirez. Ensembles of deep learning architectures for the early diagnosis of the alzheimer's disease. *International journal of neural systems*, 26(07):1650025, 2016.

[158] P. Padilla, J. Górriz, J. Ramírez, E. Lang, R. Chaves, F. Segovia, M. López, D. Salas-González, and I. Álvarez. Analysis of spect brain images for the diagnosis of alzheimer's disease based on nmf for feature extraction. *Neuroscience letters*, 479(3):192–196, 2010.

[159] S. J. Pan and Q. Yang. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359, 2009.

[160] A. Paul, T. C. Shen, N. Balachandar, Y. Tang, Y. Peng, Z. Lu, and R. M. Summers. Comesee: Cross-modality semantic embedding ensemble for generalized zero-shot diagnosis of chest radiographs. In *Interpretable and Annotation-Efficient Learning for Medical Image Computing*, pages 103–111. Springer, 2020.

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

[162] J. Peng, X. Zhu, Y. Wang, L. An, and D. Shen. Structured sparsity regularized multiple kernel learning for alzheimer's disease diagnosis. *Pattern recognition*, 88:370–382, 2019.

[163] L. Peng, W. Chen, W. Zhou, F. Li, J. Yang, and J. Zhang. An immune-inspired semi-supervised algorithm for breast cancer diagnosis. *Computer methods and programs in biomedicine*, 134:259–265, 2016.

[164] C. Plant, S. J. Teipel, A. Oswald, C. Böhm, T. Meindl, J. Mourao-Miranda, A. W. Bokde, H. Hampel, and M. Ewers. Automated detection of brain atrophy patterns based on mri for the prediction of alzheimer's disease. *Neuroimage*, 50(1):162–174, 2010.

[165] M. Plocharski, L. R. Østergaard, A. D. N. Initiative, et al. Extraction of sulcal medial surface and classification of alzheimer's disease using sulcal features. *Computer methods and programs in biomedicine*, 133:35–44, 2016.

[166] pydicom. pydicom. https://github.com/pydicom/pydicom, 2020. Accessed: 2021-04-19.

[167] A. Radford, L. Metz, and S. Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.

[168] A. Radford, L. Metz, and S. Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks, 2016.

[169] M. Raghu, C. Zhang, J. Kleinberg, and S. Bengio. Transfusion: Understanding transfer learning for medical imaging, 2019.

[170] W. Raghupathi and V. Raghupathi. Big data analytics in healthcare: promise and potential. *Health information science and systems*, 2(1):1–10, 2014.

[171] J. Ramírez, J. M. Górriz, M. López, D. Salas-Gonzalez, I. Álvarez, F. Segovia, and C. G. Puntonet. Early detection of the alzheimer disease combining feature selection and kernel machines. In *International Conference on Neural Information Processing*, pages 410–417. Springer, 2008.

[172] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.

[173] A. Reghu and R. John. Glcm based feature extraction of neurodegenerative disease for regional brain patterns. *Int J Eng Res Gen Sci*, 4:319–323, 2016.

[174] M. Ren, E. Triantafillou, S. Ravi, J. Snell, K. Swersky, J. B. Tenenbaum, H. Larochelle, and R. S. Zemel. Meta-learning for semi-supervised few-shot classification. *arXiv preprint arXiv:1803.00676*, 2018.

[175] S. Ren, K. He, R. Girshick, and J. Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *arXiv preprint arXiv:1506.01497*, 2015.

[176] T. Ross, D. Zimmerer, A. Vemuri, F. Isensee, M. Wiesenfarth, S. Bodenstedt, F. Both, P. Kessler, M. Wagner, B. Müller, et al. Exploiting the potential of unlabeled endoscopic video data with self-supervised learning. *International journal of computer assisted radiology and surgery*, 13(6):925–933, 2018.

[177] A. Sabay, L. Harris, V. Bejugama, and K. Jaceldo-Siegl. Overcoming small data limitations in heart disease prediction by using surrogate data. *SMU Data Science Review*, 1(3):12, 2018.

[178] D. Salas-Gonzalez, J. M. Górriz, J. Ramírez, M. López, I. Alvarez, F. Segovia, R. Chaves, and C. Puntonet. Computer-aided diagnosis of alzheimer's disease using support vector machines and classification trees. *Physics in Medicine & Biology*, 55(10):2807, 2010.

[179] Z. Sankari and H. Adeli. Probabilistic neural networks for diagnosis of alzheimer's disease using conventional and wavelet coherence. *Journal of neuroscience methods*, 197(1):165–170, 2011.

[180] A. Santoro, S. Bartunov, M. Botvinick, D. Wierstra, and T. Lillicrap. One-shot learning with memory-augmented neural networks. *arXiv preprint arXiv:1605.06065*, 2016.

[181] S. Sarraf, G. Tofighi, A. D. N. Initiative, et al. Deepad: Alzheimer's disease classification via deep convolutional neural networks using mri and fmri. *BioRxiv*, page 070441, 2016.

[182] A. Savio, M. García-Sebastián, C. Hernández, M. Graña, and J. Villanúa. Classification results of artificial neural networks for alzheimer's disease detection. In *International Conference on Intelligent Data Engineering and Automated Learning*, pages 641–648. Springer, 2009.

[183] J. Schmidhuber. Deep learning in neural networks: An overview. *Neural networks*, 61:85–117, 2015.

[184] D. Schmitter, A. Roche, B. Maréchal, D. Ribes, A. Abdulkadir, M. Bach-Cuadra, A. Daducci, C. Granziera, S. Klöppel, P. Maeder, et al. An evaluation of volume-based morphometry for prediction of mild cognitive impairment and alzheimer's disease. *NeuroImage: Clinical*, 7:7–17, 2015.

[185] T. Schreiber and A. Schmitz. Improved surrogate data for nonlinearity tests. *Physical review letters*, 77(4):635, 1996.

[186] F. Segovia, J. Górriz, J. Ramírez, D. Salas-González, I. Álvarez, M. López, R. Chaves, and P. Padilla. Classification of functional brain images using a gmm-based multi-variate approach. *Neuroscience Letters*, 474(1):58–62, 2010.

[187] T. Shaikhina and N. A. Khovanova. Handling limited datasets with neural networks in medical applications: A small-data approach. *Artificial intelligence in medicine*, 75:51–63, 2017.

[188] D. Shen, C.-Y. Wee, D. Zhang, L. Zhou, and P.-T. Yap. Machine learning techniques for ad/mci diagnosis and prognosis. In *Machine learning in healthcare informatics*, pages 147–179. Springer, 2014.

[189] D. Shen, G. Wu, and H.-I. Suk. Deep learning in medical image analysis. *Annual review of biomedical engineering*, 19:221–248, 2017.

[190] J. Sheng, B. Wang, Q. Zhang, Q. Liu, Y. Ma, W. Liu, M. Shao, and B. Chen. A novel joint hcpmmp method for automatically classifying alzheimer's and different stage mci patients. *Behavioural brain research*, 365:210–221, 2019.

[191] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers. Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. *IEEE transactions on medical imaging*, 35(5):1285–1298, 2016.

[192] H.-C. Shin, N. A. Tenenholtz, J. K. Rogers, C. G. Schwarz, M. L. Senjem, J. L. Gunter, K. P. Andriole, and M. Michalski. Medical image synthesis for data augmentation and anonymization using generative adversarial networks. In *International workshop on simulation and synthesis in medical imaging*, pages 1–11. Springer, 2018.

[193] A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, and R. Webb. Learning from simulated and unsupervised images through adversarial training. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2107–2116, 2017.

[194] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[195] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition, 2015.

[196] J. Snell, K. Swersky, and R. S. Zemel. Prototypical networks for few-shot learning. *arXiv preprint arXiv:1703.05175*, 2017.

[197] R. Socher, M. Ganjoo, H. Sridhar, O. Bastani, C. D. Manning, and A. Y. Ng. Zero-shot learning through cross-modal transfer. *arXiv preprint arXiv:1301.3666*, 2013.

[198] R. Socher, C. C.-Y. Lin, A. Y. Ng, and C. D. Manning. Parsing natural scenes and natural language with recursive neural networks. In *ICML*, 2011.

[199] R. Socher, A. Perelygin, J. Wu, J. Chuang, C. D. Manning, A. Y. Ng, and C. Potts. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 conference on empirical methods in natural language processing*, pages 1631–1642, 2013.

[200] R. A. Sperling, P. S. Aisen, L. A. Beckett, D. A. Bennett, S. Craft, A. M. Fagan, T. Iwatsubo, C. R. Jack Jr, J. Kaye, T. J. Montine, et al. Toward defining the preclinical stages of alzheimer's disease: recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. *Alzheimer's & dementia*, 7(3):280–292, 2011.

[201] H.-I. Suk, S.-W. Lee, and D. Shen. Latent feature representation with stacked auto-encoder for ad/mci diagnosis. *Brain Structure and Function*, 220(2):841–859, 2015.

[202] H.-I. Suk, S.-W. Lee, D. Shen, A. D. N. Initiative, et al. Hierarchical feature representation and multimodal fusion with deep learning for ad/mci diagnosis. *NeuroImage*, 101:569–582, 2014.

[203] H.-I. Suk and D. Shen. Deep learning-based feature representation for ad/mci classification. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 583–590. Springer, 2013.

[204] F. Sung, Y. Yang, L. Zhang, T. Xiang, P. H. Torr, and T. M. Hospedales. Learning to compare: Relation network for few-shot learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1199–1208, 2018.

[205] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions, 2014.

[206] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.

[207] M. Tan and Q. Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International Conference on Machine Learning*, pages 6105–6114. PMLR, 2019.

[208] F. H. K. d. S. Tanaka and C. Aranha. Data augmentation using gans. *arXiv preprint arXiv:1904.09135*, 2019.

[209] Y. Tang. Deep learning using linear support vector machines. *arXiv preprint arXiv:1306.0239*, 2013.

[210] S. Tangaro, A. Fanizzi, N. Amoroso, R. Bellotti, A. D. N. Initiative, et al. A fuzzy-based system reveals alzheimer's disease onset in subjects with mild cognitive impairment. *Physica Medica*, 38:36–44, 2017.

[211] M. Tanveer, B. Richhariya, R. Khan, A. Rashid, P. Khanna, M. Prasad, and C. Lin. Machine learning techniques for the diagnosis of alzheimer's disease: A review. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 16(1s):1–35, 2020.

[212] A. M. Taqi, A. Awad, F. Al-Azzo, and M. Milanova. The impact of multi-optimizers and data augmentation on tensorflow convolutional neural network performance. In *2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, pages 140–145. IEEE, 2018.

[213] Y. Tian, D. Krishnan, and P. Isola. Contrastive multiview coding, 2020.

[214] Y. Tian, C. Sun, B. Poole, D. Krishnan, C. Schmid, and P. Isola. What makes for good views for contrastive learning?, 2020.

[215] J. Timmer. Power of surrogate data testing with respect to nonstationarity. *Physical Review E*, 58(4):5153, 1998.

[216] A. van den Oord, Y. Li, and O. Vinyals. Representation learning with contrastive predictive coding, 2019.

[217] J. Vanschoren. Meta-learning: A survey. *arXiv preprint arXiv:1810.03548*, 2018.

[218] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. Attention is all you need, 2017.

[219] P. Vemuri, J. L. Gunter, M. L. Senjem, J. L. Whitwell, K. Kantarci, D. S. Knopman, B. F. Boeve, R. C. Petersen, and C. R. Jack Jr. Alzheimer's disease diagnosis in individual subjects using structural mr images: validation studies. *Neuroimage*, 39(3):1186–1197, 2008.

[220] O. Vinyals, C. Blundell, T. Lillicrap, K. Kavukcuoglu, and D. Wierstra. Matching networks for one shot learning. *arXiv preprint arXiv:1606.04080*, 2016.

[221] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, and X. Tang. Residual attention network for image classification, 2017.

[222] H. Wang, Y. Shen, S. Wang, T. Xiao, L. Deng, X. Wang, and X. Zhao. Ensemble of 3d densely connected convolutional network for diagnosis of mild cognitive impairment and alzheimer's disease. *Neurocomputing*, 333:145–156, 2019.

[223] L. Wang, Z. Q. Lin, and A. Wong. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Scientific Reports*, 10(1):1–12, 2020.

[224] S.-H. Wang, P. Phillips, Y. Sui, B. Liu, M. Yang, and H. Cheng. Classification of alzheimer's disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling. *Journal of medical systems*, 42(5):1–11, 2018.

[225] W. Wang, V. W. Zheng, H. Yu, and C. Miao. A survey of zero-shot learning: Settings, methods, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(2):1–37, 2019.

[226] Y. Wang, Q. Yao, J. T. Kwok, and L. M. Ni. Generalizing from a few examples: A survey on few-shot learning. *ACM Computing Surveys (CSUR)*, 53(3):1–34, 2020.

[227] Z. Wang, Q. She, and T. E. Ward. Generative adversarial networks in computer vision: A survey and taxonomy. *arXiv preprint arXiv:1906.01529*, 2019.

[228] G. H. Weissberger, J. V. Strong, K. B. Stefanidis, M. J. Summers, M. W. Bondi, and N. H. Stricker. Diagnostic accuracy of memory measures in alzheimer's dementia and mild cognitive impairment: a systematic review and meta-analysis. *Neuropsychology review*, 27(4):354–388, 2017.

[229] J. Wen, E. Thibeau-Sutre, M. Diaz-Melo, J. Samper-González, A. Routier, S. Bottani, D. Dormont, S. Durrleman, N. Burgos, O. Colliot, et al. Convolutional neural networks for classification of alzheimer's disease: Overview and reproducible evaluation. *Medical image analysis*, 63:101694, 2020.

[230] W. H. O. (WHO). WHO Coronavirus (COVID-19) Dashboard. https://covid19.who.int/, 2021. Accessed: 2021-04-09.

[231] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon. Cbam: Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)*, pages 3–19, 2018.

[232] Z. Wu, Y. Xiong, S. X. Yu, and D. Lin. Unsupervised feature learning via non-parametric instance discrimination. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[233] Y. Xian, B. Schiele, and Z. Akata. Zero-shot learning-the good, the bad and the ugly. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4582–4591, 2017.

[234] Z. Xiao, Y. Ding, T. Lan, C. Zhang, C. Luo, and Z. Qin. Brain mr image classification for alzheimer's disease diagnosis based on multifeature fusion. *Computational and mathematical methods in medicine*, 2017, 2017.

[235] R. Xu and D. Wunsch. Survey of clustering algorithms. *IEEE Transactions on neural networks*, 16(3):645–678, 2005.

[236] S.-T. Yang, J.-D. Lee, T.-C. Chang, C.-H. Huang, J.-J. Wang, W.-C. Hsu, H.-L. Chan, Y.-Y. Wai, and K.-Y. Li. Discrimination between alzheimer's disease and mild cognitive impairment using som and pso-svm. *Computational and mathematical methods in medicine*, 2013, 2013.

[237] R. Zhang, T. Che, Z. Ghahramani, Y. Bengio, and Y. Song. Metagan: An adversarial approach to few-shot learning. *NeurIPS*, 2:8, 2018.

[238] Y. Zhang, H. Jiang, Y. Miura, C. D. Manning, and C. P. Langlotz. Contrastive learning of medical visual representations from paired images and text, 2020.

[239] Y. Zhang, S. Wang, P. Phillips, Z. Dong, G. Ji, and J. Yang. Detection of alzheimer's disease and mild cognitive impairment based on structural volumetric mr images using 3d-dwt and wta-ksvm trained by psotvac. *Biomedical Signal Processing and Control*, 21:58–73, 2015.

[240] Y. Zhang, Y. Wei, Q. Wu, P. Zhao, S. Niu, J. Huang, and M. Tan. Collaborative unsupervised domain adaptation for medical image diagnosis. *IEEE Transactions on Image Processing*, 29:7834–7844, 2020.

[241] A. Zhao, G. Balakrishnan, F. Durand, J. V. Guttag, and A. V. Dalca. Data augmentation using learned transformations for one-shot medical image segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8543–8553, 2019.

[242] M. Zhao, R. H. Chan, T. W. Chow, and P. Tang. Compact graph based semi-supervised learning for medical diagnosis in alzheimer's disease. *IEEE signal processing letters*, 21(10):1192–1196, 2014.

[243] Z.-Q. Zhao, P. Zheng, S.-t. Xu, and X. Wu. Object detection with deep learning: A review. *IEEE transactions on neural networks and learning systems*, 30(11):3212–3232, 2019.

[244] X. Zhu and A. B. Goldberg. Introduction to semi-supervised learning. *Synthesis lectures on artificial intelligence and machine learning*, 3(1):1–130, 2009.

[245] C. Zhuang, A. L. Zhai, and D. Yamins. Local aggregation for unsupervised learning of visual embeddings. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2019.

Acknowledgment

I am very thankful to the Professor Kim for providing me the opportunity to work under his supervision and for always leading me to the right path. It as been a pleasure to do research together and learn from him. Thank you for believing in me. I am also grateful to my lab mate and colleague Taemin, who worked on this project alongside me. He was always there to guide me and help me in understanding things. I learned a lot from him. Finally I am grateful to my family. It is because of their continuous efforts that I am here today. I dedicate everything to them.

Curriculum Vitae

Name   : Raja Haseeb

Date of Birth : July 17, 1997

Birthplace   : Islamabad, Pakistan

Address   : KAIST, 291 Daehak-ro, Eoeun-dong, Yuseong-gu, Daejeon, South Korea

Educations

2012. 7. – 2014. 7. College

2014. 9. – 2018. 7. University

2019. 9. – 2021. 7. Master

Career

2019. 9. – 2020. 8. Opto-Electro Structural Lab, KAIST

2020. 8. – 2021. 6. Robot Intelligence Technology Lab, KAIST