Degree Project in Computer Science

Second cycle, 30 credits

An Unsupervised Exploration of the CNN Feature Space for Phenotype Structure Discovery in Alzheimer's Disease

SOFIA HARMÉN

An Unsupervised Exploration of the CNN Feature Space for Phenotype Structure Discovery in Alzheimer's Disease

SOFIA HARMÉN

Master's Programme, Computer Science, 120 credits
Date: January 27, 2026

Supervisor: Pawel Andrzej Herman
Examiner: Arvind Kumar
School of Electrical Engineering and Computer Science
Swedish title: En oövervakad utforskning av CNN representationsrymd för upptäckt av fenotypisk struktur vid Alzheimers sjukdom

Abstract

Alzheimer's disease (AD) is a progressive neurodegenerative disorder and the leading cause of dementia worldwide. Early detection and understanding of the disease heterogeneity are crucial for improving the quality of life for patients. Even though convolutional neural networks (CNNs) have shown strong performance in classifying AD from structural MRI scans, most existing work focuses on supervised classification rather than exploring disease subtypes and associated phenotypes. Unsupervised learning can be used to explore such structures and can consequently help deepen our understanding of disease progression and support treatments tailored to individual patient needs. To address the issue, this exploratory study used unsupervised learning to investigate whether a CNN-learned feature space can reveal clinically meaningful patterns related to AD progression. A ResNet-18 CNN was trained on structural MRI data from the ADNI dataset, and the learned feature representations were analysed using unsupervised clustering and complementary statistical methods in relation to clinical variables. The results indicate that the CNN learned a structured and clinically meaningful feature space. The feature representations captured both genetic and environmental influences embedded in brain structure, and the unsupervised clustering recovered three distinct subgroups that aligned strongly with diagnostic categories but primarily reflected the underlying phenotypic variability. Overall, the findings show the potential of unsupervised deep learning approaches for discovering phenotypes from neuroimaging and deepen the understanding of AD heterogeneity.

Keywords

Alzheimer's disease, Magnetic resonance imaging, Convolutional neural networks, Deep learning, Unsupervised learning, Clustering, Phenotype discovery

Sammanfattning

Alzheimers sjukdom (AD) är en progressiv neurodegenerativ sjukdom och den vanligaste orsaken till demens i världen. En tidig upptäckt och bättre förståelse av sjukdomens heterogenitet är avgörande för att förbättra livskvaliteten hos patienter. Även om konvolutionella neurala nätverk (CNN) har visat starka resultat för klassificeringen av AD från strukturell MR-data, fokuserar majoriteten av tidigare studier på övervakad klassificering snarare än på att utforska sjukdomens undergrupper och fenotyper. Oövervakad inlärning kan användas för att analysera sådana strukturer, vilket kan bidra till en djupare förståelse av sjukdomens progression och stödja utvecklingen av mer individanpassade behandlingar. I denna studie användes oövervakad inlärning för att undersöka om CNN-baserade representationsrymder kan framhäva kliniskt meningsfulla mönster relaterade till progressionen av AD. En ResNet-18 CNN arkitektur tränades på strukturella MR-bilder från databasen ADNI och därefter analyserades de inlärda representationerna med hjälp av oövervakad klustring och kompletterande statistiska metoder i relation till kliniska variabler. Resultaten visar att CNN-modellen lärde sig en strukturerad och kliniskt meningsfull representationsrymd. Representationerna fångade både genetiska och miljömässiga förändringar inbäddade i hjärnans struktur och den oövervakade klustringen identifierade tre distinkta undergrupper som hade en stark överensstämmelse med de diagnostiska kategorierna men som framförallt speglade en underliggande fenotypisk variation. Sammanfattningsvis visar dessa resultat potentialen hos oövervakade djupinlärningsmetoder för att upptäcka fenotyper från neuroavbildningar samt för att få en fördjupad förståelse av heterogeniteten hos AD.

Nyckelord

Alzheimers sjukdom, Magnetisk resonanstomografi, Konvolutionella neurala nätverk, Djupinlärning, Oövervakad inlärning, Klustring, Fenotypupptäckt

Contents

List of acronyms and abbreviations

AD	Alzheimer's disease
ADNI	Alzheimer's Disease Neuroimaging Initiative
APOE	Apolipoprotein E
ARI	Adjusted Rand Index
CAD	Computer-Aided Diagnosis
CN	Cognitive Normal
CNN	Convolutional Neural Network
CSF	Cerebrospinal Fluid
FSL	FMRIB Software Library
GMM	Gaussian Mixture Model
KDE	Kernel Density Estimation
MCI	Mild Cognitive Impairment
MLR	Multinomial Logistic Regression
MNI	Montreal Neurological Institute
MRI	Magnetic Resonance Imaging
PC	Principal Component
PCA	Principal Component Analysis
PET	Positron Emission Tomography
ResNet	Residual Network
t-SNE	t-Distributed Stochastic Neighbor Embedding
UMAP	Uniform Manifold Approximation and Projection
VIF	Variance Inflation Factor
WCSS	Within-Cluster Sum of Squares

Chapter 1

Introduction

1.1 Background

Alzheimer's Disease (AD) is a progressive and irreversible neurodegenerative disorder that affects memory and cognitive abilities. AD is the most common cause of dementia worldwide, where the symptoms worsen over time. However, it is possible to slow the progression through early treatment, which can greatly improve the quality of life for a patient with AD [1].

Magnetic Resonance Imaging (MRI) is one of the most widely used biomarkers for AD diagnosis [2]. It is a medical scan that uses a powerful magnetic field and radio waves to create a detailed 3D representation of internal organs in the body. Raw MRI scans typically undergo baseline preprocessing techniques including intensity normalisation, skull stripping, affine transformation, and image registration. Then, the MRI can be either kept as a full 3D volume or divided into 2D slices of axial, coronal, and sagittal planes, where all or one plane can be used. In AD research, the axial plane is the most commonly used view. One of the most common databases for obtaining AD MRI data is the Alzheimer's Disease Neuroimaging Initiative (ADNI) [3].

With the recent advances in deep learning, computer-aided diagnostics (CAD) have become increasingly important in helping early and accurate diagnosis of AD. Convolutional Neural Networks (CNNs) are deep neural networks widely used for image analysis inspired by the visual cortex of the brain [4]. It is one of the most successful deep models for extracting features from 2D and 3D images to use for AD classification. CNNs identify visual features in images on increasingly complex and abstract levels throughout their convolutional layers, which makes them effective for neuroimaging tasks. The ResNet-18 model is a variant of the Residual Network (ResNet) CNN architecture with 18 layers, introduced by He et al. in 2015 [5]. The novel idea was to use residual blocks, which are connections that allow the network to bypass layers as a solution to the vanishing gradient problem occurring in deep networks.

AD progression is commonly described using three baseline diagnostic categories. These are Cognitive Normal (CN), Mild Cognitive Impairment (MCI), and Alzheimer's Disease (AD). There are several clinical factors that are known to be associated with disease risk and progression, including age, sex, educational level, and specific genes [2]. Age is one of the strongest risk factors for AD. Although dementia is not a normal consequence of ageing, the likelihood of developing AD increases with age as structural changes occur in the brain [6]. Sex differences have also been observed, with approximately twice as many women developing AD compared to men, though the underlying causes are not fully understood. Potential explanations include lifespan expectancies, hormonal changes, and historical disparities in gender roles, such as educational opportunities, which could negatively affect cognitive reserve [7]. Cognitive reserve is believed to decrease the risk of developing AD, as mental stimulation, such as higher educational levels, complex work, and social engagement, enhances the brain's ability to maintain function later in life [8]. Apolipoprotein E (APOE) is the primary risk gene associated with AD, with three different allele types associated with different levels of risk. The ε3 allele is the most common and is considered neutral, the ε4 allele is associated with increased disease risk, and the ε2 allele is thought to have a decreased risk. Since each individual inherits one allele from each parent, there are different allele combinations of these types [9].

1.2 Problem

Phenotypes are observable traits that arise from genetic and environmental influences and, in the context of neuroimaging, can be reflected as distinct patterns in brain structure [10]. In AD research, deep learning models are commonly used to learn such patterns from MRI data for diagnostic classification rather than for explicit analysis. Consequently, there is an issue with the lack of unsupervised approaches that aim to discover phenotypic structures directly within learned feature spaces. Therefore, exploring unsupervised deep learning approaches for studying AD has the potential to deepen the understanding of disease progression and to support the development of treatments tailored to individual patient needs.

1.3 Purpose

This exploratory study aims to train a ResNet-18 CNN architecture on 2D MRI slices from the ADNI dataset to extract feature representations for unsupervised clustering. Through complementary statistical analyses, the study investigates whether the learned feature space captures clinically meaningful patterns related to AD progression and associated clinical variables. This research aim is addressed through three exploratory objectives:

(*i*) To investigate whether unsupervised clustering of the learned MRI features reveals a clinically meaningful structure.

(*ii*) To explore how clinical variables are reflected within the feature space.

(*iii*) To assess whether the CNN captures phenotypic variability beyond the traditional diagnostic categories.

1.4 Scope

This study is conducted within the computer science scope and focuses on the application of deep learning and unsupervised learning methods for exploratory analysis of neuroimaging data. With this in mind, a few delimitations are established. The work is limited to exploring phenotypic structure within the learned feature space and does not aim to define new clinical phenotypes or diagnostic subtypes. Furthermore, only a single CNN architecture is used for feature extraction, as the primary focus is on analysis rather than on model optimisation or performance improvement. Finally, the data used are exclusively from the ADNI database and therefore reflect a relatively narrow demographic and are based on axial 2D MRI slices.

Chapter 2

Background

2.1 CNNs in AD classification tasks

Several studies have demonstrated the potential of deep learning methods for classifying AD from MRI, with CNNs showing particularly strong performance in AD classification tasks using both 3D and 2D convolutions. In 2015, Payan et al. [11] developed a CNN pre-trained with a sparse autoencoder for classification. Their novel approach was using 3D convolutions to capture a volumetric brain image from MRI and then compared its performance to using 2D convolutions on MRI slices. The 3D approach outperformed the 2D approach in classifying the three baseline AD classes, although with a small margin. Building on this work, Hosseini-Asl et al. [12] developed a 3D adaptable CNN with a 3D convolutional autoencoder in 2016 for feature extraction, pre-trained on the CADDementia dataset and fine-tuned with MRI from ADNI. The results indicated significantly higher accuracy compared to previous traditional classifiers. Kundaram et al. [13] and Salehi et al. [14] trained CNNs on MRI slices from ADNI for 3-way classification, achieving a classification accuracy of 98.57% and 99% respectively in 2020 and 2021. Overall, these studies highlight the potential of CNNs for neuroimaging tasks.

2.2 Transfer Learning in AD classification

Several works have also explored transfer learning using established CNN architectures. Commonly used models include ResNet, GoogleNet, LeNet-5, and VGGNet [15]. Sarraf et al. [16] applied LeNet and GoogleNet to pre-processed MRI data for binary classification between AD and NC, with GoogleNet achieving the highest accuracy. Farooq et al. [17] compared 3-way classification accuracies between ResNet and GoogleNet, where both models received similar results. Stoleru and Iftene [18] compared a ResNet model to an AlexNet model trained on different planes of pre-processed MRI from ADNI and reported that ResNet-152 achieved the highest accuracy of 99.96% in a binary classification between AD and CN. Based on these findings, the ResNet architecture demonstrates a strong performance across different neuroimaging tasks, which motivates its use for feature extraction in this study.

2.3 Diagnostic variability

CN, MCI, and AD are the standard diagnostic categories for AD. However, given that MCI has great internal variability due to its transitional nature, several studies have extended classification tasks to include MCI subtypes, distinguishing between a stable MCI state and a converting MCI state. Basaia et al. [19] performed binary classification between stable and converting MCI, achieving an accuracy of 75%, and also reported high accuracy between AD and CN classification on data augmented MRI from ADNI. Farooq et al. [17] performed 4-way classification, including the two transitional MCI states, and achieved an accuracy of 98.8% with the GoogleNet architecture. Helaly et al. [20] compared a 2D CNN, 3D CNN, and a transfer learning VGG-19 method for 4-way classification, and reported similar results between the different approaches but achieved the overall highest accuracy with the VGG-19 method. These studies motivate the exploration of underlying phenotypic structure beyond predefined diagnostic labels.

2.4 Integration of clinical variables

Some studies have also highlighted the importance of integrating clinical data in their research rather than relying on a single imaging modality. Venugopalan et al. [21] combined features from multiple data modalities, including MRI, genetic, and clinical data from ADNI, to improve classification accuracy. Estarellas et al. [22] used multiple pre-defined biomarkers in their disease progression framework to discover subtypes of AD and model their trajectory through associations with clinical variables. These studies show that AD is a heterogeneous disease and suggest that clinical variables may contribute to underlying patterns that are not fully captured by diagnostic labels alone.

2.5 Unsupervised learning in AD research

With the strong focus on AD classification, comparatively few studies have applied unsupervised learning to MRI-based AD analysis. Bi et al. [23] proposed an unsupervised approach using feature extraction based on PCANet followed by K-means clustering, which was then evaluated through classification performance. In their work, they compared using a single MRI slice from one plane to using all slices from all three orthogonal planes. They reported similar 3-way classification accuracies for these methods, with the more informative MRI method achieving only a slightly higher accuracy, indicating the potential of using slices from a single plane. While their approach incorporated unsupervised clustering, the focus remained on classification.

2.6 Study motivation

Existing literature includes relatively few studies that have focused on using unsupervised learning approaches beyond AD classification. Therefore, this study aims to use unsupervised learning to investigate whether a CNN-learned feature space can reveal clinically meaningful patterns related to AD progression. Specifically, it will be explored how the feature representations captured from MRI relate to clinical variables, providing insight into the information that is implicitly captured by the CNN. Based on previous research, a ResNet-18 architecture is trained on pre-processed 2D MRI slices from a single plane, with the three diagnostic categories used as a baseline reference for exploratory analysis.

Chapter 3

Methods

3.1 Research process

The methodology follows a research process consisting of the following steps:

1. Collecting MRI and registry data from the ADNI database.
2. Pre-processing MRI data.
3. Training a ResNet-18 CNN and extracting feature vectors.
4. Performing dimensionality reduction, unsupervised clustering, and visualisation of the learned feature space.
5. Evaluating and statistically analysing the results.

An overview of the methodological pipeline is seen in Figure 3.1.

3.2 Data

3.2.1 MRI images

The data used in this study were obtained from the ADNI database [3]. ADNI was funded in 2004 as a research initiative organization by the National Institute of Biomedical Imaging and Bioengineering led by Principal Investigator Michael W. Weiner, MD. The goal of the organisation is to evaluate the progression of AD. Researchers have to go through an application process in order to access ADNI data, which follow a strict framework of ethical guidelines. All participants were informed and provided consent for

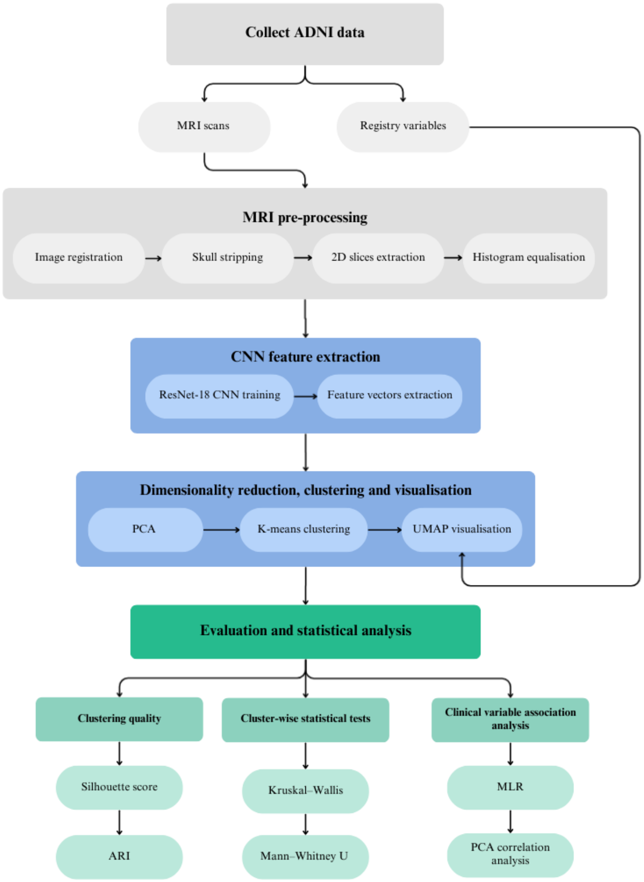

Figure 3.1: Methodological pipeline of the research process.

the use of their data for research purposes, and the shared data are fully anonymised, ensuring that no personally identifiable information is accessible to researchers.

The ADNI-1 collection consists of T-1 weighted volumetric 3D MRI scans acquired at 1.5T. There are 639 subjects and a total of 2294 images in the dataset, where each image has a diagnosis label of CN, MCI or AD. In the set, there are 476 AD scans, 1113 MCI scans, and 705 CN scans. The scans were collected within a one-year time span. Each image has the NIfTI data format.

3.2.2 Registry variables

Five external clinical variables were selected from the ADNI registry and manually extracted for this study. These variables were merged into a single

CSV file containing patient data.

- Diagnosis (CN, MCI, AD): Diagnosis represents the baseline clinical categorisation provided by ADNI and is used as a reference label for external validation.

- Genotype (APOE ε2, ε3, ε4 alleles): The APOE genotype is one of the most well-established genetic risk factors for AD and represents a biologically constant variable [9].

- Education (in years): Education is included as an environmental variable commonly used in AD research to study cognitive reserve [8]. It is measured as the number of years of formal education, where 12 years will generally refer to high school level and 16 years will refer to college level.

- Age (in years): Age is a major risk factor for AD and reflects both biological and environmental processes over time. Given that many symptoms relating to dementia can be mediated through age, it is important to analyse eventual dependencies [6].

- Sex (female/male): Sex is a biological variable relevant to AD research due to the differences in biological ageing processes between males and females [7].

The selected registry variables represent a combination of biological and environmental factors, both constant and time-dependent, and were included to explore their association with disease progression and potentially discover meaningful phenotypes. The variables were merged into the CSV file using pandas, which were then cleaned and fixed for missing values. In this study, the abbreviation “AD” is used both to denote Alzheimer’s disease as a condition and the AD diagnostic category within ADNI.

3.3 Pre-processing

3.3.1 ADNI corrections and data organisation

Prior to download, the ADNI dataset has undergone several standard image correction procedures [24]. These include GradWarp correction, B1 Non-Uniformity correction, and N3 Non-Uniformity correction. The application of these corrections depends on the scanner manufacturer and radiofrequency coil configuration used when acquiring the MRI.

GradWarp correction is applied to correct geometric distortions caused by gradient nonlinearity in MRI scanners. B1 Non-Uniformity correction normalizes the intensity across the images if there has been a problem with variability in the signal intensity during scanning. Finally, N3 Non-Uniformity correction is used for histogram peak sharpening and is applied to all images after the previous corrections. Additionally, the image intensities are standardised to ensure that they are scaled to a consistent range.

Before the manual pre-processing was performed, the dataset was organised by sorting the MRI scans into directories corresponding to their diagnostic labels, and were then re-named for simplicity [25].

3.3.2 MRI pre-processing

In this work, the pre-processing pipeline follows the approach proposed by Ammari [26]. Here, the FMRIB Software Library (FSL) [27] provides tools for processing and analysis of imaging data, and has been utilised for the initial steps.

Since brain anatomy varies across individuals, image registration is commonly performed to spatially align images to a common coordinate system. As a first step, affine transformations were applied to linearly align each MRI scan to a reference template to limit variations in brain size, shape, orientation, and position [2]. The Montreal Neurological Institute (MNI) template [28] was used as the reference and is shown in Figure 3.2.

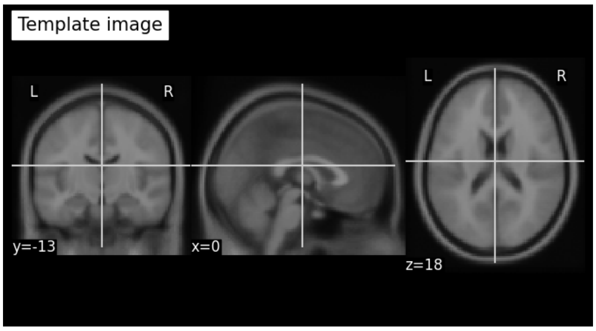

Figure 3.2: MNI template.

Following registration, skull stripping was performed to extract brain tissue to remove unnecessary information. The brain mask derived from the

MNI template was applied to each image to extract the relevant brain regions. The result of these pre-processing steps is shown in Figure 3.3.

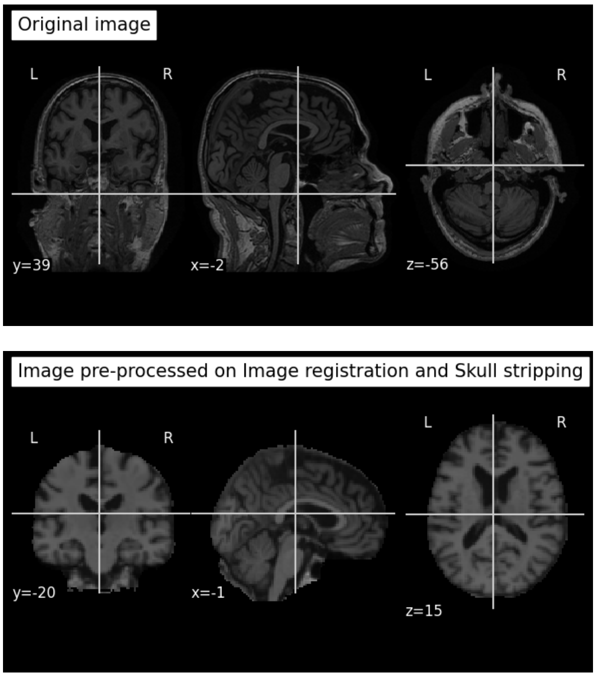
Original image
L
R
L
R
y=39
x=-2
z=-56
Image pre-processed on Image registration and Skull stripping
L
R
L
R
y=-20
x=-1
z=15

Figure 3.3: Example of an MRI scan before (top) and after (bottom) image registration and skull stripping.

Using 2D slices instead of full 3D MRI volumes reduces network complexity and computational requirements, which is beneficial when computational resources are limited. However, this approach comes at the cost of losing spatial dependencies between adjacent slices, which highlights the importance of choosing the most informative slices [2]. Typically, central regions of the brain contain more structural information than the edges.

To identify the most informative slices, Shannon entropy was computed for each axial slice, as the axial plane is the most commonly used view in 2D MRI studies [29]. Entropy was computed from the intensity histogram of each slice, where higher entropy values indicate greater intensity variation and detail, and are therefore considered more informative. The 16 slices with the highest entropy values were selected from each scan and arranged into a 4×4 grid to form a single composite 2D image.

Finally, histogram equalisation was applied to enhance the contrast of each image. An example of a final pre-processed image can be seen in Figure 3.4.

To ensure that the pre-processing pipeline preserved relevant information, a small-scale CNN was trained on a simple classification task. This validation

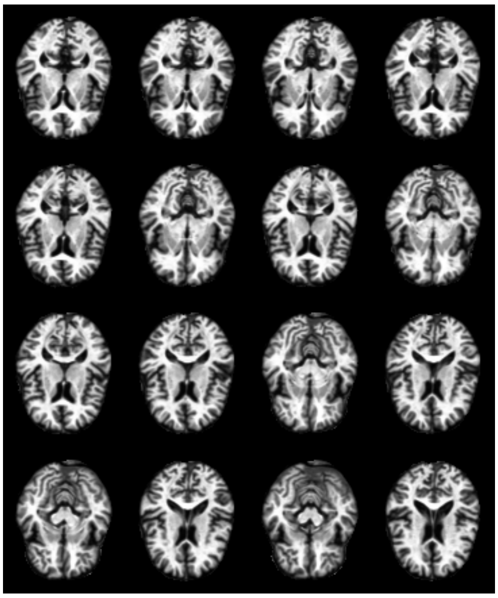

Figure 3.4: Final pre-processed MRI image.

step served as a sanity check to confirm that the pre-processing did not negatively affect the data and that it resulted in improved performance compared to using raw images.

3.3.3 Dataset preparations

The dataset was divided into an 80/20 stratified split for training and validation, where the stratification ensured equal representation of each diagnostic class in both splits. As there was a slight class imbalance in the data, class weights were computed to be used in the loss function during training to prevent the model from being biased toward the most frequent class.

Additionally, to improve generalisation and reduce overfitting, subtle augmentation was applied to the training images after pre-processing, where small affine transformations were used to simulate minor variations in the images.

3.4 Model architecture

3.4.1 ResNet-18 CNN

The model used in this work was based on a 2D ResNet-18 architecture, which is a well-established CNN consisting of 18 layers and originally pre-trained on the ImageNet dataset to utilise transfer learning [5]. It was selected as a balance between model capacity and computational efficiency, and due to ResNet-18 containing fewer parameters compared to deeper residual networks, the risk of overfitting is reduced when training on relatively limited medical imaging datasets such as ADNI.

To adapt the model to grayscale MRI, the first convolutional layer was modified to accept single-channel inputs. Furthermore, the original fully connected layer was replaced with a new classifier head consisting of a dropout layer with $p = 0.2$ to reduce overfitting, followed by a linear layer with three output units corresponding to the three diagnostic classes.

After experimenting with different hyperparameter values, the final model was trained using the Adam optimiser with a learning rate of 1×10^{-4}. A weighted cross-entropy loss function was used to account for class imbalance, where the weights were inversely proportional to the class frequencies in the dataset. The model was trained for a maximum of 50 epochs with early stopping applied when the validation loss failed to improve for eight consecutive epochs. The batch size was set to 32.

Training performance was monitored across epochs by tracking both accuracy and loss for the training and validation sets. The model with the lowest validation loss was saved as the final model.

3.4.2 Feature extraction

Once CNN training was completed, feature extraction was performed on the best model. The classification head of the model was removed, and a 512-dimensional feature vector was extracted for each image from the final activation layer. The feature representations were stored alongside the corresponding clinical variables for unsupervised learning and analysis.

Using the trained CNN as a feature extractor rather than as a classifier enables exploration of the learned feature space without enforcing predefined diagnostic labels. Even though the CNN is trained in a supervised manner and the learned representations are therefore shaped by diagnostic labels, unsupervised clustering and analysis are performed to explore and potentially reveal additional structure and phenotypic variability beyond the diagnostic boundaries. This approach allows the discovery of potential imaging-derived phenotypes associated with AD progression in a high-dimensional feature space.

3.5 Dimensionality reduction, clustering and visualisation

3.5.1 Principal component analysis

Principal Component Analysis (PCA) was applied to reduce the dimensionality of the high-dimensional feature vectors extracted from the CNN. PCA transforms the original feature space into a set of orthogonal principal components (PCs) to decrease computational complexity and extract the features that retained the maximum variance in the data.

To determine an optimal number of PCs, the cumulative explained variance was plotted as a function of the number of components, where a threshold between 85% and 90% explained variance was considered [30], as shown in Figure 3.5. Based on this criterion, 30 PCs were selected, representing the smallest number of components required to explain more than 85% of the total variance in the data.

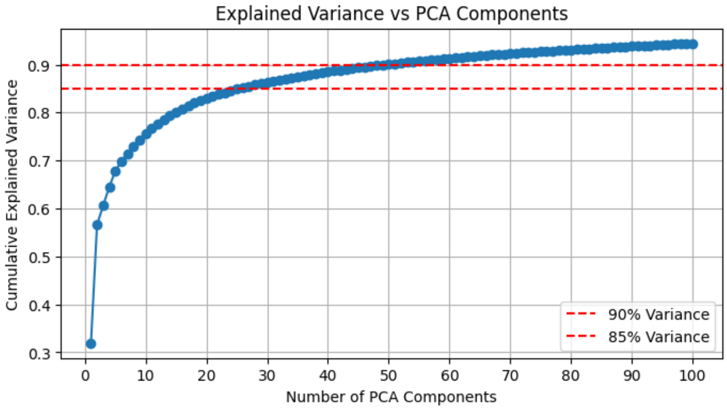

Figure 3.5: Cumulative explained variance as a function of the number of PCs.

3.5.2 Clustering algorithm

Clustering algorithms are a class of unsupervised learning methods used to identify structure in unlabeled data by grouping similar data points together [31]. In this study, clustering was applied to the PCA-reduced feature representations to explore the structure within the learned feature space.

K-means clustering was selected as the primary clustering algorithm. K-means partitions the data into K clusters by minimising the variance within clusters and is widely used due to its simplicity and computational efficiency. Gaussian Mixture Models (GMMs) were also considered as an alternative clustering approach, as they can model more complex cluster shapes. However, comparison of the two methods using Silhouette scores (see Section 3.6.1.1) and visual inspection of the resulting clusters revealed minimal differences in performance. However, K-means achieved a slightly higher Silhouette score and was therefore selected for the final analysis.

The number of clusters, K, was determined using the Within-Cluster Sum of Squares (WCSS) criterion [32]. The optimal value of K was identified using the Elbow method, where a noticeable change in the WCSS curve indicates a balanced number of clusters. Based on this analysis, *K=3* was selected as the optimal number of clusters, as shown in Figure 3.6.

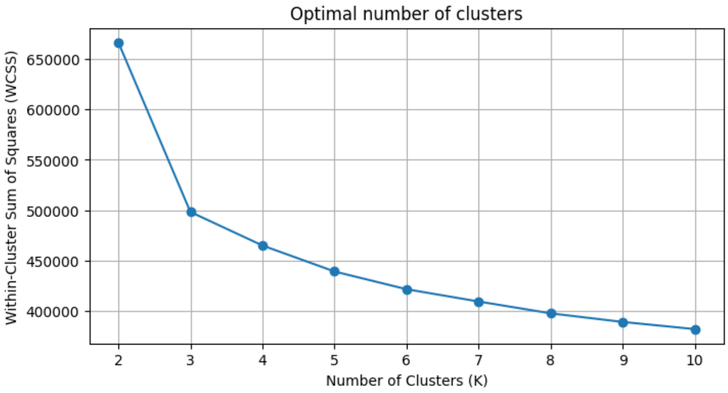

Figure 3.6: WCSS curve illustrating the Elbow method for cluster selection.

3.5.3 UMAP visualisation

Uniform Manifold Approximation and Projection (UMAP) is a dimensionality reduction technique commonly used to visualise high-dimensional data in a low-dimensional space [33]. Compared to techniques such as t-Distributed

Stochastic Neighbor Embedding (t-SNE), UMAP is computationally efficient and better at preserving both local neighbourhood structure and broader global relationships in the data.

UMAP was applied to the PCA-reduced feature representations to visualise the structure of the learned feature space. The cluster assignments obtained from K-means were overlaid onto UMAP to visualise the cluster structure and separation.

The UMAP hyperparameters *n_neighbors* and *min_dist* control the balance between local and global structure in the visualisation space. Multiple parameter combinations were evaluated by assessing cluster quality both visually and quantitatively using the Silhouette score to determine appropriate values.

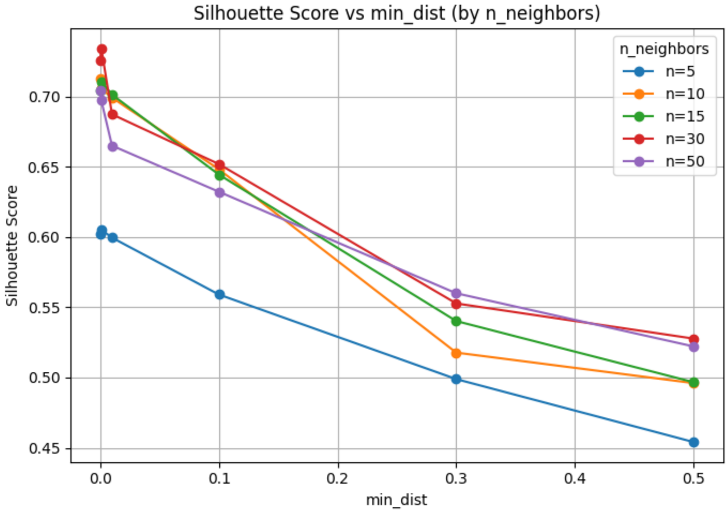

Figure 3.7: Silhouette score for different UMAP parameter configurations.

The highest Silhouette score was achieved with *n_neighbors=30* and *min_dist=0.0*, as shown in Figure 3.7. Using this configuration, separate UMAP plots were generated for each clinical variable, where data points were coloured according to the corresponding variable values for a qualitative interpretation of the clustering structure.

It should be noted that distances between points and clusters in UMAP only approximate relationships in the original high-dimensional feature space. UMAP is primarily used here as a visualisation tool, and conclusions about cluster separation are supported by complementary quantitative analyses.

3.6 Statistical evaluation and analysis

3.6.1 Clustering quality

To evaluate the quality of the clustering results, both internal and external validation metrics were implemented. Specifically, the Silhouette score and the Adjusted Rand Index (ARI) were used to assess how well the clustering algorithm grouped the data.

3.6.1.1 Silhouette score

The Silhouette score is an internal clustering validation metric that measures how similar a data point is to its assigned cluster compared to other clusters [32]. The score ranges from -1 to 1, where higher values indicate that data points are well matched to their own cluster and well separated from neighbouring clusters.

For a given data point i, the Silhouette score is computed as follows:

- $a(i)$: the average distance between i and all other data points within the same cluster.
- $b(i)$: the average distance between i and all data points in the nearest neighbouring cluster.

The Silhouette score for data point i is then defined as:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \tag{3.1}$$

The overall Silhouette score for the clustering is obtained by averaging $s(i)$ across all data points.

3.6.1.2 Adjusted Rand Index

ARI is an external clustering evaluation metric used to compute the agreement between two different partitions of the same dataset [34]. In contrast to internal metrics such as the Silhouette score, ARI compares the clustering results to a set of reference labels while correcting for agreement that may occur by chance. In this study, one partition corresponds to the three cluster assignments obtained from K-means clustering, while the other partition corresponds to the clinical diagnostic labels CN, MCI, and AD provided by the ADNI dataset.

The formula for ARI is based on the Rand Index R, which computes the proportion of data point pairs that are assigned to either the same group or different groups in both partitions among all possible pairs of data points. The ARI then adjusts this measure by accounting for the expected R under random cluster assignment, which is denoted by E. The maximum possible Rand Index is $\max(R) = 1$. Therefore, the ARI is defined as:

$$ARI = \frac{R - E}{\max(R) - E} \tag{3.2}$$

The ARI score ranges from -1 to 1, where a value of 1 indicates perfect agreement between the clustering and the reference labels, values close to 0 indicate agreement comparable to random assignment, and negative values indicate an agreement worse than random.

3.6.2 Clusterwise statistical tests

Clusterwise statistical tests were performed to evaluate whether the clusters differed significantly with respect to external clinical variables. Specifically, the Kruskal–Wallis test was used to determine if there were significant differences across clusters, and the Mann-Whitney U Test was used as a post-hoc analysis to assess which specific clusters show significant differences. For all statistical tests, exact p-values are reported when available. For values that were numerically shown as 0, the results are reported as $p < 10^{-16}$.

3.6.2.1 Kruskal–Wallis test

The Kruskal-Wallis H-Test is a non-parametric statistical hypothesis test used to determine whether there are statistically significant differences between multiple groups with respect to a single variable [35]. In this study, the Kruskal–Wallis test was used to evaluate whether the distributions of each clinical variable differed significantly across the three clusters. The test returns an H-statistic that measures differences between group distributions and a p-value to assess statistical significance.

The hypotheses are defined as:

- H_0: The clinical variable has the same distribution across all clusters.
- H_1: The clinical variable differs in distribution across at least one cluster.

If the resulting p-value falls below the chosen significance threshold, the null hypothesis H_0 is rejected, indicating that at least one cluster differs significantly. The significance level was set to the commonly used threshold of $p < 0.05$.

3.6.2.2 Post-hoc Mann–Whitney U Test

The Mann-Whitney U test is a non-parametric statistical hypothesis test used as a post-hoc analysis when the Kruskal-Wallis test has indicated that at least one cluster differs significantly from the others with respect to a clinical variables [36]. The test is then used to determine which specific pairs of clusters differed significantly.

For each pairwise comparison, the hypotheses were defined as:

- H_0: The clinical variable has the same distribution in the two clusters.
- H_1: The clinical variable differs in distribution between the two clusters.

The null hypothesis H_0 was rejected if the p-value was below the threshold of $p < 0.05$.

3.6.3 Clinical variable association analysis

Multinomial Logistic Regression (MLR) and PCA correlation analysis were used to explore relationships between clinical variables and the learned feature representations.

3.6.3.1 Multinomial Logistic Regression

MLR is a parametric statistical model that describes the relationship between independent variables and an outcome with multiple categories [37]. In this study, MLR was used to assess how well the clinical variables could explain cluster membership and to identify variables that were independently associated with the clustering structure.

MLR probabilities are computed through the log-odds of each outcome cluster relative to a baseline cluster. The model was specified as:

$$\text{Cluster} \sim \text{Diagnosis} + \text{Education} + \text{Genotype} + \text{Age} + \text{Sex} \tag{3.3}$$

A variable was considered to significantly influence cluster membership if its associated p-value was below the threshold of $p < 0.05$. The process was repeated with different baseline clusters so that all pairwise differences between the clusters could be captured.

3.6.3.2 MLR assumptions

Before performing MLR, the assumptions of the model were evaluated to ensure the validity of the analysis [38].

First, MLR assumes independence of observations, and in this study, each MRI scan was treated as an independent observation.

Second, the dependent variable must be categorical and unordered. This assumption is satisfied, as cluster membership represents mutually exclusive categories obtained through unsupervised clustering and does not imply an inherent ordering.

Third, MLR assumes no multicollinearity, stating that there should be no correlations between the independent variables. The Variance Inflation Factor (VIF) was computed for each variable to ensure that this assumption holds, and was backed up by a correlation matrix of the predictors.

Finally, the assumption of linearity between the predictors and the log-odds of the outcome was evaluated using scatterplots to ensure that the relationship between the predictors and the log-odds was approximately linear.

3.6.3.3 PCA correlation analysis

After the PCA dimensionality reduction, a correlation analysis was performed to investigate how individual PCs relate to the clinical variables [39]. In doing this, it can be explored whether the features space captures clinically relevant information and how variance associated with different variables is distributed across the latent dimensions.

The non-parametric Spearman's rank correlation was computed between the first ten PCs, which together account for the majority of the variance, and each clinical variable [40]. Spearman's correlation coefficient, ρ, and the corresponding p-value were used to assess the strength and significance of these associations, thus identifying which components were most strongly correlated with each variable.

3.6.4 Visualisation techniques

Visualisation methods were used to increase the interpretability and as a qualitative evaluation of the clustering results and learned feature representations, complementing the quantitative analyses.

UMAP cluster plots were used to project the high-dimensional feature space into 2D for visual interpretation. These plots served as the primary visualisation tool and qualitative evaluation method for exploring the clustering structure and for assessing how the learned feature space aligned with clinical variables.

Additionally, distribution plots were generated to visualise how clinical variables were distributed across the three diagnostic groups. Kernel density estimation (KDE) plots were used for continuous variables, such as education and age, while bar plots were used for categorical variables, such as genotype and sex. All plots were normalised to ensure fair comparison of distributions independently of group size.

Finally, feature maps were extracted from the early convolutional layers of the CNN to provide qualitative insight into its learning process. Visualising these feature maps highlights how the CNN responds to different patterns in the MRI slices and provides insight into the hierarchical feature extraction performed by the model.

3.6.5 Evaluation and analysis method choices

Due to the high dimensionality and complexity of the feature representations, multiple methods were used to ensure a comprehensive and robust interpretation of the learned feature space. Each method addresses a different aspect of the analysis, and together they provide a deeper understanding that would not be achievable through any single approach alone.

Cluster quality was evaluated through both qualitative and quantitative methods. UMAP visualisations served as the primary qualitative tool for exploring the structure of the feature space. Simultaneously, the Silhouette score provided an internal quantitative evaluation of cluster groupings, while ARI offered an external quantitative evaluation of the agreement between the clusters and reference labels. Together, these metrics provided complementary perspectives on cluster compactness, separation, and agreement with diagnostic categories.

To evaluate differences in clinical variables across clusters, non-parametric univariate statistical tests were applied. The Kruskal–Wallis test was used to identify whether the distribution of a clinical variable differed across clusters, and the Mann–Whitney U test was implemented as a post-hoc analysis to determine which specific cluster pairs showed significant differences. However, univariate tests do not account for dependencies between variables. To address this limitation, MLR was included as a parametric multivariate analysis to investigate whether the variables can independently predict cluster membership while controlling for the influence of other variables. Consequently, apparent associations observed in the univariate tests can be evaluated in a multivariate setting to determine whether the relationships are direct or mediated through other factors.

Finally, while the other tests assess the relationships between variables, PCA correlation analysis provides insight into how variance associated with the clinical variables is embedded within the learned representation. This analysis complements clustering results by revealing which variables shape the structure of the high-dimensional feature space.

Overall, the combination of qualitative visualisation, quantitative clustering metrics, univariate and multivariate statistical testing, and feature space correlation analysis provides a robust and multidimensional evaluation framework.

3.7 Experimental setup

All experiments were conducted on an Apple MacBook Pro equipped with an M1 Pro processor and 16 GB of RAM. The development environment used was Visual Studio Code, and all experiments were conducted on the CPU.

All code was written in Python. The pre-processing pipeline utilised tools from the FSL library. Model development and training were performed using PyTorch, while NumPy and Pandas were used for data handling. Scikit-learn and SciPy were used for statistical analysis and clustering. Matplotlib, Seaborn, and UMAP-learn were used for visualisation and plotting.

Chapter 4

Results

4.1 Model training performance

The CNN was trained for a maximum of 50 epochs with early stopping based on validation loss. Training and validation loss and accuracy across epochs are shown in Figure 4.1.

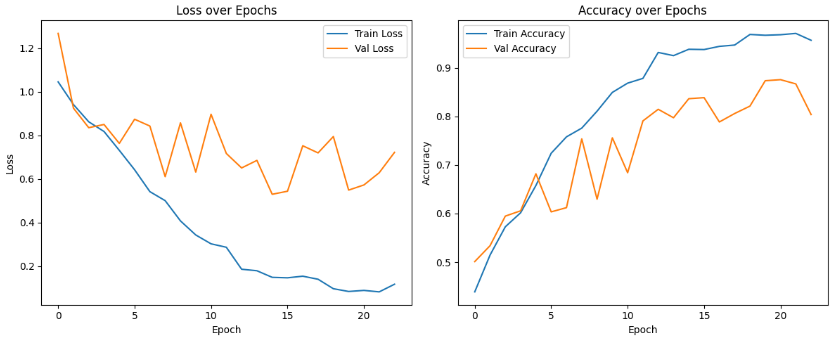

Figure 4.1: CNN training performance for the first 23 epochs.

Training stopped after 23 epochs due to early stopping. As shown in Figure 4.1, the training loss decreased steadily throughout training, while the validation loss had greater variability. Validation accuracy increased over epochs, although with some fluctuations. The lowest validation loss was observed at epoch 15, with a value of 0.5296, corresponding to a validation accuracy of 83.66%. At this epoch, the training accuracy was 93.84% with a training loss of 0.1487.

Although a gap between training and validation performance was observed, early stopping, data augmentation, and dropout helped to limit overfitting. Overall, the steadily decreasing training loss and improving validation accuracy suggests that the model was able to extract relevant and meaningful features from the MRI, and generalise well to unseen data.

4.2 Feature maps for CNN activations

Feature maps from selected convolutional layers of the CNN were visualised to examine how spatial features were captured during training.

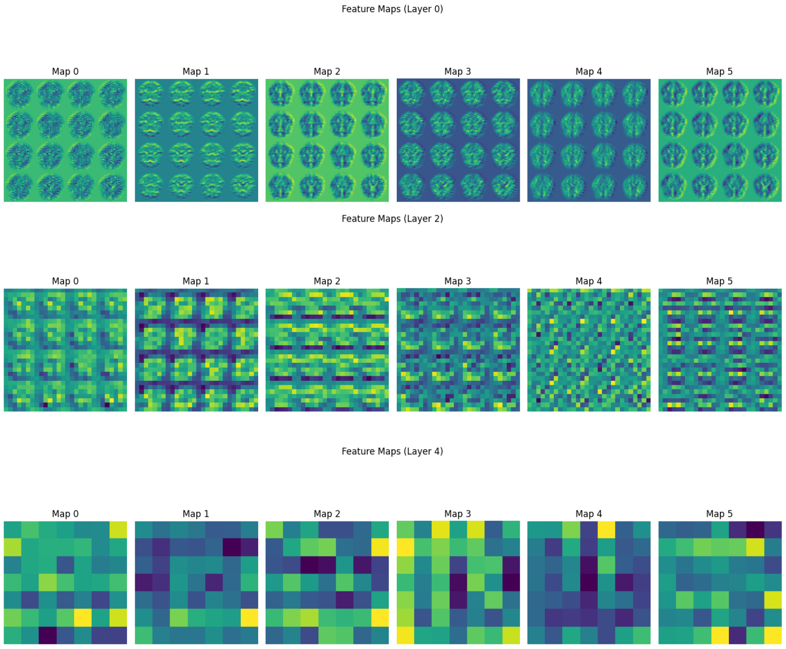

Figure 4.2: Feature maps extracted from convolutional layers 0, 2, and 4 of the trained CNN.

Figure 4.2 shows feature maps from early, intermediate, and deeper convolutional layers of the CNN. In the initial convolutional layer, layer 0, the activations primarily reflect low-level structural features, such as edges and contours of brain anatomy. In layer 2, the feature maps capture more spatially localised patterns, indicating increasingly complex representations. By layer 4, the activations appear more abstract, reflecting higher-level feature representations. The progression of these feature maps highlights the CNN's capacity to learn increasingly complex imaging features from MRI, which may reflect disease progression.

4.3 Distribution plots

To characterise the three diagnostic groups before any modelling or clustering, distribution plots were generated from the baseline ADNI data. These plots visualise the distributions of clinical variables across the diagnostic groups AD, MCI, and CN, and provide an intuitive overview of differences in the dataset.

4.3.1 Education

Figure 4.3 shows the distribution of education levels across diagnostic groups. The AD group is skewed toward fewer years of education, with a peak around 12 years, corresponding approximately to a high school education. In contrast, the CN and MCI groups show higher education levels, both peaking around 16 years, which typically corresponds to a college education. The MCI group follows a similar distribution to the CN group but displays a slightly broader spread.

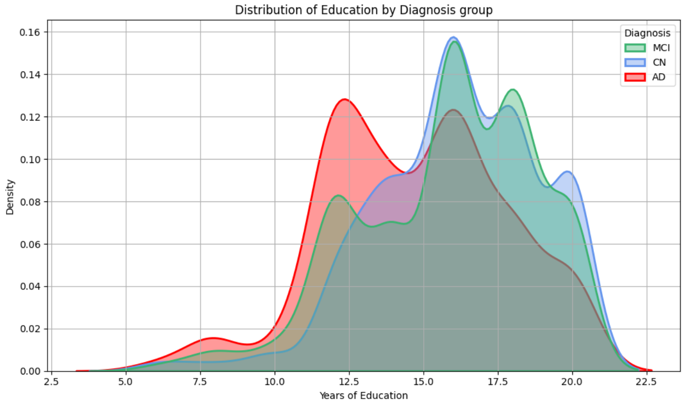

Figure 4.3: Distribution of education levels by diagnostic groups.

4.3.2 Genotype

Figure 4.4 shows the distribution of APOE genotype variants across the diagnostic groups. The largest differences are observed between the AD and CN groups. The AD group is mainly characterised by containing ε4 genotypes, with peaks at the 3/4 and 4/4 combinations. In contrast, the CN group shows higher proportions of the ε2 genotype, with peaks at 2/3 and 2/4.

The MCI group overlaps with both AD and CN, showing a distribution with peaks at 2/4 and 3/4, reflecting its transitional state between the two diagnostic groups.

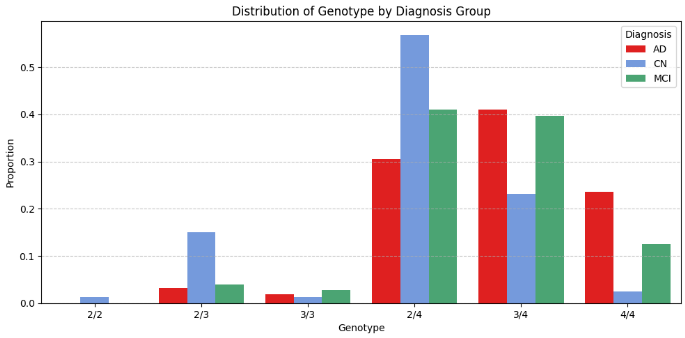

Figure 4.4: Distribution of genotype by diagnostic groups.

4.3.3 Age

Figure 4.5 shows the age distributions across the diagnostic groups. The distributions are highly similar, with all groups peaking between approximately 70 and 80 years of age. Notably, the CN group has a slightly narrower age range compared to the MCI and AD groups.

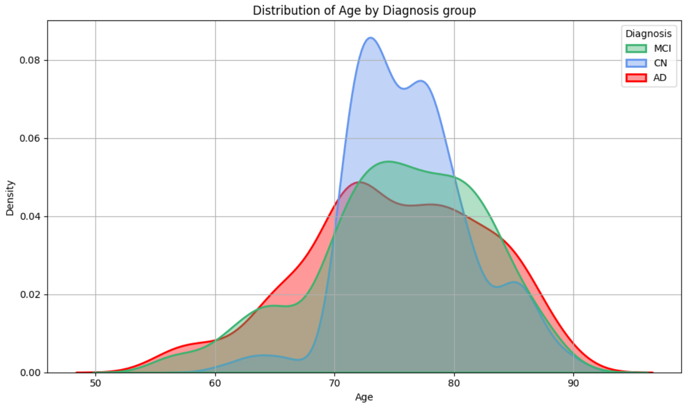

Figure 4.5: Distribution of age by diagnostic groups.

4.3.4 Sex

Figure 4.6 illustrates the proportion of female and male participants within each diagnostic group. The CN and AD groups show relatively balanced sex distributions, while the MCI group displays a noticeably higher proportion of male patients. Therefore, within this dataset, male patients are more strongly represented in the MCI group compared to the other diagnostic categories.

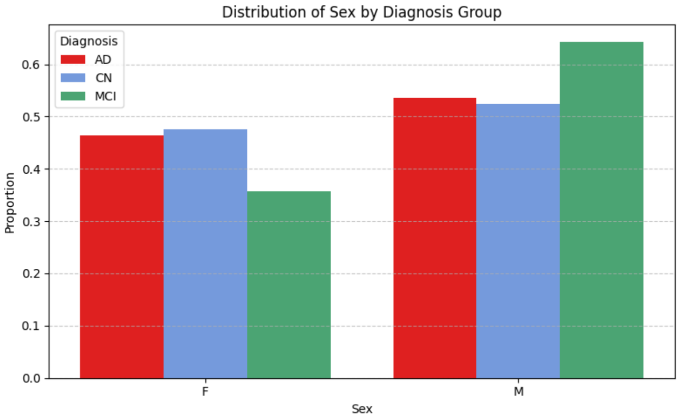

Figure 4.6: Distribution of sex by diagnostic groups.

4.4 Unsupervised clustering of the CNN feature space

The UMAP visualisations of the K-means clustering results are shown in Figures 4.7 – 4.12, using UMAP hyperparameters set to *n_neighbors = 30* and *min_dist = 0.0*.

4.4.1 Unsupervised clusters

Figure 4.7 presents the UMAP visualisation of the CNN feature space coloured by K-means cluster assignment. Three distinct clusters are observed, with cluster centres marked by “×”. The clusters are well separated in the low-dimensional space, with minimal overlap between groups.

Clusters 0 and 1 show relatively compact structures, indicating low variability within each cluster. In contrast, Cluster 2 appears more diffuse, suggesting a higher intra-cluster variability. A small number of data points

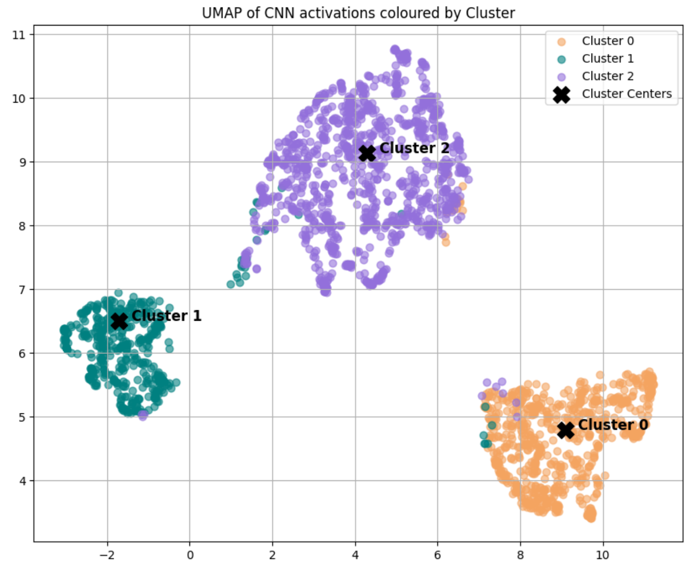

Figure 4.7: UMAP visualisation of CNN activations coloured by cluster assignment.

from other clusters are observed near the outer regions of Cluster 2 and Cluster 0, while Clusters 1 show minimal overlap with other groups.

Overall, the clustering results indicate that the feature space contains a well-defined structure that can be separated into distinct groups using unsupervised learning.

4.4.2 Diagnosis

To investigate the relationship between the unsupervised clusters and clinical diagnosis, the same UMAP visualisation is coloured by diagnostic labels AD, CN, and MCI, as seen in Figure 4.8. A strong overlap is observed between the clusters and diagnostic categories. Specifically, Cluster 0 aligns with the CN group, Cluster 1 aligns with the AD group, and Cluster 2 with the MCI group.

However, the alignment is not perfect. Within each cluster, some data points are coloured with alternative diagnostic labels, indicating overlap between diagnostic groups in the learned feature space. This is particularly evident within the AD-aligned cluster, which contains a small number of CN and MCI data points.

Although clustering was performed without access to diagnostic labels, the resulting alignment indicates that the feature representations capture structure

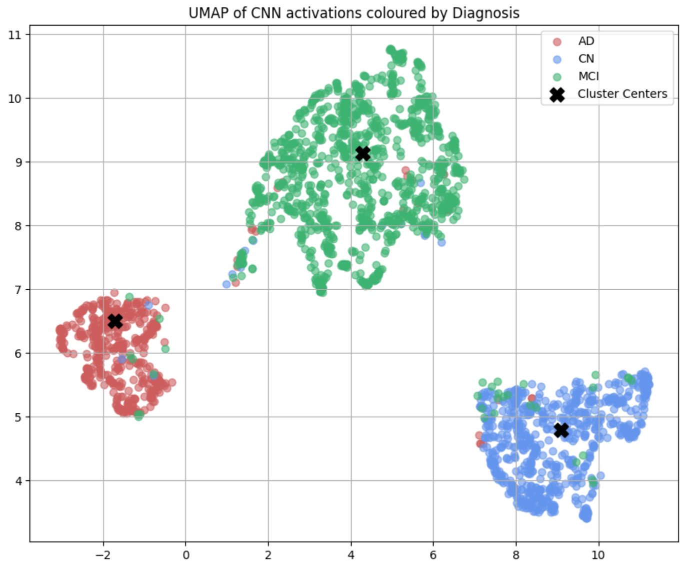

Figure 4.8: UMAP visualisation of CNN activations coloured by diagnosis.

closely associated with clinical diagnostic categories.

4.4.3 Education

Figure 4.9 shows the UMAP visualisation coloured by years of education, where lighter colours correspond to higher educational levels. The distribution appears relatively scattered across the feature space, however, a gradual trend can be observed across clusters. The AD-aligned cluster contains a higher proportion of data points associated with fewer years of education, while the CN-aligned cluster contains comparably higher education values. The MCI-aligned cluster shows a broader distribution, containing patients with both lower and higher levels of education. Although the clusters are not distinctly separated by education, these patterns suggest that some educational differences are reflected within the learned feature space.

4.4.4 Genotype

Figure 4.10 shows the UMAP visualisation coloured by APOE genotype variants. Six genotype combinations are represented, with lighter colours indicating ε4 genotypes, which are associated with increased risk of AD. The AD-aligned cluster shows a distinct distribution of ε4 genotypes, while

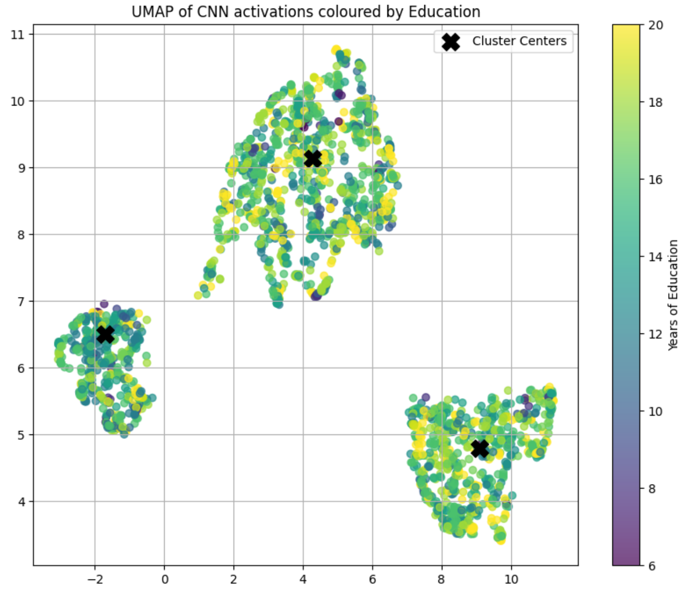

Figure 4.9: UMAP visualisation of CNN activations coloured by years of education.

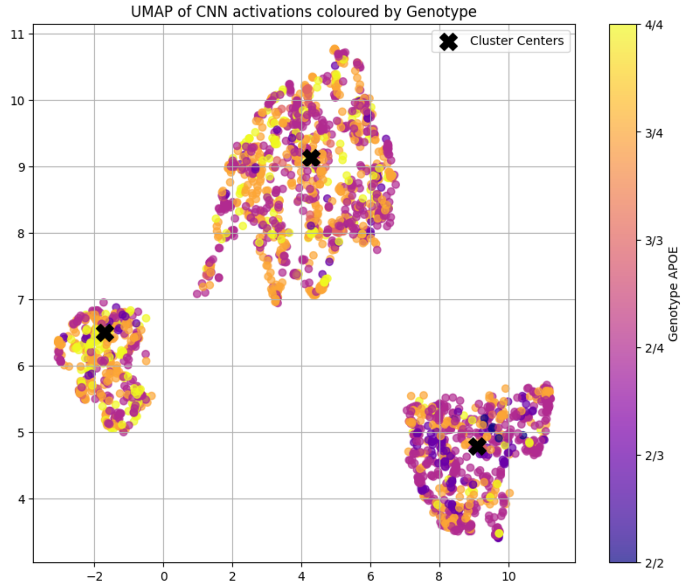

Figure 4.10: UMAP visualisation of CNN activations coloured by genotype variant.

the CN-aligned cluster primarily contains ε2 variants, which are associated with a reduced risk of AD. The MCI-aligned cluster shows an intermediate distribution, containing a mixture of ε2, ε3, and ε4 genotypes. These patterns indicate a strong alignment between genetic risk variants and the cluster structure.

4.4.5 Age

Figure 4.11 shows the UMAP visualisation coloured by age, ranging from approximately 55 to 90 years. There appears to be a broad distribution of ages across the learned feature space, with no clear separation or gradient corresponding to the cluster structure. The CN-aligned cluster shows a slightly narrower pool of ages around 70–80 years, while the AD- and MCI-aligned clusters cover broader age ranges. Overall, these patterns indicate that age-related information is only weakly reflected in the feature space and does not appear to be a dominant factor shaping the clustering structure.

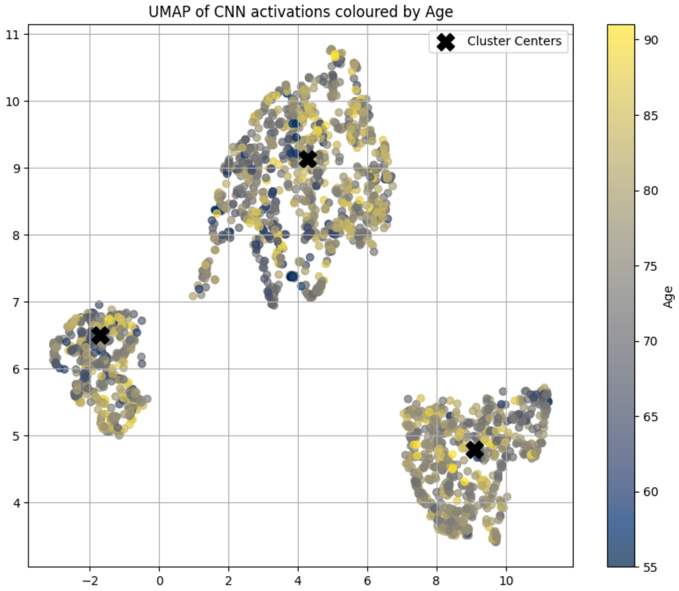

Figure 4.11: UMAP visualisation of CNN activations coloured by age.

4.4.6 Sex

Figure 4.12 shows the UMAP visualisation coloured by sex. Female and male data points largely overlap across all clusters, with no clear separation. A slight pattern is observed where the CN- and AD-aligned clusters contain a higher proportion of female patients, while the MCI-aligned cluster shows a higher proportion of male participants. However, this pattern is very subtle and the overall distribution suggests that sex is not a strong determinant of the clustering structure.

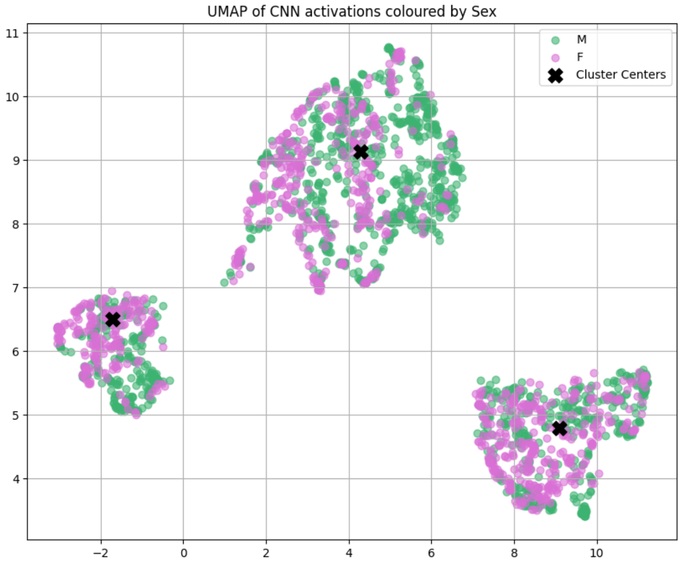

Figure 4.12: UMAP visualisation of CNN activations coloured by sex.

4.5 Clustering quality evaluation

Clustering quality was evaluated using the Silhouette score and ARI to assess both the internal structure of the clustering and its agreement with clinical diagnostic labels.

4.5.1 Silhouette score

The Silhouette score was computed for the clusters derived from the feature space. The resulting average Silhouette score was 0.323, indicating a moderate degree of separation between the clusters. This suggests that data points are on average closer to their assigned cluster than to neighbouring clusters, although with some overlap between clusters. Overall, the score supports the presence of a meaningful clustering structure in the learned feature space, with moderate compactness and separation.

4.5.2 ARI result

The ARI was computed to evaluate the agreement between the unsupervised cluster assignments and the clinical diagnostic categories CN, MCI, and AD. The resulting ARI score was 0.889, which indicates a very strong agreement between the cluster structure and diagnostic groups. These results suggest that the clusters derived from the feature space effectively capture patterns relevant to disease progression.

4.6 Clusterwise statistical tests

Clusterwise differences were evaluated using the Kruskal–Wallis test, followed by the post-hoc Mann–Whitney U test.

4.6.1 Kruskal–Wallis test results

The Kruskal–Wallis H-test was performed to assess whether the distributions of clinical variables differed significantly across the clusters.

Table 4.1: Kruskal–Wallis results for differences in clinical variables across clusters.

Variable	**H-value**	**p-value**
Diagnosis	2003.72	$p < 10^{-16}$
Education	67.51	2.19×10^{-15}
Genotype	255.72	2.96×10^{-56}
Age	5.52	0.063
Sex	28.03	8.19×10^{-7}

As shown in Table 4.1, all tested variables except age showed statistically significant differences between clusters. This indicates that age was relatively consistent between clusters, while diagnosis, genotype, education, and sex differed significantly. As expected, diagnosis showed an extremely strong effect, confirming the strong alignment between clusters and diagnostic groups observed in previous results. Genotype also showed a very strong association with cluster membership, while education and sex demonstrated comparatively weaker but statistically significant differences between clusters.

4.6.2 Post-hoc Mann–Whitney U Test results

Based on the Kruskal–Wallis test results, post-hoc pairwise cluster comparisons were performed using the Mann–Whitney U test for all variables except age.

Table 4.2: Mann–Whitney U test p-values for pairwise cluster comparisons.

Variable	**0 vs 1**	**0 vs 2**	**1 vs 2**
Diagnosis	1.66×10^{-196}	$p < 10^{-16}$	2.36×10^{-294}
Education	2.96×10^{-12}	0.184	3.09×10^{-9}
Genotype	4.21×10^{-44}	6.70×10^{-38}	2.63×10^{-6}
Sex	0.423	1.01×10^{-4}	1.56×10^{-5}

As seen in Table 4.2, the diagnosis variable showed extremely significant differences across all pairwise cluster comparisons. This further confirms the strong alignment between the unsupervised clusters and the diagnostic groups. Based on these consistent results, it can be concluded that Cluster 0 shows the strongest correspondence with CN, Cluster 1 with AD, and Cluster 2 with MCI.

For education, significant differences were observed between the CN- and AD-aligned clusters, as well as between the AD- and MCI-aligned clusters, while no significant difference was observed between the CN- and MCI-aligned clusters. This indicates that the AD-aligned cluster differs more strongly with respect to education.

The genotype variable showed highly significant differences between the CN-aligned cluster and both other groups, while a weaker but still significant difference was observed between the AD- and MCI-aligned clusters. This indicates that the CN-aligned cluster is genetically more distinct from the other two clusters than AD and MCI are from each other.

For the sex variable, significant differences were found between the CN- and MCI-aligned clusters and between the AD- and MCI-aligned clusters, while no significant difference was observed between the CN- and AD-aligned clusters. This suggests that sex-related differences were primarily associated with the MCI-aligned cluster.

Overall, the post-hoc analysis demonstrated that several clinical variables differed significantly across cluster pairs. The strength of these differences varied, indicating that some clinical characteristics had a stronger influence on cluster structure than others.

4.7 Clinical variable association analysis

Association analyses were performed to investigate how individual clinical variables align with the learned feature space and to identify which latent dimensions capture clinically relevant information.

4.7.1 MLR results

MLR was performed to assess which clinical variables independently predicted cluster membership while controlling for the influence of all variables together. The resulting p-values for each pairwise cluster comparison are shown in Table 4.3.

Table 4.3: MLR p-values for pairwise cluster comparisons for all variables.

Predictor	**0 vs 1**	**0 vs 2**	**1 vs 2**
Diagnosis	$p < 10^{-16}$	$p < 10^{-16}$	$p < 10^{-16}$
Education	0.066	0.149	0.720
Genotype	0.001	$p < 10^{-16}$	0.871
Age	0.164	0.217	0.060
Sex	0.626	0.448	0.375

As expected, diagnosis was highly significant across all cluster comparisons, reinforcing its dominant role in shaping the clustering structure. Genotype was also a significant predictor when comparing the CN-aligned cluster with both the AD- and MCI-aligned clusters, but no longer showed a significant effect when comparing the AD- and MCI-aligned clusters.

In contrast, education, age, and sex did not show statistically significant effects in any of the pairwise comparisons when controlling for all other variables. Education and age approached significance in some comparisons, but their effects were considerably weaker once diagnosis and genotype were accounted for.

Overall, these results show that diagnosis is the strongest independent predictor of cluster membership, while genotype mainly helps to distinguish CN from disease states. Other clinical variables contribute more indirectly and do not independently predict cluster membership. This highlights the importance of multivariable analysis, as associations seen in univariate analyses may be mediated by more dominant variables.

4.7.2 PCA correlation analysis

To investigate how clinical variables aligned with the learned feature space, Spearman rank correlations were computed between each variable and the top ten PCs explaining the most variance.

The strongest correlation was observed between diagnosis and the first PC ($\rho = -0.836, p < 10^{-16}$), indicating that the dominant axis of variation in the feature space primarily reflects disease progression. Genotype also showed a moderate correlation with PC1 ($\rho = -0.310$, $p = 2.01 \times 10^{-52}$), suggesting that genetic risk partially overlaps with the same latent dimension. Education also displayed a weaker but significant correlation with PC1 ($\rho = 0.138$, $p = 3.65 \times 10^{-11}$), indicating a small association with the diagnostic structure.

Age showed its strongest correlation with PC4 ($\rho = 0.232$, $p = 1.64 \times 10^{-29}$), suggesting that age-related variation is captured along a separate latent dimension of the feature space. Sex displayed only weak correlations with later PCs (PC8–PC10), indicating that sex-related effects are represented in more subtle, higher-dimensional ways in the feature space.

Overall, these results indicate that diagnosis and genotype primarily shape the dominant structure of the feature space, while education contributes more weakly. In contrast, age and sex influence the learned representation but do not meaningfully shape the clustering structure.

Chapter 5

Discussion

5.1 Overview and main findings

This study explored whether unsupervised clustering of the CNN feature representations could reveal clinically meaningful patterns related to AD progression beyond traditional diagnostic classification. By combining deep learning for feature extraction with unsupervised clustering and statistical analysis, these results indicate that the feature representations capture both genetic and environmental influences embedded in brain structure and are grouped based on phenotypic variability. Overall, the findings support the use of unsupervised deep learning approaches for discovering neuroimaging-based phenotype structures in AD research.

5.2 Structure of unsupervised clusters

Unsupervised K-means clustering of the feature representations resulted in three visually well-separated clusters that showed strong alignment with the clinical diagnostic categories CN, MCI, and AD. This alignment was supported by a high ARI, moderate Silhouette score, and by highly significant differences in diagnosis across all cluster pairs in both univariate and multivariate analyses. This clustering structure was further supported by the PCA correlation analysis, which showed that the dominant PC primarily reflected diagnostic variation. Although the CNN was trained in a supervised manner, the unsupervised clustering was performed independently of diagnostic labels, indicating that the learned feature representations capture meaningful disease-related patterns from MRI data.

The unsupervised clusters also managed to capture the typical intra-cluster behaviour of each diagnostic group. While the CN- and AD-aligned clusters were compact with data points close to their centres, the MCI-aligned cluster displayed more intra-cluster variability. Due to MCI being a transitional state and therefore having more variability in its disease trajectory, these results align with known clinical expectations.

Although there is a strong agreement between the unsupervised cluster assignments and diagnostic groups, it is not a perfect alignment. In the UMAP visualisations, the unsupervised clusters appear more homogeneous than the clinically defined categories. These findings suggest that unsupervised clustering of feature representations may group patients more consistently according to their underlying brain-based phenotypes rather than traditional diagnostic groupings, which can potentially reveal more natural and meaningful subgroups within the AD trajectory.

5.3 Genetic influences on the feature space

Genotype appeared as one of the strongest variables reflected in the feature space. The UMAP visualisation showed strong separation between clusters, where the distribution of alleles aligned with clinical expectations. Both the univariate statistical tests and the multivariate analysis showed that genotype significantly distinguished the CN-aligned cluster from both the AD- and MCI-aligned clusters. The PCA correlation analysis further showed that genotype was moderately aligned with the dominant PC associated with disease progression. This suggests that genetic risk factors primarily separate CN individuals from those along the disease trajectory, which aligns well with existing literature stating that APOE ε4 is strongly associated with increased risk of developing AD.

The fact that genotype-related differences were reflected in the feature space highlights the CNN's ability to capture genetically relevant brain patterns from MRI data without explicit genetic input. This finding further supports the potential of deep learning and unsupervised neuroimaging approaches to discover biologically meaningful phenotypes embedded in brain structure.

5.4 Educational effects and cognitive reserve

Throughout the analyses, education showed consistent but moderate associations with the clustering structure. The UMAP visualisation showed gradual differences across clusters, which reflected the distribution plot, where the CN group was centered around higher education levels, while the AD group was skewed toward fewer years of education. The univariate analyses revealed significant differences in education, mainly when comparing the AD-aligned cluster to the CN- and MCI-aligned clusters. These findings are consistent with the concept of cognitive reserve, which states that higher educational accomplishment can delay or even mitigate neurodegeneration. However, education did not remain a significant independent predictor in the multivariate analysis once diagnosis and genotype were accounted for. The weak but significant correlation between education and the primary PC further supported the idea that education partially overlaps with disease-related variation in the feature space.

These results suggest that education contributes indirectly to the clustering structure, likely through its relationship with diagnosis and genetic risk rather than acting as an independent determinant. Additionally, the subtle influence of education on the feature representations indicates that the protective effect of cognitive reserve may be reflected in brain patterns.

5.5 Age and sex as indirect influences on the feature space

Although age is a well established risk factor for AD, it revealed relatively weak associations to the clustering structure. Age did not significantly differ across clusters in univariate tests and did not independently predict cluster membership in the MLR analysis. However, PCA correlation analysis showed that age was moderately associated with a later PC, suggesting that age-related variation is captured along a separate dimension of the feature space.

Similarly, sex showed only weak effects across the analysis. The sex distribution plot showed that the MCI group had a higher proportion of males, and a significant difference was found between the MCI-aligned cluster and the other two clusters in the univariate tests. However, sex did not remain a significant predictor in the MLR analysis and was only weakly correlated with the later PCs. The discrepancy between findings in the literature, which states that there is a higher proportion of females with AD compared to males, and the relatively balanced sex distribution observed in this study likely reflects demographic imbalances in the dataset rather than sex-related differences.

Overall, these results indicate that age and sex influence the feature space in subtle, high-dimensional ways, but do not directly shape the clustering structure.

5.6 Complementary analytical approaches

An important strength of this study lies in the combination of complementary analytical methods. Visualisation techniques such as UMAP provided qualitative insight into the structure of the learned feature space, while quantitative clustering metrics computed objective measures of cluster quality and alignment with clinical labels. Univariate non-parametric tests identified distributional differences across clusters, while multivariate MLR clarified which variables independently contributed to cluster membership. Finally, PCA correlation analysis provided insight into how clinical variables were embedded within the latent dimensions of the feature space.

The contrast between univariate and multivariate results was particularly informative. Variables such as education and sex showed apparent associations with clustering in univariate analyses but lost significance in the multivariate model, which revealed how their effects are mediated through more dominant variables like diagnosis and genotype. This highlights the importance of using multiple analytical perspectives when interpreting high-dimensional data, as each method captured a different aspect of the relationship between the feature space and the clinical data, and revealed new information that would otherwise be hidden.

Chapter 6

Conclusions

6.1 Conclusions

This study investigated whether unsupervised clustering of the CNN feature representations derived from MRI could reveal clinically meaningful patterns related to AD progression. The results demonstrated that unsupervised clustering of the feature representations recovered a clinically meaningful structure that aligned closely with diagnostic categories, indicating that the learned feature space encodes disease-relevant information from MRI data.

Diagnosis proved to be the dominant factor shaping the feature space, with genotype providing a biologically meaningful contribution, particularly in distinguishing CN individuals from disease groups. Education showed a moderate influence that was largely mediated through diagnosis and genotype, with effects of cognitive reserve. In contrast, age and sex displayed weak associations, influencing the feature space in more subtle ways.

Although the unsupervised clusters aligned closely with the diagnostic groupings, they also showed differences in compactness and overlap. This suggests that the feature space groups patients more consistently based on captured phenotypic variability that is not fully reflected by traditional diagnostic labels.

Overall, these findings confirm that the feature space captures both genetic and environmental influences embedded in brain structure. By combining deep learning with unsupervised clustering and complementary analyses, this study moves beyond diagnostic classification toward a deeper exploration of AD heterogeneity. The results address the three objectives of the research aim and support the potential of unsupervised deep learning approaches as a valuable tool for neuroimaging-based phenotype discovery in AD research.

6.2 Limitations and Future work

While this exploratory study provides promising results for the use of unsupervised deep learning to investigate AD progression, several limitations should be acknowledged.

Firstly, although the ADNI dataset is widely used in AD research, it represents a relatively narrow demographic as the participants are primarily from the United States and Canada. Therefore, the findings may not generalise well to diverse populations. Future work should therefore include datasets with broader demographic and clinical diversity, as well as variations when acquiring the MRI data.

Additionally, the number of clinical variables included in this study was limited. Incorporating a wider range of variables could provide deeper insight into what influences disease progression. Future work can also experiment with additional biomarkers, such as Cerebrospinal Fluid (CSF) or Positron Emission Tomography (PET) scans, as well as longitudinal data to better understand disease trajectories over time.

Another limitation is the potential for data leakage. The ADNI dataset includes multiple visits from some participants, and each MRI scan was treated as an independent observation, which may introduce bias. However, clustering was performed on feature representations rather than raw images, and the focus was on analysis rather than supervised classification, making this an acceptable assumption for the exploratory nature of the study. Future work should consider constructing datasets with independent patient samples or explicitly modelling longitudinal relationships to strengthen the validity of the findings.

Furthermore, the conversion of 3D MRI volumes into 2D composite slices likely reduced spatial information, which may limit the CNN’s ability to capture finer anatomical patterns. Using a full volumetric 3D CNN could therefore reveal more informative feature representations, though at the cost of increased computational complexity. Future work could also explore alternative architectures beyond ResNet to further improve feature quality.

Finally, alternative clustering strategies and references could be explored. For example, the diagnostic groupings could be expanded by separating MCI into stable and progressive subtypes, or clustering could be performed and evaluated without using diagnostic categories as reference labels. By exploring these factors, the model has the potential to reveal novel subgroups and expand the discovery of phenotypes within the AD disease trajectory.

References

[1] Alzheimer's Association, "What is alzheimer's disease?" 2025. [Online]. Available: https://www.alz.org/alzheimers-dementia/what-is-alzheimers

[2] M. A. Ebrahimighahnavieh, S. Luo, and R. Chiong, "Deep learning to detect Alzheimer's disease from neuroimaging: A systematic literature review," *Computer Methods and Programs in Biomedicine*, vol. 187, p. 105242, 2020. doi: https://doi.org/10.1016/j.cmpb.2019.105242. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169260719310946

[3] Alzheimer's Disease Neuroimaging Initiative, "Sharing alzheimer's research data with the world," 2024. [Online]. Available: https://adni.loni.usc.edu

[4] R. Yamashita, M. Nishio, R. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights into Imaging*, vol. 9, Jun. 2018. doi: 10.1007/s13244-018-0639-9

[5] 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*, 2016, pp. 770–778.

[6] Alzheimer's Research UK, "Dementia and ageing," 2024. [Online]. Available: https://www.alzheimersresearchuk.org/dementia-information/dementia-risk/dementia-and-ageing

[7] Alzheimer's Society, "Why is dementia different for women?" 2025. [Online]. Available: https://www.alzheimersresearchuk.org/dementia-information/dementia-risk/dementia-and-ageing

[8] Alzheimer's Research UK, "Cognitive reserve and dementia risk," 2024. [Online]. Available: https://www.alzheimersresearchuk.org/dementia-information/dementia-risk/cognitive-reserve-and-dementia-risk

[9] C.-C. Liu, T. Kanekiyo, H. Xu, and G. Bu, "Apolipoprotein E and Alzheimer disease: Risk, mechanisms and therapy," *Nature reviews. Neurology*, vol. 9, Jan. 2013. doi: 10.1038/nrneurol.2012.263

[10] National Human Genome Research Institute, "Phenotype," 2025. [Online]. Available: https://www.genome.gov/genetics-glossary/Phenotype

[11] A. Payan and G. Montana, "Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks," *CoRR*, vol. abs/1502.02506, 2015, arXiv: 1502.02506. [Online]. Available: http://arxiv.org/abs/1502.02506

[12] E. Hosseini-Asl, R. Keynton, and A. El-Baz, "Alzheimer's disease diagnostics by adaptation of 3D convolutional network," in *2016 IEEE International Conference on Image Processing (ICIP)*, 2016. doi: 10.1109/ICIP.2016.7532332 pp. 126–130.

[13] S. S. Kundaram and K. C. Pathak, "Deep Learning-Based Alzheimer Disease Detection," in *Proceedings of the Fourth International Conference on Microelectronics, Computing and Communication Systems*, V. Nath and J. K. Mandal, Eds. Singapore: Springer Singapore, 2021. ISBN 978-981-15-5546-6 pp. 587–597.

[14] W. Salehi, P. Baglat, A. Upadhya, B. Sharma, and G. Gupta, "A CNN Model: Earlier Diagnosis and Classification of Alzheimer Disease using MRI," Nov. 2020.

[15] M. Swapna, Y. K. Sharma, and B. Prasadh, "CNN Architectures: Alex Net, Le Net, VGG, Google Net, Res Net," *International Journal of Recent Technology and Engineering*, vol. 8, no. 6, pp. 953–960, 2020.

[16] S. Sarraf and G. Tofighi, "Classification of Alzheimer's Disease Structural MRI Data by Deep Learning Convolutional Neural Networks," *CoRR*, vol. abs/1607.06583, 2016, arXiv: 1607.06583. [Online]. Available: http://arxiv.org/abs/1607.06583

[17] 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), 2017. doi: 10.1109/IST.2017.8261460 pp. 1–6.

[18] G. I. Stoleru and A. Iftene, "Transfer Learning for Alzheimer's Disease Diagnosis from MRI Slices: A Comparative Study of Deep Learning Models," *Procedia Computer Science*, vol. 225, pp. 2614–2623, 2023. doi: https://doi.org/10.1016/j.procs.2023.10.253. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050923014114

[19] S. Basaia, F. Agosta, L. Wagner, E. Canu, G. Magnani, R. Santangelo, and M. Filippi, "Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks," *NeuroImage: Clinical*, vol. 21, p. 101645, 2019. doi: https://doi.org/10.1016/j.nicl.2018.101645. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2213158218303930

[20] H. A. Helaly, M. Badawy, and A. Y. Haikal, "Deep Learning Approach for Early Detection of Alzheimer's Disease," *Cognitive Computation*, vol. 14, no. 5, pp. 1711–1727, Sep. 2022. doi: 10.1007/s12559-021-09946-2. [Online]. Available: https://doi.org/10.1007/s12559-021-09946-2

[21] J. Venugopalan, L. Tong, H. R. Hassanzadeh, and M. Wang, "Multimodal deep learning models for early detection of Alzheimer's disease stage," *Scientific Reports*, vol. 11, p. 3254, Feb. 2021. doi: 10.1038/s41598-020-74399-w

[22] M. Estarellas, N. Oxtoby, J. Schott, D. Alexander, and A. Young, "Multimodal subtypes identified in Alzheimer's Disease Neuroimaging Initiative participants by missing-data-enabled subtype and stage inference," *Brain Communications*, vol. 6, Jun. 2024. doi: 10.1093/braincomms/fcae219

[23] X. Bi, S. Li, B. Xiao, Y. Li, G. Wang, and X. Ma, "Computer aided Alzheimer's disease diagnosis by an unsupervised deep learning technology," *Neurocomputing*, vol. 392, pp. 296–304, 2020. doi: https://doi.org/10.1016/j.neucom.2018.11.111. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231219304709

[24] Alzheimer's Disease Neuroimaging Initiative, "MRI Pre-processing," 2024. [Online]. Available: https://adni.loni.usc.edu/data-samples/adni-data/neuroimaging/mri/mri-pre-processing/

[25] M. F. Amin, "Organize ADNI," Feb. 2024. [Online]. Available: https://github.com/FahimFBA/organize-ADNI

[26] Y. Ammari, "PFE alzheimer disease classification," 2022, GitHub repository. [Online]. Available: https://github.com/yacineammari/PFE-alzheimer-disease-classification

[27] FMRIB Software Library, "FSL," 2025. [Online]. Available: https://fsl.fmrib.ox.ac.uk/fsl

[28] NeuroImaging and Surgical Technologies Lab, "MNI ICBM152 non-linear," 2025. [Online]. Available: https://nist.mni.mcgill.ca/mni-icbm152-non-linear-6th-generation-symmetric-average-brain-stereotaxic-registration-model

[29] M. Hon and N. Khan, "Towards alzheimer's disease classification through transfer learning," 05 2023. doi: 10.32920/22734329.v1

[30] M. Greenacre, P. Groenen, T. Hastie, A. Iodice D'Enza, A. Markos, and E. Tuzhilina, "Principal component analysis," *Nature Reviews Methods Primers*, vol. 2, p. 100, 12 2022. doi: 10.1038/s43586-022-00184-w

[31] S. Chander and P. Vijaya, *Unsupervised learning methods for data clustering*, 01 2021, pp. 41–64. ISBN 9780128206010

[32] H. Belyadi and A. Haghighat, *Unsupervised machine learning: clustering algorithms*, 01 2021, pp. 125–168. ISBN 9780128219294

[33] L. McInnes, J. Healy, and J. Melville, "Umap: Uniform manifold approximation and projection for dimension reduction," *arXiv preprint arXiv:1802.03426*, 2018.

[34] J. E. Chacón and A. I. Rastrojo, "Minimum adjusted rand index for two clusterings of a given size," *Advances in Data Analysis and Classification*, vol. 17, no. 1, pp. 125–133, 2023.

[35] E. Ostertagova, O. Ostertag, and J. Kováč, "Methodology and application of the kruskal-wallis test," *Applied mechanics and materials*, vol. 611, pp. 115–120, 2014.

[36] Y. Cheng, W. Jia, R. Chi, and A. Li, "A clustering analysis method with high reliability based on wilcoxon-mann-whitney testing," *IEEE Access*, vol. 9, pp. 19 776–19 787, 2021. doi: 10.1109/ACCESS.2021.3053244

[37] Q. S. Cheng Hua, Dr. Youn-Jeng Choi, *Companion to BER 642: Advanced Regression*. Bookdown, 2021. [Online]. Available: https://bookdown.org/chua/ber642_advanced_regression/multinomial-logistic-regression.html

[38] J. Frost, "Multinomial logistic regression: Overview and example," 2025. [Online]. Available: https://statisticsbyjim.com/regression/multinomial-logistic-regression

[39] D. Granato, J. S. Santos, G. B. Escher, B. L. Ferreira, and R. M. Maggio, "Use of principal component analysis (PCA) and hierarchical cluster analysis (HCA) for multivariate association between bioactive compounds and functional properties in foods: A critical perspective," *Trends in Food Science & Technology*, vol. 72, pp. 83–90, 2018.

[40] V. Higgins, S. Hooshmand, and K. Adeli, "Principal component and correlation analysis of biochemical and endocrine markers in a healthy pediatric population (caliper)," *Clinical Biochemistry*, vol. 66, pp. 29–36, 2019.

TRITA – XXX-EX 2025:0000
Stockholm, Sweden 2026

www.kth.se

€€€€ For DIVA €€€€

{
”Author1”: { ”Last name”: ”Harmén”,
”First name”: ”Sofia”,
”E-mail”: ”sharme@kth.se”,
”organisation”: {”L1”: ”School of Electrical Engineering and Computer Science”,
}
},
”Cycle”: ”2”,
”Course code”: ”DA231X”,
”Credits”: ”30.0”,
”Degree1”: {”Educational program”: ”Master’s Programme, Computer Science, 120 credits”
,”programcode”: ”TCSCM”
,”Degree”: ”Degree of Master of Science in Engineering”
,”subjectArea”: ”Computer Science”
},
”Title”: {
”Main title”: ”An Unsupervised Exploration of the CNN Feature Space for Phenotype Structure Discovery in Alzheimer’s Disease”,
”Language”: ”eng” },
”Alternative title”: {
”Main title”: ”En oövervakad utforskning av CNN representationsrymd för upptäckt av fenotypisk struktur vid Alzheimers sjukdom”,
”Language”: ”swe”
},
”Supervisor1”: { ”Last name”: ”Andrzej Herman”,
”First name”: ”Pawel”,
”organisation”: {”L1”: ”School of Electrical Engineering and Computer Science”,
}
},
”Examiner1”: { ”Last name”: ”Kumar”,
”First name”: ”Arvind”,
”organisation”: {”L1”: ”School of Electrical Engineering and Computer Science”,
}
},
"National Subject Categories": "ddddd, ddddd",
"SDGs": "XXX, XXX",
”Other information”: {”Year”: ”2026”, ”Number of pages”: ”x,??”},
”Copyrightleft”: ”copyright”,
”Series”: { ”Title of series”: ”TRITA – XXX-EX” , ”No. in series”: ”2025:0000” },
”Opponents”: { ”Name”: ”XXX”},
”Presentation”: { ”Date”: ”2022-03-15 13:00”
,”Language”:”XXX”
,”Room”: ”via Zoom https://kth-se.zoom.us/j/ddddddddddd”
,”Address”: ”Isafjordsgatan 22 (Kistagången 16)”
,”City”: ”Stockholm” },
”Number of lang instances”: ”2”,
”Abstract[eng]”: €€€€

,

Alzheimers disease (AD) is a progressive neurodegenerative disorder and the leading cause of dementia worldwide. Early detection and understanding of the disease heterogeneity are crucial for improving the quality of life for patients. Even though convolutional neural networks (CNNs) have shown strong performance in classifying AD from structural MRI scans, most existing work focuses on supervised classification rather than exploring disease subtypes and associated phenotypes. Unsupervised learning can be used to explore such structures and can consequently help deepen our understanding of disease progression and support treatments tailored to individual patient needs. To address the issue, this exploratory study used unsupervised learning to investigate whether a CNN-learned feature space can reveal clinically meaningful patterns related to AD progression. A ResNet-18 CNN was trained on structural MRI data from the ADNI dataset, and the learned feature representations were analysed using unsupervised clustering and complementary statistical methods in relation to clinical variables. The results indicate that the CNN learned a structured and clinically meaningful feature space. The feature representations captured both genetic and environmental influences embedded in brain structure, and the unsupervised clustering recovered three distinct subgroups that aligned strongly with diagnostic categories but primarily reflected the underlying phenotypic variability. Overall, the findings show the potential of unsupervised deep learning approaches for discovering phenotypes from neuroimaging and deepen the understanding of AD heterogeneity.

”Abstract[swe]”: €€€€

Alzheimers sjukdom (AD) är en progressiv neurodegenerativ sjukdom och den vanligaste orsaken till demens i världen. En tidig upptäckt och bättre förståelse av sjukdomens heterogenitet är avgörande för att förbättra livskvaliteten hos patienter. Även om konvolutionella neurala nätverk (CNN) har visat starka resultat för klassificeringen av AD från strukturell MR-data, fokuserar majoriteten av tidigare studier på övervakad klassificering snarare än på att utforska sjukdomens undergrupper och fenotyper. Oövervakad inlärning kan användas för att analysera sådana strukturer, vilket kan bidra till en djupare förståelse av sjukdomens progression och stödja utvecklingen av mer individanpassade behandlingar. I denna studie användes oövervakad inlärning för att undersöka om CNN-baserade representationsrymder kan framhäva kliniskt meningsfulla mönster relaterade till progressionen av AD. En ResNet-18 CNN arkitektur tränades på strukturella MR-bilder från databasen ADNI och därefter

analyserades de inlärda representationerna med hjälp av oövervakad klustring och kompletterande statistiska metoder i relation till kliniska variabler. Resultaten visar att CNN-modellen lärde sig en strukturerad och kliniskt meningsfull representationsrymd. Representationerna fångade både genetiska och miljömässiga förändringar inbäddade i hjärnans struktur och den oövervakade klustringen identifierade tre distinkta undergrupper som hade en stark överensstämmelse med de diagnostiska kategorierna men som framförallt speglade en underliggande fenotypisk variation. Sammanfattningsvis visar dessa resultat potentialen hos oövervakade djupinlärningsmetoder för att upptäcka fenotyper från neuroavbildningar samt för att få en fördjupad förståelse av heterogeniteten hos AD.

€€€€,
”Keywords[eng]”: €€€€
Alzheimer’s disease, Magnetic resonance imaging, Convolutional neural networks, Deep learning, Unsupervised learning, Clustering, Phenotype discovery €€€€,
€€€€,
”Keywords[swe]”: €€€€
Alzheimers sjukdom, Magnetisk resonanstomografi, Konvolutionella neurala nätverk, Djupinlärning, Oövervakad inlärning, Klustring, Fenotypupptäckt €€€€,
}

acronyms.tex

```
%%% Local Variables:
%%% mode: latex
%%% TeX-master: t
%%% End:
\setabbreviationstyle[acronym]{long-short}

\newacronym{AD}{AD}{'Alzheimers disease}
\newacronym{ADNI}{ADNI}{'Alzheimers Disease Neuroimaging Initiative}
\newacronym{APOE}{APOE}{Apolipoprotein E}
\newacronym{ARI}{ARI}{Adjusted Rand Index}
\newacronym{CAD}{CAD}{-ComputerAided Diagnosis}
\newacronym{CN}{CN}{Cognitive Normal}
\newacronym{CNN}{CNN}{Convolutional Neural Network}
\newacronym{CSF}{CSF}{Cerebrospinal Fluid}
\newacronym{FSL}{FSL}{FMRIB Software Library}
\newacronym{GMM}{GMM}{Gaussian Mixture Model}
\newacronym{KDE}{KDE}{Kernel Density Estimation}
\newacronym{MCI}{MCI}{Mild Cognitive Impairment}
\newacronym{MLR}{MLR}{Multinomial Logistic Regression}
\newacronym{MNI}{MNI}{Montreal Neurological Institute}
\newacronym{MRI}{MRI}{Magnetic Resonance Imaging}
\newacronym{PC}{PC}{Principal Component}
\newacronym{PCA}{PCA}{Principal Component Analysis}
\newacronym{PET}{PET}{Positron Emission Tomography}
\newacronym{ResNet}{ResNet}{Residual Network}
\newacronym{tsne}{t-SNE}{-tDistributed Stochastic Neighbor Embedding}
\newacronym{UMAP}{UMAP}{Uniform Manifold Approximation and Projection}
\newacronym{VIF}{VIF}{Variance Inflation Factor}
\newacronym{WCSS}{WCSS}{Within-Cluster Sum of Squares}
```