UNIVERSITAT POLITÈCNICA DE VALÈNCIA

_ TELECOM ESCUELA TÉCNICA VLC SUPERIOR DE INGENIERÍA DE TELECOMUNICACIÓN

BRAIN CONNECTIVITY NETWORK ESTIMATION BASED ON CORRELATION AND SPECTRAL COHERENCE FROM FMRI TO SUPPORT ALZHEIMER DIAGNOSIS

Yuling Sang

Tutor: Dr. Addisson Salazar

Cotutor: Prof. Luis Vergara

Bachelor's Thesis presented at Escuela Técnica Superior de Ingenieros de Telecomunicación of Universitat Politècnica de València,

Academic Year 2019-2020

Valencia, 5 de March de 2020

Escuela Técnica Superior de Ingeniería de Telecomunicación
Universitat Politècnica de València
Edificio 4D. Camino de Vera, s/n, 46022 Valencia
Tel. +34 96 387 71 90, ext. 77190
www.etsit.upv.es

VLC/ CAMPUS
VALENCIA, INTERNATIONAL CAMPUS OF EXCELLENCE

EMAS
Gestión ambiental verificada
REG.NO. ES-CV- 000030

Resumen

La enfermedad de Alzheimer (EA) es una enfermedad neurológica degenerativa con un inicio desconocido y desarrollo progresivo. La enfermedad afecta principalmente al nervio central del cerebro de los ancianos, causando cambios degenerativos, lo que a su vez hace que los ancianos tengan un evidente deterioro de la memoria. Este proyecto realiza un análisis de datos de imágenes de resonancia magnética funcional, implementando la correlación y coherencia espectral, con el fin de proporcionar información útil para el diagnóstico precoz de la enfermedad de Alzheimer (EA). La red de conectividad cerebral ("Brain Connectivity Network", BCN) representa las formas de conexión entre diferentes regiones del cerebro, cuya estructura se alteraría de acuerdo con el estado de la enfermedad. Por lo tanto, la tarea principal de este proyecto es hacer una estimación de BCNs que podría reflejar la diferencia entre los pacientes con EA y el grupo de control normal. Se extraen varias características de las BCNs estimadas que se utilizan en la etapa de clasificación del procesado. Finalmente, la clasificación implementa el análisis discriminante lineal ("Linear Discriminant Analysis", LDA) y el análisis discriminante cuadrático ("Quadratic Discriminant Analysis", QDA). Los resultados se analizan utilizando la tasa de precisión y matriz de confusión de la clasificación.

Palabras clave: resonancia magnética funcional (fMRI), red de conectividad cerebral (BCN), correlación y coherencia espectral, LDA, QDA

Resum

La malaltia d'Alzheimer (EA) és una malaltia neurològica degenerativa amb un inici desconegut i desenvolupament progressiu. La malaltia afecta principalment al nervi central del cervell dels ancians, causant canvis degeneratius, la qual cosa al seu torn fa que els ancians tinguen una evident deterioració de la memòria. Aquest projecte realitza una anàlisi de dades d'imatges de ressonància magnètica funcional, implementant la correlació i coherència espectral, amb la finalitat de proporcionar informació útil per al diagnòstic precoç de la malaltia d'Alzheimer (EA). La xarxa de connectivitat cerebral ("Brain Connectivity Network", BCN) representa les formes de connexió entre diferents regions del cervell, l'estructura del qual s'alteraria d'acord amb l'estat de la malaltia. Per tant, la tasca principal d'aquest projecte és fer una estimació de BCNs que podria reflectir la diferència entre els pacients amb EA i el grup de control normal. S'extrauen diverses característiques de les BCNs estimades que s'utilitzen en l'etapa de classificació del processament. Finalment, la classificació implementa l'anàlisi discriminant lineal ("Linear Discriminant Analysis", LDA) i l'anàlisi discriminant quadràtica ("Quadratic Discriminant Analysis", QDA). Els resultats s'analitzen utilitzant la taxa de precisió i matriu de confusió de la classificació.

Palabras clave: resonancia magnética funcional (fMRI), red de conectividad cerebral (BCN), correlación y coherencia espectral, LDA, QDA

Abstract

Alzheimer's disease (AD) is a neurological degenerative disease with a concealed onset and progressive development. The disease mainly affects the central nerve of the brain of the elderly, causing degenerative changes, which in turn causes the elderly to have obvious memory decline. This project will make analysis of functional magnetic resonance imaging (fMRI) data, with the implementation of correlation and spectral coherence, in order to provide useful information for the early diagnosis of Alzheimer's disease (AD). Brain connectivity network (BCN) represents the forms of connection between different regions inside the brain, and its structure would alter according to the state of the disease. Thus, the primary task of this project is to make BCNs estimation which could reflect the difference between the patients with AD and normal control group. Several features would be extracted from these estimated BCNs for the classification step. Finally, linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) are two general automatic method we implemented for classification, and the results would be analyzed through accuracy rate, together with their confusion matrix.

Keywords: functional magnetic resonance imaging (fMRI), brain connectivity network (BCN), correlation and spectral coherence, LDA, QDA

Index

Capítulo 1. Introduction

With the gradual entry into an aging society, dementia of the elderly has received worldwide attention, and the most common form of it is Alzheimer 's disease(AD). This disease mainly affects the central nervous system of the elderly and causes degenerative changes. Although AD is quite common among the elderly in the world (50% to 60% of the total number of people with dementia), the rate of development of the disease is very slow, and it is often difficult to detect[1]. However, with the increasing age of the elderly, the clinical symptoms of AD gradually show up, such as memory decline, cognitive decline, and vague expressions, which make the elderly's normal life and social activities unable to proceed normally.

At present, the aging of the population has attracted widespread attention in various countries, and the incidence of AD has shown a rapid growth trend. The prevention and treatment of AD has caused widespread attention from governments and medical communities around the world. For instance, the Interdisciplinary Network for Dementia Utilizing Current Technology (INDUCT) is a collaborative research, composed of European Union and other leading organization, aims to make use of newest technologies to provide help to the patients and their carers. Also, the Alzheimer's Disease International(ADI) has always cooperated with the World Health Organization (WHO) to put the dementia patients at the centre of policies among the world[2].

The development of AD is a long and slow process. The lesions may begin decades before the clinical symptoms appear. The early clinical symptoms are mild and people tend to ignore them. Later in the disease, especially after dementia, treatment can be performed, although it can delay the process of cognitive decline, it cannot inhibit its development and reverse the existing damage. Therefore, the early diagnosis of AD is the key, and early intervention is of great significance, which means intervention before the occurrence of irreversible brain damage, to delay the progress of dementia or prevent the occurrence of dementia. The clinical diagnosis of AD patients currently mainly depends on clinical symptoms and neuropsychological tests, lacking objective biological markers[3]. Traditional Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) only reflect changes in the brain morphology of AD patients, and in the early stage, that is, before the obvious changes in brain atrophy, the physiological and functional metabolism of the brain area has changed[3]. Functional Magnetic Resonance Imaging (fMRI) provides an effective method to achieve functional imaging of the brain, by detecting blood oxygen level-dependent signals related to brain functional activity.

In our project, we proposed a method of constructing a brain functional network based on fMRI data. On the whole brain scale, the brain is divided into 116 functional regions, and each region is used as a node of the brain network. Then based on fMRI data which can express neuronal activity, we

calculated the correlation and coherence value between the standard brain region time series to measure the functional connection between the brain network nodes. After this, the network density was set, generating a series of threshold correlation and coherence matrix, to describe the brain functional network. Final step is using two general method, LDA and QDA, to finish the classification process, calculating the accuracy and making some comparison analysis.

The report includes six sections:

1. Introduction. This part will introduce the definition of Alzheimer's Disease and current method of diagnosis for this disease, followed by the overall process of our application of analysing functional magnetic resonance imaging data.
2. Objective of the thesis. This part includes the final result management method of our project.

   The Gantt Chart that reflects the task distribution and timetable is also included.
3. Background. This part will introduce the basic principle of the technologies used in our project, such as fMRI, together with the detailed illustration about the method we selected and implemented in each step of our project.
4. Design and implementation. This part will show the three main process of our project, including network estimation, feature extraction and classification. The implementations of method mentioned in background are clarified and test combinations are also included in this part.
5. Results and discussion. This part includes the results and analysis of results of each process. The connectivity matrix will be demonstrated and the network graph will also be shown in this part. The classification results include the accuracy and confusion matrix is included as well.
6. Conclusion and future work. This part will give the overall conclusion of our project and our results. Possible future work that will improve the performance is also mentioned.

Capítulo 2. Objectives of the thesis

2.1 Management

The primary goal of this project is to develop a system that is able to perform automatic diagnosis of AD patients from their fMRI images. There are three outcomes that should be measurable at the end of the project:

1. Software of the implementation of the brain connectivity networks based on correlation and spectral coherence
2. Software of the implementation of the feature extraction and general classification methods
3. Software for obtaining results: network diagrams, classification accuracy, confusion matrices

2.2 Task Distribution and Time Diagram

The detailed task distribution and according time diagram is shown below:

Work Plan (Gantt Chart)

	Nov 1-15	Nov 16-30	Dec 1-15	Dec 16-31	Jan 1-15	Jan 16-31	Feb 1-15	Feb 16-29	Mar 1-15	Mar 16-31	Apr 1-15	Apr 16-30
Task 1 [Study of the diagnosis of Alzheimer disease. Study of methods for estimating brain connectivity networks from fMRI based on correlation and spectral coherence: preprocessing, feature extraction, network estimation, and classifiers]												
Study of the diagnosis of Alzheimer's disease	X	X	X									
Study of the knowledge of functional magnetic resonance imaging (fMRI)	X	X	X	X	X	X	X					
Study brain connectivity network and method of estimating it			X	X	X	X	X					
Study of features of Alzheimer's disease and methods of classifying them						X	X					
Task 2 [Design and implementation of the procedures of network estimation, feature extraction, and classification (general methods: LDA, QDA, ...)]												
Design and implementation of the procedures of pre-processing						X						
Design and implementation of the procedures of network estimation						X	X	X				
Design and implementation of the procedures of feature extraction								X	X	X		
Design and implementation of the procedures of classification										X	X	
Task 3 [Experimentation: definition of the database; tuning and debugging of the methods; implementation of figures of merit.]												
Definition of the database						X						
Tuning and debugging of the methods						X	X	X	X	X	X	
Implementation of figures of merit						X	X	X	X	X	X	
Task 4 [Evaluation and reporting of the results]												
Evaluation and reporting on the network diagram						X	X	X				
Evaluation and reporting on the classification accuracy										X	X	
Evaluation and reporting on the confusion matrices						X	X	X	X	X	X	
Completing and revising the final report	X	X	X	X	X	X	X	X	X	X	X	X

Capítulo 3. Dataset and basic method

3.1 Functional Magnetic Resonance Imaging(fMRI)

Functional magnetic resonance imaging(fMRI) is applied to measure brain activity by detecting changes related to blood flow. The technology is based on the fact that cerebral blood flow and neuronal activation are combined. When a brain area is used, blood flow to that area also increases.

The generating principle behind functional MRI is the blood oxygenation level dependent (BOLD) contrast[4]. Oxy-hemoglobin, is a composite which is synthesized due to the respiration process, when oxygen is combined with the protein hemoglobin. Deoxy-hemoglobin is the same protein hemoglobin which is not bound to oxygen. The magnetic properties of blood are dependent on the oxygen levels: deoxy-hemoglobin suppresses the magnetic resonance signal while oxy-hemoglobin does not. When a region in our brain is more active, the inflow of blood which is oxygenated to that region would increase, which should have a negative influence on the deoxy-hemoglobin, resulting in the increase in the magnetic signal. The process diagram of fMRI can be shown as the figure below.

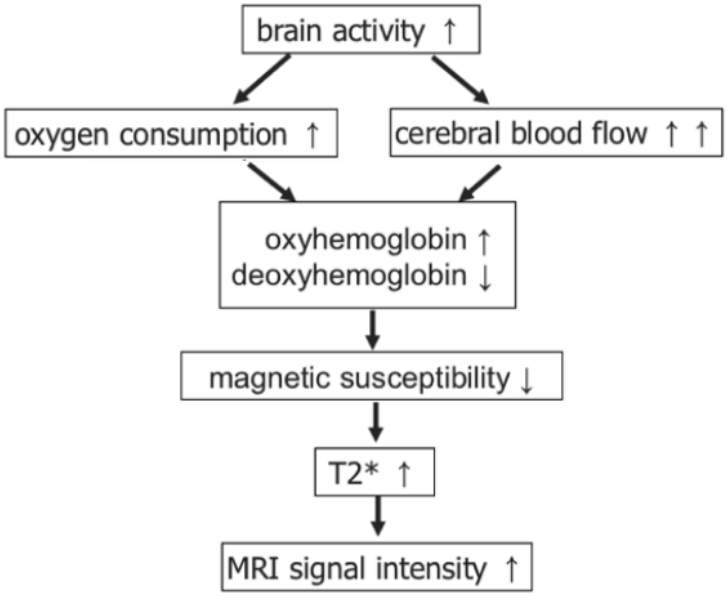

Figure 1 Process flowchart of functional MRI

3.2 Pre-processing method of fMRI

The pre-processing stage can be divided into four steps, including slice timing, realignment, normalization and smoothing[4]. The main goal of this stage is to minimize the negative effect during data collection process and physiological artefacts produced during fMRI generation. Also this stage is useful to verify statistical hypotheses and standardize the location of brain regions in different subjects, so that the accuracy and effectiveness would be increased in group analysis.

The detailed necessity of each step in pre-processing is shown below:

1. Slice timing: the assumption that the overall fMRI image of the scanned brain is measured simultaneously. However, in real operation, the 3D fMRI brain is composed of slices of the brain sequentially, which leads to temporal shift in similar time courses. The principle of slice timing is shown below[4]:

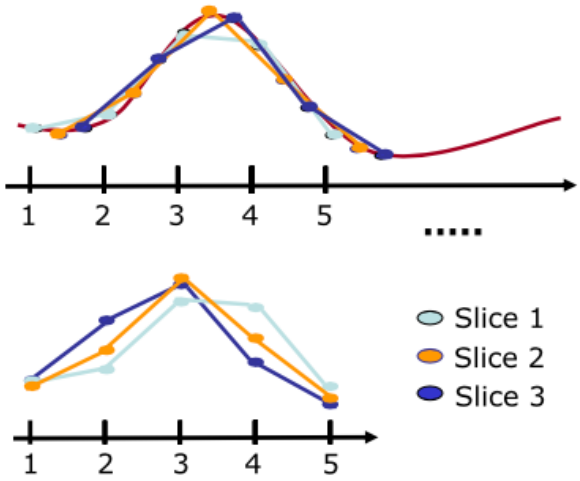

Figure 2 Principle of slice timing

In figure 2, the top line graph is the time series of three slices in fMRI data and the bottom one is the same three series after slice timing. As we can see from the top row, since different voxels are sampled at different times, what we need to do is making a shift to the time courses, so that these slices can be considered as scanned at the same time.

2. Realignment: The aim is primarily to remove movement artefact in fMRI time-series. It is important that a small movement could lead to significant error during further steps, therefore this step is necessary to estimate the head motion estimation and make related correction. The process of realignment is shown below[5]:

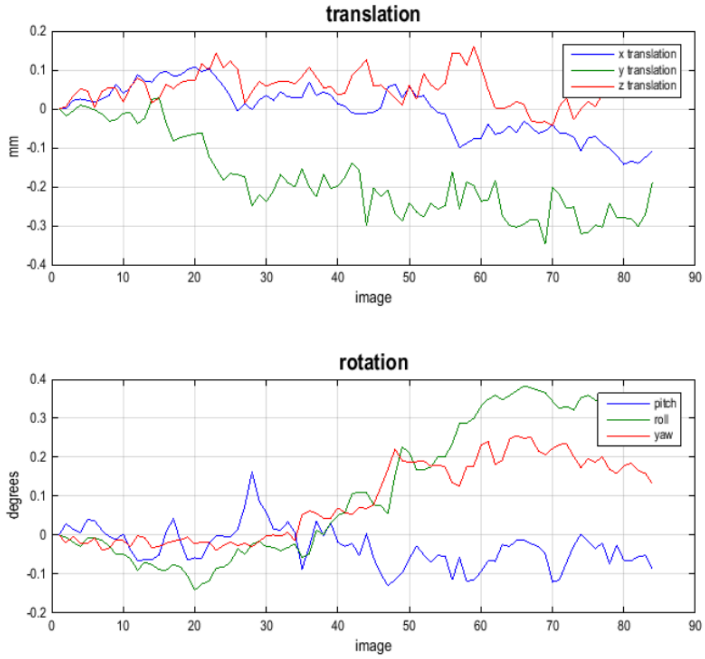

Figure 3 Realignment process, including translation and rotation

What we need to do is to find the alignment between each slice and reference slice, which is usually set as the first slice. We have six head dynamic parameter estimates, including three horizontal translation parameters and three axial rotation parameters, in order to align the brain voxels at different sampling time points of the same subject.

3. Normalization: Normalization means we provides a same brain structure for different individual subject. To realize this, we can warp the fMRI data of our test subject on a template brain structure[4]:

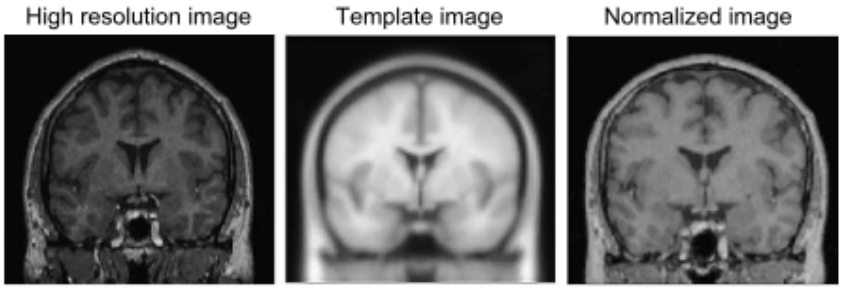

Figure 4 The process of normalization

With this step, the results are able to be utilized to a larger population and make it possible to compare the results across different studies and other subjects.

4. Smoothing: this step usually takes advantages of a Gaussian kernel to smooth the fMRI data. The figures below shows the data with and without smoothing step[5]:

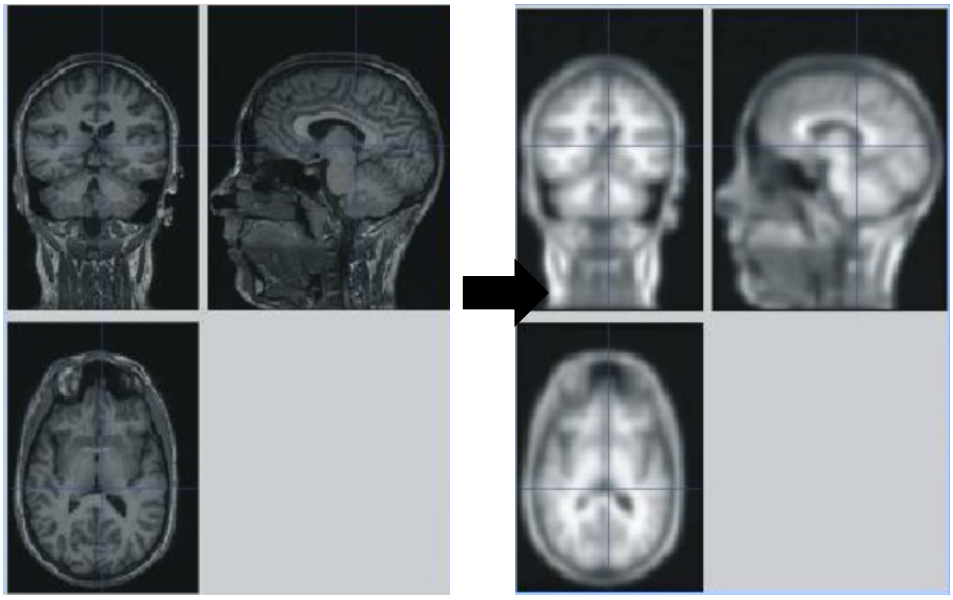

Figure 5 The process of smoothing(The left one is the figure without smoothing, the right one is the figure after smoothing)

With this step, we are able to reduce the remaining individual differences after normalization, as well as improve the signal-to-noise ratio.

3.3 Method of Network estimation

The network estimation is based on the analysis of the brain connectivity. Brain connectivity is based on three different connectivity: anatomical connectivity, functional connectivity, effective connectivity. fMRI data is better used to perform in the region of functional connectivity.

Functional Connectivity(FC)[6] means the temporal dependency between neuronal activation patterns. It is a statistical concept which can be seen as the correlation relationship between the oscillatory rate of tissue related to neuron, which we can deduce from fMRI.

The computational method this project selects for brain functional connectivity is known as knowledge-based method. This kind of method make selection between different regions of interest(ROI). And also it takes application of some matrix which is predefined such as correlation and coherence, to determine whether there is a connection between one ROI and another, so that we are supposed to generate the connectivity map of the brain.

3.3.1 Correlation

First, this metrics can be defined by correlation analysis. Pearson's linear correlation coefficient[6] ρ_{qs} is the most commonly used linear correlation coefficient.

$$\rho_{qs} = \frac{\sigma_{qs}(\tau)}{\sqrt{\sigma_q \cdot \sigma_s}} = \frac{\langle (a_q(t+\tau) - \langle a_q(t) \rangle)(a_s(t) - \langle a_s(t) \rangle) \rangle}{\sqrt{\langle (a_q(t+\tau) - \langle a_q(t) \rangle)^2 \rangle \sqrt{\langle (a_s(t) - \langle a_s(t) \rangle)^2}}} \quad (1)$$

where τ denotes a predefined time lag, σ_i denotes the variance of the neuronal activity in the query region $i = q$ or the seed region $i = s$, and $\sigma_{qs} = \langle \big(a_q(t+\tau) - \langle a_q(t) \rangle\big)(a_s(t) - \langle a_s(t) \rangle) \rangle$ denotes the covariance of the fluctuations in neuronal activity in the query and seed regions, respectively.

The expected range of correlation coefficient should be -1 to +1, which means that if the value of correlation coefficient is closer to -1, these two ROI has more perfect negative correlation . And in the similar way, when the result is closer to +1, more perfect positive correlation should be estimated. At the same time, if the value is equal to 0, it means there is no correlation between the query and seed region. We could select a threshold ρ_0, which means if the result $\rho_{qs} > \rho_0$, the functional connectivity could be defined.

3.3.2 Spectral coherence

Unlike the correlation coefficient used in time domain, spectral coherence expresses the similarity of two time series in frequency domain.

Coherence[7], $Coh_{xy}(k)$, measures the linear time-invariant (LTI) relationship between two time series, x and y, at frequency k. It is defined as

$$Coh_{xy}(\lambda) = \left|R_{xy}(\lambda)\right|^2 = \frac{\left|f_{xy}(\lambda)\right|^2}{f_{xx}(\lambda) f_{yy}(\lambda)} \quad (2)$$

Where $R_{xy}(\lambda)$ is the complex valued coherency of x and y, $f_{xy}(\lambda)$ is the cross-spectrum of x and y, and $f_{xx}(\lambda)$ is the power spectrum of x. Coherence is a positive function, symmetric to x and y (e.g. $Coh_{xy} = Coh_{yx}$), and bounded by 0 and 1, where 0 indicates that x and y have no linear relationship, and 1 indicates that x can perfectly predict y in a linear fashion.

3.4 Method of Feature Extraction

One of the main parts of the project will be the definition and implementation of several features to characterize the graphs. MATLAB includes a toolbox on Graph and Network Algorithms, which can be used to plot graphs and help with feature calculation.

Graphs are usually used to describe the connections in a network, which currently has multiple use in the field of information system, medical imaging and etc[8]. Nodes and edges are two significant concept when it comes to the graph model. An example graph representing the connections is shown in Figure 6, a point with a number is a node or vertex, which means an object, while the lines between the vertices are the edges, which indicate the relationship between things.

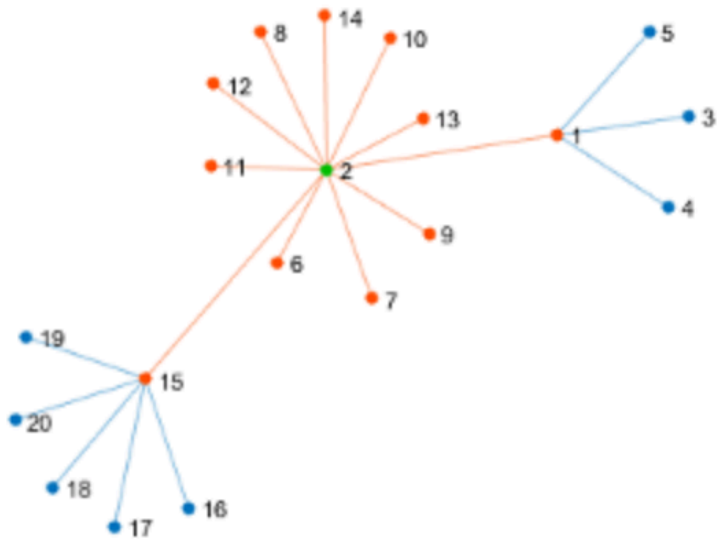

Figure 6 Example of graph shown in MATLAB

Based on the network estimation of above procedures, we are able to generate the graph variable for feature extraction. There are many analyzable features that can effectively describe different angles of the network graph. For instance, to analyze the structure, we can calculate the centrality of the graph, which represents the importance of each node. To analyze the traversals, features like distances and maximum flow can be measured and when it comes to the node details, the information of the node degree and its neighbors can be extracted from the graph. We can select some of these features to describe the brain connectivity network for the test of next classification process.

3.5 Method of Classification

Supervised learning used a set of labelled samples to build its model to predict or estimate the class value for a new subject. The supervised model makes attempts to find out whether there is a connection between the extracted attributes and the labels of each subject. Regression and classification are two popular method in the field of deep learning.

In the classification technique, the labelled data are represented by a set of known values or classes, which are used to build a classification model. This classification model is then used to assign a class label to an unknown sample[9][10].

Discriminant Analysis (DA) classifier was introduced by R. Fisher and it was used in many classification problems[11]. DA classifier is one of the basic and simple classifiers. There are two types of DA classifier, namely, linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) classifiers. In LDA classifier, the decision surface is linear, while the decision boundary in QDA is nonlinear.

3.5.1 Linear discriminant analysis

In LDA, each class k∈{1,…,K} is assigned a prior π such that $\sum_{i=1}^{k} \pi_k = 1$[12]. According to Bayes' rule, the posterior probability related to the feature sample *x* is

$$\Pr(G = k|X = x) = \frac{f_k(x)\pi_k}{\sum_{i=1}^{K} f_i(x)\pi_i} \tag{3}$$

Where $f_k(x)$ is the density of *X* conditioned on *K*. The maximum-a-posteriori(MAP) estimator simplifies to the following equation because the denominator is identical for all classes.

$$G(x) = arg\max_k \Pr(G = k|X = x) = arg\max_k f_k(x)\pi_k \tag{4}$$

The LDA method assumes that the density is Gaussian:

$$f_k(x) = |2\pi\Sigma_k|^{-\frac{1}{2}} e^{\left(-\frac{1}{2}(x-\mu_k)^T\Sigma_k^{-1}(x-\mu_k)\right)} \tag{5}$$

Where Σ_k is the covariance matrix for the samples from class *k*, and LDA assumes that all classes have the same covariance matrix, for example, $\Sigma_k = \Sigma, \forall k$. When we put the f_k into the classification method, we are able to achieve the classification function:

$$G(x) = arg\max_k \delta_k(x) \tag{6}$$

Where

$$\delta_k(x) = x^T\Sigma^{-1}\mu_k - \frac{1}{2}\mu_k^T\Sigma^{-1}\mu_k + \log\pi_k \tag{7}$$

is the discriminant function for class *k*. With this, we are able to realize the classifier.

3.5.2 Linear discriminant analysis

QDA is a variant of LDA in which an individual covariance matrix is estimated for every class of observations. QDA is particularly useful if there is prior knowledge that individual classes exhibit distinct covariances.

In QDA, we need to estimate Σ_k for each class $k \in \{1,\ldots,K\}$ rather than assuming $\Sigma_k = \Sigma$ as in LDA[12]. So the discriminant function of LDA is quadratic in x:

$$\delta_k(x) = -\frac{1}{2}\log|\Sigma_k| - \frac{1}{2}(x-\mu_k)^T\Sigma_k^{-1}(x-\mu_k) + \log\pi_k \quad (8)$$

Since QDA estimates a covariance matrix for each class, it has a greater number of effective parameters than LDA.

The figure below shows the visualization of comparison of LDA and QDA result[9].

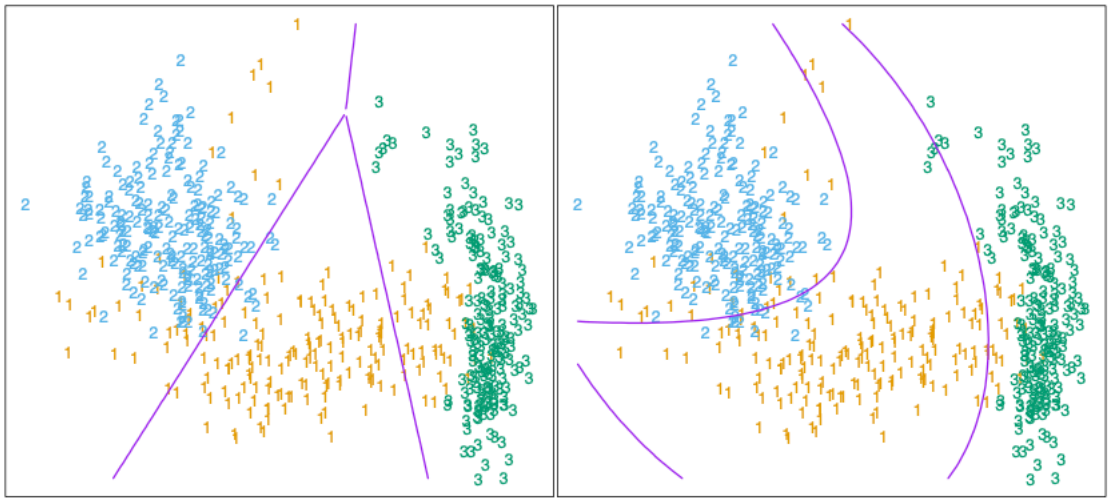

Figure 7 The left figure shows the classification of three classes using LDA while the right one shows the result using QDA

Capítulo 4. Development and Implementation

The process of the whole project can be divided into four steps:

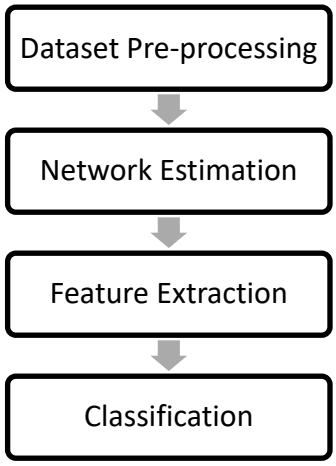

Figure 8 Flow chart of the whole project

The four steps are implemented through the MATLAB and its related toolbox.

4.1 Dataset

Our project uses experimental data to verify the effectiveness of the brain function network construction method. The data was selected from the fMRI data of the brains of 5 patients who have been diagnosed as AD, and fMRI data of the brains of 5 normal people as the control group.

All the data are collected from the Alzheimer's Disease Neuroimaging Initiative (ADNI)[13], which is a multicenter study which is aimed at developing clinical, imaging, genetic, and biochemical biomarkers in order to make contributions to the early detection and tracking of Alzheimer's disease.

The experimental scanning machine is the 3.0T superconducting magnetic resonance imaging system from Philips. The functional MRI image scanning uses an echo-planar imaging(EPI) sequence.

The Statistical Parametric Mapping(SPM)[5] software is used to pre-process fMRI data, including time correction, spatial registration, normalization, and smoothing using a Gaussian filter of diameter equal to 8 voxels. Finally, the whole brain time series data is obtained by filtering to remove high frequency physiological noise and low frequency signal drift.

4.2 Network Estimation

After pre-processing, one of the steps required for operation with fMRI data is converting the 4D array of values (size [X, Y, Z, time]) to a matrix of values (size [#voxels, time]). This change can be accomplished through the reshape function.

```
>> size(volumeInput)

ans =

    64    64    36   137
```

Figure 9 Size of fMRI data before reshape

```
>> size(volumeInput_2D)

ans =

      147456         137
```

Figure 10 Size of fMRI data after reshape

Therefore, with the 2D matrix to represent the fMRI information, we are able to get the time series of each voxel during the given time. The figure below shows the time series of the first voxel of the fMRI data:

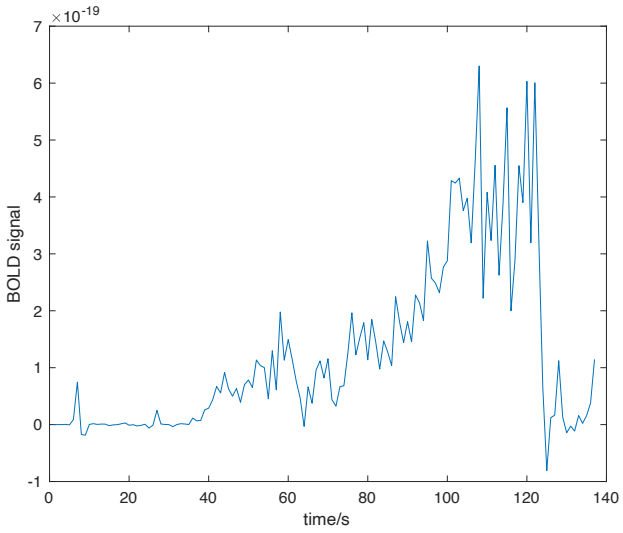

Figure 11 Time series of the first voxel

To calculate the connectivity map, it is necessary to separate the brain into several regions. From the 2D converted volume we are able to estimate a connectivity matrix between the different regions of the brain defined in the Anatomical Automatic Labelling(AAL)[14] template. AAL is a software package and digital atlas of the human brain, which is used to obtain neuroanatomical labels for the locations in 3-dimensional space.

The whole brain is divided into 116 standard functional regions, and all the voxel signals in the brain region collectively represent the neuron activity in this region.

After matching the voxels with their corresponding labels, the mean value of the time series corresponding to the voxels in each brain region is calculated, which is used to represent the time series and brain activity function attributes of the brain region.

Therefore, we are able to get a '116x137' matrix to represent each time series in each brain region:

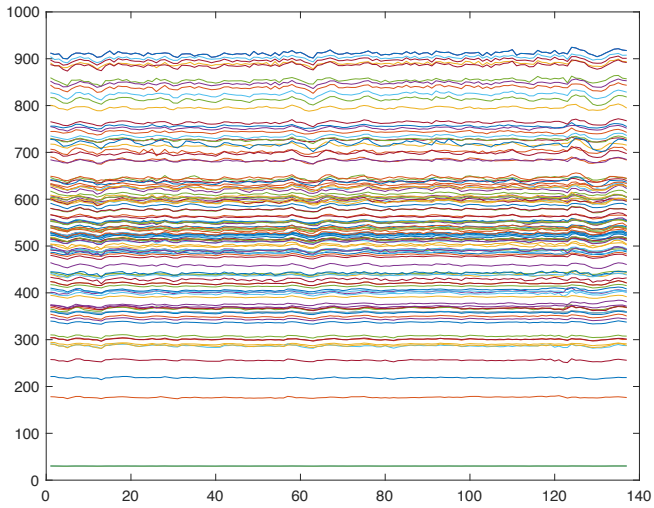

Figure 12 Times series of 116 brain regions

Then, with the application of equation(1) and equation(2), we are able to calculate the related correlation coefficient and coherence coefficients between the time series of these 116 regions. So we are supposed to achieve a '116x116' matrix to represent connectivity network of each method.

4.3 Feature Extraction

After we achieve the connectivity matrix, a threshold is set to determine whether there is a connection between two nodes. An example threshold set in this stage is 0.94, which means if the correlation or coherence coefficient between two regions is over 0.94, we can consider that there is a connection between them. With this step, we are able to get a Boolean matrix representing the network. With the application of 'graph' function in MATLAB, we are supposed to achieve the connectivity map, where we could perform the feature extraction operation.

When it comes to graph, we need to take the node and edge of this graph into consideration. In our brain, we can see the 116 different regions in our brain as the nodes, and the connections between each region achieved by connectivity network as the edges.

At current stages, we select two kinds of features used for classification step, which both can describe the conditions of current connectivity network. The first one is the average degree of the nodes in the graph. The degree is the number of edges connected to each node, we calculate the average value to represent the overall quality. Another feature is the average distance between nodes. The distance feature shows the length of the shortest path between two different node, and the mean value is also calculated for classification.

For each subject's connectivity map, we are able to extract these two kind of features. During the next classification step, we need to group these features of each subject, dividing them into training group and test group.

4.4 Classification

4.4.1 Realization of classification method

The classification of the graph features extracted from the connectivity graph will be approached using discriminant analysis-based classifiers. Those classifiers can be implemented using the "classify" (Discriminant Analysis) MATLAB function.

```
>> [class, err] = classify(sample,training,group)
```

This function takes process of the features in the sample, and make the classification process. About the input parameter, in this project, 'sample' is the matrix which includes the average degree and average distance of the test subject, while 'training' is the matrix that has the same column with the 'sample', holding the features of other nine training subjects. We will allocate a value to each training subject to represent their group, which in this case should be a Boolean value, either '0' or '1' according to the fact whether the subject has Alzheimer's disease or not.

About the output parameters, 'class' is the classification result, which should be one of the value in the 'group' matrix. 'err' is also a return value, which presents the estimated misclassification rate, based on the information of the nine pairs of training feature.

Using the parameter 'type', several discriminant-based classifiers can be implemented using different types of discriminant function:

```
>> [class, err] = classify(sample,training,group,'type')
```

The parameter 'type' could be chosen from the following values: 'linear', 'diaglinear'; 'quadratic' 'diagquadratic' and etc. 'linear' is the default value, which takes the application of equations from (3) to (7) to classify features that hold a normal density. 'diaglinear' is a method that has similar principle with 'linear', but its estimate is with diagonal covariance. 'quadratic' makes the use of the equation (8) to make classification. The relationship between 'diaquadratic' and 'quadratic' is the same as the first two type parameter.

Therefore, in our project, the linear discriminant analysis(LDA) can be implemented by choosing type as 'linear' or using the default value, while the quadratic discriminant analysis can be realized by setting type as 'quadratic' .

4.4.2 'Leave-one-out' test method

The leave-one-out (LOO) method is a useful method to train and test the classifier[15]. All sections within the dataset would be set as training set and test set in different classification iterations. For instance, if we have *n* sample dataset, in total LOO method would experience *n* iterations. And during each iteration, *n*-1 pairs of feature would be set as the training group, the remaining one would be the test group.

Compared to the test set approach, LOO approach has many advantages. First of all, it is not affected by the division method of the test set and training set, because each data has been tested separately. At the same time, it uses n-1 data to train the model, and almost all the data would be used to ensure that the bias of the model is smaller.

Therefore, in our project, we performed the same principle to realize the leave-one-out method. The features of the first subject will be setting as the test set and the remaining features make up the training set, then perform the same process to the second subject. After ten iteration, we are able to get the validation result of this classifier.

4.4.3 Design of the classification combination

Since we have two different connectivity network gained by correlation and spectral coherence, and we are also able to extract two different features from the network, along with two different classifier, LDA and QDA, we performed five combination of the process of classification.

1. Two features in total, both of them extracted from connectivity network of correlation, and followed by the method of LDA.

2. Two features in total, both of them extracted from connectivity network of correlation, and followed by the method of QDA.
3. Two features in total, both of them extracted from connectivity network of spectral coherence, and followed by the method of LDA.
4. Two features in total, both of them extracted from connectivity network of spectral coherence, and followed by the method of QDA.
5. Four features in total, two of them extracted from connectivity network of correlation, two of them extracted from connectivity network of spectral coherence, and followed by the method of LDA.

Capítulo 5. Results and Discussion

5.1 Connectivity matrix

5.1.1 Connectivity matrix using correlation

The functional relationship between brain regions is represented by the synchronization between brain region time series. The method used is to calculate the Pearson correlation coefficient between the time series of each brain region.

We used a loop to calculate the correlation coefficient between each brain regions, and at the same time we took use of the function 'abs()' to obtain the absolute value from the output, for negative correlation simply means that the growth trends of the two regions are opposite, so it is necessary to look at the absolute value of the correlation coefficient to determine the strength of the correlation.

The figures 13 and 14 shows comparison of the connectivity matrix and the distribution of correlation coefficient between a Alzheimer's disease patient and a normal person selected from our dataset:

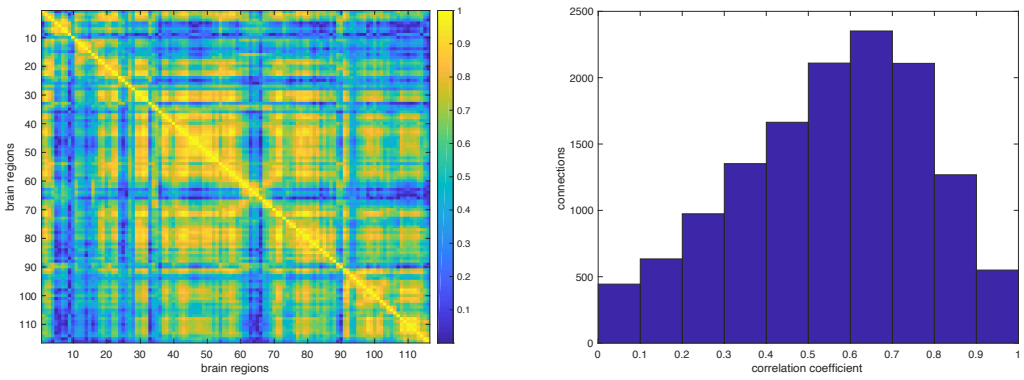

Figure 13 the connectivity matrix gained by correlation method(left) and the distribution of correlation coefficient(right) of Alzheimer's disease patient

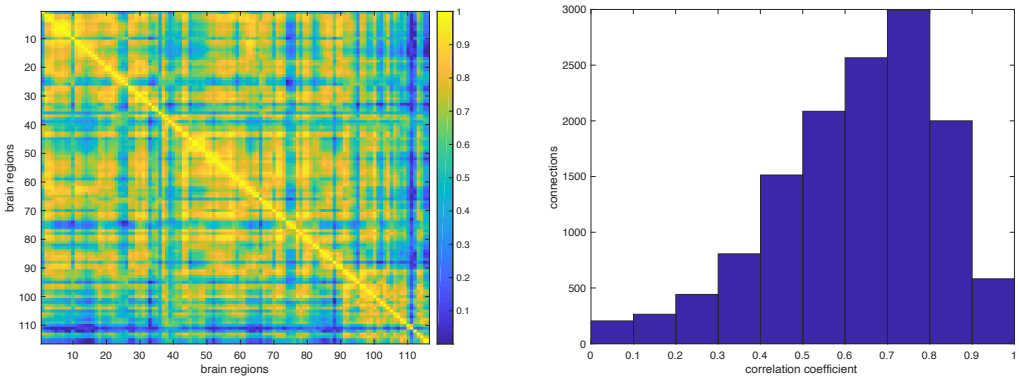

Figure 14 the connectivity matrix gained by correlation method(left) and the distribution of correlation coefficient(right) of normal people

It is clear to see from the figures above, in the one of Alzheimer's disease patient, the majority of the correlation coefficient is distributed between 0.6 and 0.7, while in the one of normal person, most coefficient is in the range of 0.7 and 0.8. Therefore, the coefficient of normal people is higher than the patient's. And also we should take the part between 0.9 and 1.0 into consideration, because we usually set the effective threshold above 0.9, which means patient with Alzheimer's disease has relatively less connections than normal ones.

5.1.2 Connectivity matrix using spectral coherence

We are able to estimate the connectivity matrix in the frequency domain using the equation(2) to calculate the spectral coherence between each brain region.

The spectral coherence of each regions at different frequency is also different. For instance, the figure below represents the spectral coherence between the first and second region of different frequency:

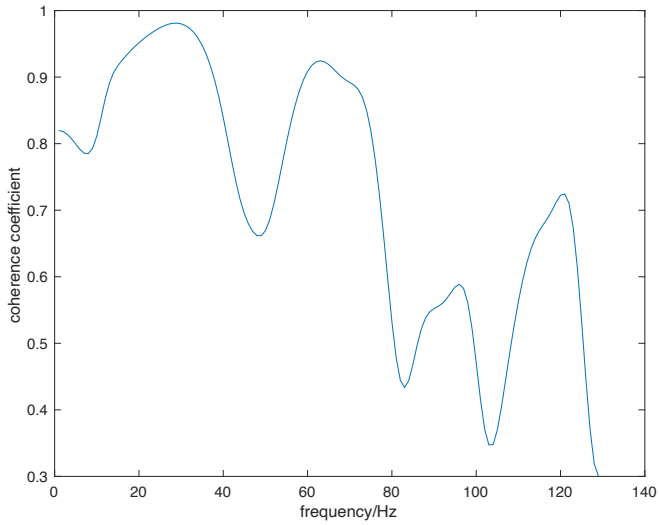

Figure 15 Coherence coefficient of first and second brain region

To achieve the estimated network, we calculated the mean value of the coherence of different frequency value to represent the coefficient of this region. Similar with the correlation method, we got the comparison of the connectivity matrix and the distribution of coherence coefficient between a Alzheimer's disease patient and a normal person selected from our dataset:

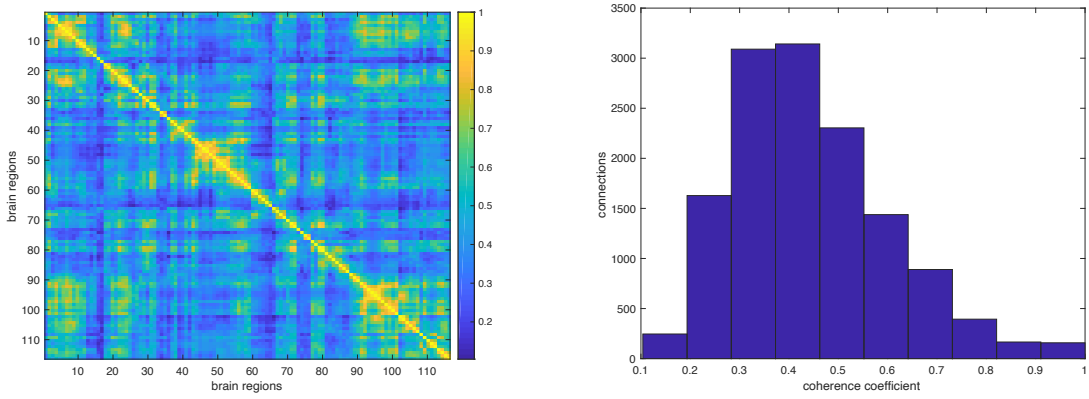

Figure 16 the connectivity matrix gained by coherence method(left) and the distribution of coherence coefficient(right) of Alzheimer's disease patient

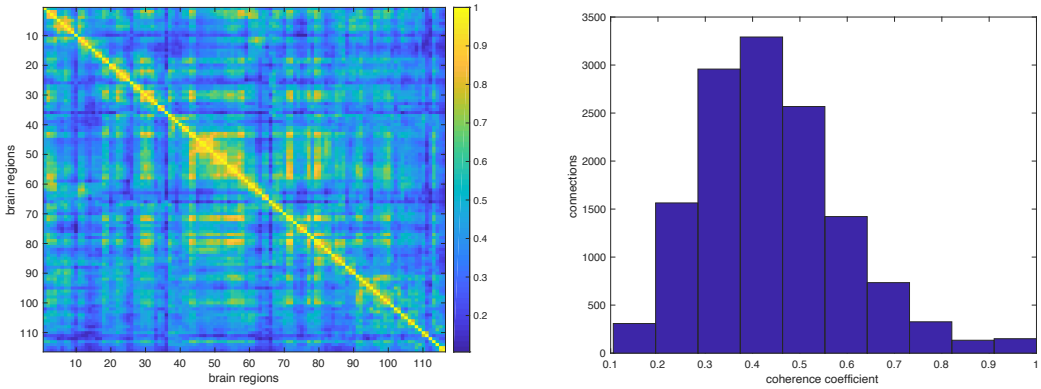

Figure 17 the connectivity matrix gained by coherence method(left) and the distribution of coherence coefficient(right) of the normal people

From the figure 16 and 17, we could see that the coherence coefficients achieved from normal people's network are generally higher than the Alzheimer's disease patient's, for they are distributed more in the range of 0.4 to 0.5, which means the connectivity network of normal people is more closely connected.

5.2 Extracted features

We have already got the connectivity matrix, so with the 'graph' function in MATLAB, we achieved the connectivity map of each subject. The picture below shows the connectivity map of the first subject:

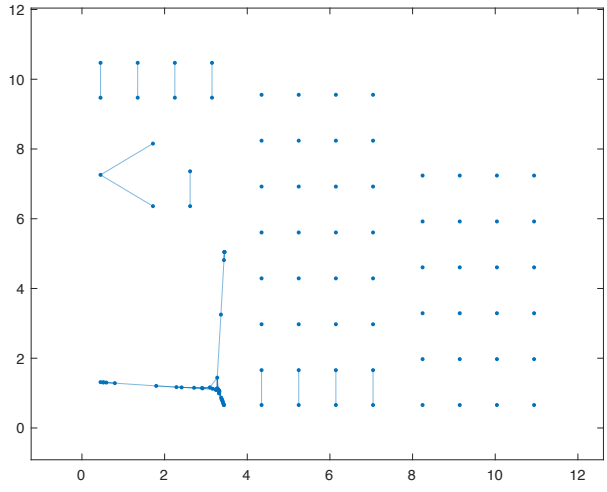

Figure 18 the connectivity map of the first subject

We calculated the average value of each feature to show the overall condition of the subject's brain. So we got two figures representing the two feature:

	1	2
1	4.5172	3.1057

Figure 19 The two features of the first subject's brain network

In figure 16, the first feature is the average degree of the nodes, which shows the centrality of the brain connectivity network, while the second feature represents the average distance between each node in the network.

After iteration, we can get the distribution of these two features of all ten subjects:

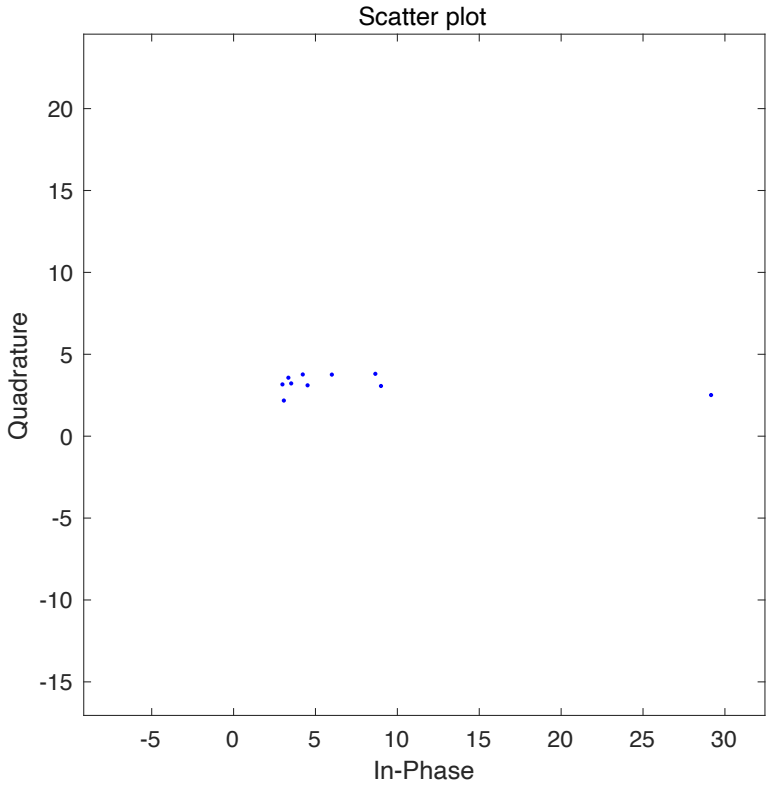

Figure 20 The distribution of two features of ten subjects

5.3 Classification Results

Since we only have two classes to be classified into, the patient with Alzheimer's Disease and normal person, we can easily set the patient with AD as class '1', and the normal one as class '0'.

The order of the dataset is that first five subjects are defined as patient with Alzheimer's Disease and the next five subjects are defined as normal people. So the ideal classification result is "1111100000", which will be used in the accuracy calculation of different combination mentioned above. The accuracy rate of each combination is shown below:

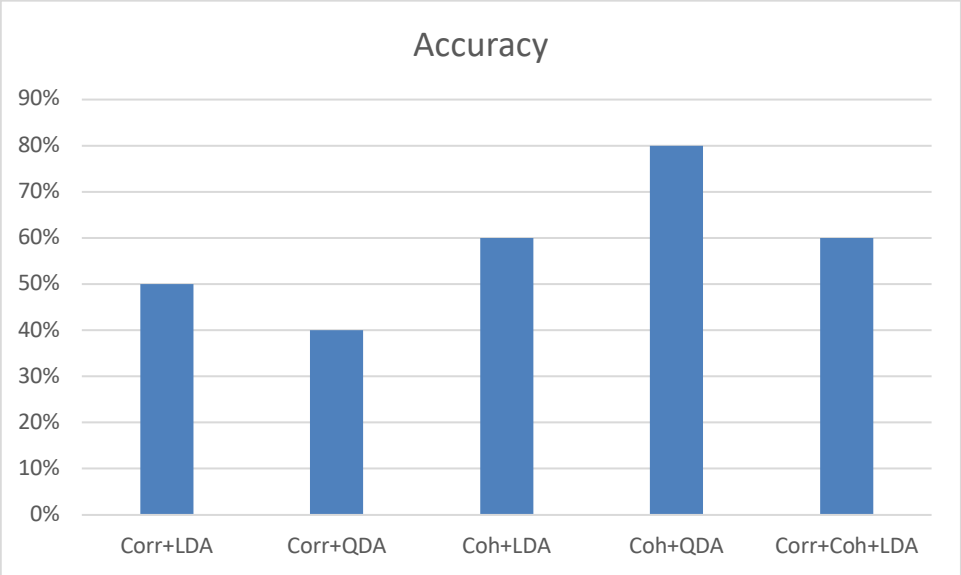

Figure 21 The accuracy of the five different test combination

From the result, we are able to tell that the combination 4, which represents the using of QDA classifier and features gained from coherence connectivity matrix, have the best performance when classifying the subjects.

We can also achieve the confusion matrix of each combination to take deep insight of them. In the field of machine learning and data analysis, the confusion matrix[16] is a useful visualization tool to show the result of estimation. The confusion matrix is a 2x2 situation analysis table that shows the number of the following four sets of records: positive records that made correct judgments (true positives), positive records that made incorrect judgments (false negatives), negative records that made correct judgments (true Negative) and negative records (false positives) that made wrong judgments. The confusion matrix of each combination are shown below:

		True condition	
		with AD	Normal
Predicted condition	with AD	2 True Positives	2 False Positives
	Normal	3 False negatives	3 True Negatives

Table 1 The confusion matric of Corr + LDA combination

		True condition	
		with AD	Normal
Predicted condition	with AD	1 True Positives	2 False Positives
	Normal	4 False negatives	3 True Negatives

Table 2 The confusion matric of Corr + QDA combination

		True condition	
		with AD	Normal
Predicted condition	with AD	2 True Positives	1 False Positives
	Normal	3 False negatives	4 True Negatives

Table 3 The confusion matric of Coh + LDA combination

		True condition	
		with AD	Normal
Predicted condition	with AD	4 True Positives	1 False Positives
	Normal	1 False negatives	4 True Negatives

Table 4 The confusion matric of Coh + QDA combination

		True condition	
		with AD	Normal
Predicted condition	with AD	3 True Positives	2 False Positives
	Normal	2 False negatives	3 True Negatives

Table 5 The confusion matric of Corr + Coh + LDA combination

When we look at the Table 1, we can see that this combination shows the result is not as ideal as what we expected, it is difficult to classify both two class.

Table 2 reveals similar with Table1, this combination tends to recognize even more subject as the normal class.

Table 3 is more interesting, which reveals that the recognition rate of normal class is much higher and this combination have a higher true negative rate than the first two combination.

Table 4 still shows the best results among these five combination. This Coh+QDA combination has the highest true positive rate and true negative rate.

Table 5 shows that with the features of coherence connectivity network, compared with combination 1, the combination 5 has higher recognition rate for AD class, and also has a higher true positive rate and true negative rate.

Capítulo 6. Conclusion and future work

In conclusion, best results are obtained by using coherence to define the connectivity edges and QDA as classification algorithm. This may explain that the underlying models fit some constraints: for coherence is a measure of the linear relationship between two signals, it seems that the signals corresponding to different regions are approximately linearly related, i.e., one channel could be reasonably obtained by passing the other channel through an unknown linear transfer function. On the other hand, QDA is optimal when the multivariate densities corresponding to the feature vectors of both classes are Gaussian with different means and covariance matrices.

Thus the obtained results are useful not only because of the classification performance but also because of the insights we get about the signal models, which may be useful information to understand the brain activity.

We can also conclude that centrality and average distance are two suitable features to describe that brain connectivity network, for the average distance measures brain resource integration ability, and the centrality coefficient quantifies the brain's ability to process local information. The information reflected from these features should also be useful in the diagnosis of Alzheimer’s disease.

About future work that may help to improve the overall performance, more features can be discovered to help describe our brain connectivity network, which may represent more detailed features between our brain regions. Also we hold the view that instead of studying all 116 brain regions in our project, we may select the brain regions that are more connected to cognitive function, since the correlation and coherence between other regions will also have influence to our key point. What’s more, the dataset we applied in this project is relatively too small, in spite of the benefits of Leave-one-out approach, it will take much more time if we want to perform larger group analysis.

Last but not least, the software and overall thought for our project is not restricted to Alzheimer’s Disease. This method may also have good development prospects in many other clinical diagnosis related to brain dysfunction, such as schizophrenia, hyperactivity in children, depression, etc.

Reference

[1] Blennow K, de Leon MJ, Zetterberg H. (2006). Alzheimer's disease[J]. The Lancet, (368): 387-403.

[2] Patterson, C. (2018). World Alzheimer report 2018: the state of the art of dementia research: new frontiers. Alzheimer's Disease International (ADI): London, UK.

[3] Chen L, Wang X, Zhang C. (2010). Progress of magnetic resonance imaging in Alzheimer disease and mild cognitive impairment. J Med Imaging, 20(4):574-576

[4] Lindquist, M. A. (2008). The statistical analysis of fMRI data. Statistical science, 23(4), 439-464.

[5] The FIL Methods Group(2019). SPM12 Manual. Retrieved April 13, 2020, from University College London Web site: https://www.fil.ion.ucl.ac.uk/spm/

[6] Lang, E. W., Tomé, A. M., Keck, I. R., Górriz-Sáez, J. M., & Puntonet, C. G. (2012). Brain connectivity analysis: a short survey. Computational intelligence and neuroscience, 2012.

[7] Sun, F. T., Miller, L. M., & D'Esposito, M. (2004). Measuring interregional functional connectivity using coherence and partial coherence analyses of fMRI data. Neuroimage, 21(2), 647-658.

[8] Mathworks, (2019). Graph and network algorithms. Retrieved February 26, 2020, from Mathworks Web site: https://www.mathworks.com/help/matlab/graph-and-network-algorithms.html?category=graph-algorithms&s_tid=CRUX_gn_documentation_graph-and-network-algorithms

[9] Duda, R.O., Hart, P.E. and Stork, D.G. (2012) Pattern Classification, 2nd ed., John Wiley & Sons, New York, NY, ISBN: 978-0-471-05669-0.

[10] Hastie, T., Tibshirani, R., Friedman, J. and Franklin, J. (2005). The elements of statistical learning: data mining, inference and prediction. The Mathematical Intelligence, Vol. 27, No. 2, pp.83–85.

[11] Guo, Y., Hastie, T. and Tibshirani, R. (2007). Regularized linear discriminant analysis and its application in microarrays. Biostatistics, Vol. 8, No. 1, pp.86–100.

[12] Matthias Döring. (2018). Linear, Quadratic, and Regularized Discriminant Analysis, Retrieved April 01, 2020, from Data Science Blog Web site: https://www.datascienceblog.net/post/machine-learning/linear-discriminant-analysis/

[13] Jack Jr, C. R., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., ... & Dale, A. M. (2008). The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine, 27(4), 685-691.

[14] Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., ... & Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage, 15(1), 273-289.

[15] Cawley, G. C., & Talbot, N. L. (2003). Efficient leave-one-out cross-validation of kernel fisher discriminant classifiers. Pattern Recognition, 36(11), 2585-2592.

[16] Stehman, S. V. (1997). Selecting and interpreting measures of thematic classification accuracy. Remote sensing of Environment, 62(1), 77-89.