Towards the Computer-aided Diagnosis of Dementia based on the Geometric and Network Connectivity of Structural MRI Data

Garry Smith[1,2], Zhivko Stoyanov[1,2], Danica Greetham[3], Peter Grindrod[4] and Doug Saddy[2]

[1]School of Systems Engineering; [2]Centre for Integrative Neuroscience and Neurodynamics;
[3]Centre for the Mathematics of Human Behaviour – University of Reading, UK
[4]Mathematical Institute, University of Oxford, UK

Introduction

- Fully automated, computer-aided diagnosis techniques have potential to rapidly increase diagnosis rates and reduce cost
- We present an intuitive geometric algorithm for analysing the structure of T1-weighted structural MRI scans using the highest available resolution
- Network Theory is employed to derive networks and test their fragility
- The analysis uses a fragility threshhold to classify structural MRI scans into three categories: Alzheimer's disease (AD); Mild Cognitive Impairment (MCI); Controls (CN)

The geometric and network structure of MRI data

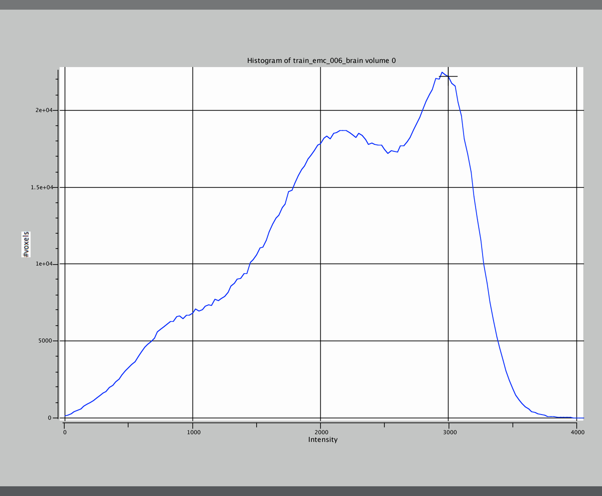
- A **3D** T1-weighted MRI image consists of $\mathbf{n_1} \times \mathbf{n_2} \times \mathbf{n_3}$ voxels and $\mathbf{f(i,j,k)} \geq \mathbf{0}$ is the level of T1-weighted signal recorded in voxel $\mathbf{(i,j,k)}$
  - We normalise the recorded signal for each brain so that we end up with $\mathbf{0} \leq \mathbf{g(i,j,k)} \leq \mathbf{1}$ for all voxels
- We assume negative changes in T1 signal gradients are a feature of neural degeneration
  - We use the normalised signal gradient to trace a path of similarity over long distances in the brain
- We focus on voxels for which the signal is above a certain threshold, θ
  - Starting at $\theta = \mathbf{0.6}$ allows us to generate connectivity networks based on primarily white matter values (see Figures)
- For each theshholded brain, we consider the **3**D set $\mathbf{A}_\theta$ and compute its surface area, $\mathbf{S}_\theta$, and its volume, $\mathbf{V}_\theta$
  - We then compute a measure of the fragility of its structure, $\mathbf{f}_\theta$, i.e how close $\mathbf{A}_\theta$ is to "breaking" apart into smaller components

Computing fragility

- Apart from being a geometrical **3**D object, we can think of $\mathbf{A}_\theta$ as a network, denoted by $\mathbf{N}_\theta$, in which two voxels are connected if they share a face or an edge (but not a corner)
- The advantage of interpreting $\mathbf{A}_\theta$, as a graph, or a network $\mathbf{N}_\theta$, is that we can apply techniques from Spectral Graph Theory
- Each graph/network can be represented with a matrix
  - Computing eigenvalues of such a matrix gives us a spectrum - an array of values that describes some structural characteristics of the given graph
- Zero eigenvalues correspond to the number of connected components
- The smallest positive eigenvalue of the Laplacian matrix, called *algebraic connectivity* , is an indicator of the robustness of the graph to vertex and edge failures and to betweenness in networks
- If $\mathbf{A}_\theta$ is split into $\mathbf{m}$ disjoint parts, this will correspond to $\mathbf{N}_\theta$ consisting of $\mathbf{m}$ connected components, which in turn corresponds to $\mathbf{m}$ eigenvalues equal to zero in the normalised Laplacian spectrum of $\mathbf{N}_\theta$
- The eigenvalues close to zero (around the second smallest normalised Laplacian eigenvalue) give us an indication of the fragility of $\mathbf{A}_\theta$
- The larger the number of eigenvalues that are close to zero, the more fragile (i.e. sensitive to breaking apart) $\mathbf{A}_\theta$ is

Distribution of tissue density

A histogram showing the distribution of the intensities across a brain; the two peaks roughly indicate the range of intensities for white and grey matter. We can see that by choosing $\theta \geq \mathbf{0.6}$ we predominantly select white matter



Project partners

Coronal view of selected tissue as θ increases

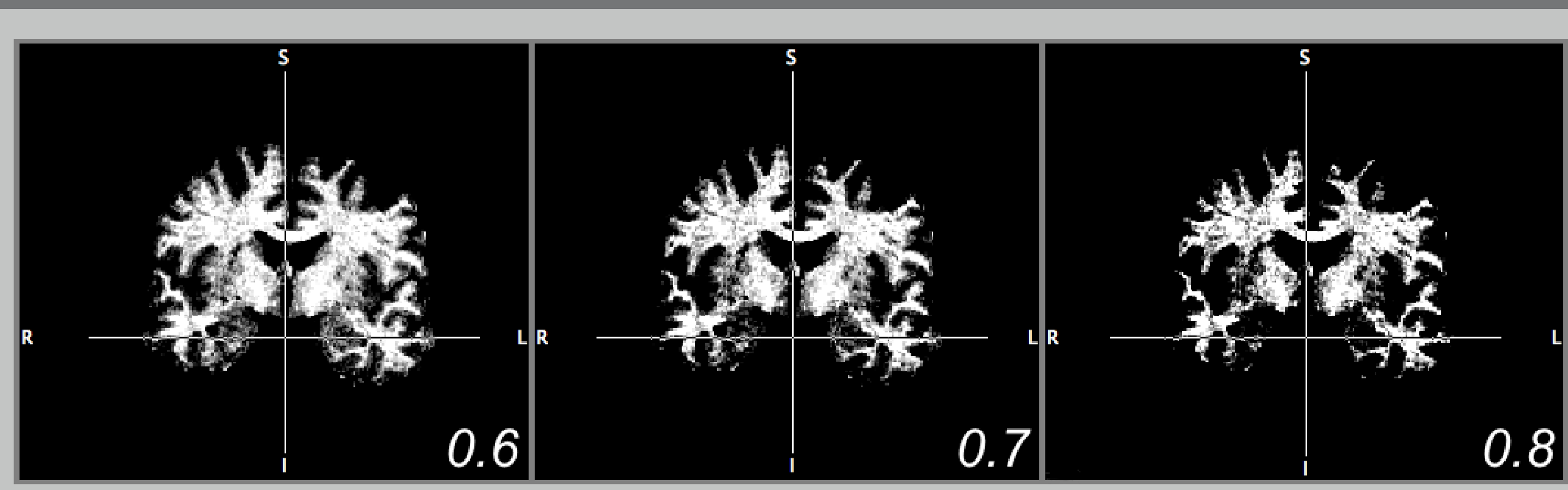


Figure: Coronal view of a single brain depicting the tissue that is selected as θ increases from 0.6 to 0.8

Results from training data

We calibrated the algorithm against the CADDementia training set as well as data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database, by combining $\mathbf{S}_\theta$ (surface area), $\mathbf{V}_\theta$ (volume) and $\mathbf{f}_\theta$ (fragility), with the age of the subject and used these four features (numbers) as *predictors* for the stage of neural degeneration (CN, MCI or AD). We firstly used gender to split the subjects apart into two groups.

Table: Partial output from the MATLAB function `mnrval` (multinomial logistic regression) applied to the group of female subjects on the CADDementia training dataset, $\theta = \mathbf{0.66}$

subject ID	diagn.	predict.	$\mathbf{p_{CN}}$	$\mathbf{p_{MCI}}$	$\mathbf{p_{AD}}$
train_emc_002	2	1	**0**	**0.78**	**0.22**
train_emc_003	0	0	**0.99**	**0.005**	**0.003**
train_emc_008	0	0	**0.89**	**0.0008**	**0.1**
train_emc_009	2	2	**0**	**0**	**1**
train_emc_011	1	1	**0**	**0.87**	**0.13**
train_up_001	2	2	**0**	**0.004**	**0.995**

- Classifications achieved (consistently) on training data appear promising:
  - CADDementia train (**30** subjects): $<$**20%** incorrect predictions
  - ADNI dataset (**189** subjects): $<$**35%** incorrect predictions
- Used up to **120** CPU cores to process subjects in a data parallel way
- Processing time for CADDementia test data (**354** subjects) *generally* between **7** to **25** minutes per subject
  - *However*, **26** outliers present
    - Processing time is dependent on the number of voxels left in the set $\mathbf{A}_\theta$ and on how fragile or connected $\mathbf{A}_\theta$ is as a **3**D structure

Conclusions

- A step towards employing Network Theory in the analysis and classification of neural diseases
- Agnostic to underlying tissue properties as well as the nature of the signal
  - We have previously applied a similar approach to resting state fMRI data, see *Grindrod et al (2014), Primary evolving networks and the comparative analysis of robust and fragile structures, Journal of Complex Networks, doi:10.1093/comnet/cnt015*
- For the CADDementia competition we intentionally biased the algorithm in favour of white matter by stepping up the threshold values
  - Stepping down would capture properties of grey matter
- Can be extended to include more sophisticated techniques from Network Theory as well as targeting brain regions for specific structural changes
- Workflow can be fully automated and scaled to massive numbers of CPUs, provisioned on demand, through private and/or public Cloud providers
  - Thereby potentially, allowing health authorities to offer wide-spread and frequent screening

Acknowledgements

We gratefully acknowledge funding from the University of Reading Research Endowment Trust Fund as well EPSRC grant EP/G065802/1 'The Horizon Hub'. The computational aspects of this work were based on a foundation previously laid down during the EPSRC NeuroCloud project under grant EP/I016856/1. Data collection and sharing partly funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (NIH Grant U01 AG024904) and DOD ADNI (DOD award W81XWH-12-2-0012).

g.m.smith@reading.ac.uk