Discovery of Discriminating Neural Regions for MRI Classification

Abstract

Machine learning methods can be applied to MRI scans of the brain in order to classify patients as having, or not having, particular characteristics, such as Alzheimer's Disease, advanced age, or a high level of education. Accurate classification effectively demonstrates that some set of image features can be used to determine the correct classification, and are therefore correlated with the classification in question. This work presents the Graph Neural Analyzer, a combination of new and existing components that can discover these features for a variety of potential classifications including age, level of education, gender, socioeconomic status, ethnicity, and Alzheimer's Disease.

1. Introduction

Driving a taxi in London has been found to affect brain structure (Maguire et al., 2000). Not only do London taxi drivers have recognizable structural changes compared to the general population, these changes occur upon beginning taxi driving, and do not appear to be the result of any innate navigational ability (Maguire et al., 2003). Discovering this fact from magnetic resonance imaging data required hard work and dedicated effort by knowledgeable researchers, to discover the neural regions which can be used to discriminate one class from the other. Interpreting the meaning of the discriminating neural regions (DNRs) for a a particular classification requires human knowledge and creativity at present. However, discovering the DNRs automatically in order to assist the process may be possible.

Performing machine learning on MR images of the brain, such as in (Long & Holder, 2012a), has the potential for automation in this regard. It is exploited to some extent in (Long & Holder, 2012b), where correlations are discovered using a graph-based representation of 3D structure, and displayed as an overlay on MR images. We use this representation along with tools for processing it as a component for the Graph Neural Analyzer (GNA).

The Brain Extraction Tool (BET) (Woolrich et al., 2009) attempts to isolate the brain from the rest of the head in structural MR images. Classifying based on the entire head raises an issue in that structures other than the brain may provide a more accurate means to make certain classifications (Long & Holder, 2012b). Use of BET may prevent this, and has the potential to increase accuracy by allowing limited processing resources to be focused on only the brain. BET is used as another component of GNA.

Separation of the brain into major structures rather than approaching the brain as a whole may avoid problems encountered in (Long & Holder, 2012b) due to varying locations of neural structures. This is possible using FIRST (Patenaude et al., 2011), which is also made available by the FSL library. This is used as another component of GNA.

Finally, inconsistent orientation of images may have an effect on classification. In order to evaluate the effect of orientation, we mark two points on each image, and level all the images according to these two points. These are marked in the midsaggital plane, and effect rotation about one axis only. Manually marking these points allows assessment of the maximum benefit of automating the marking procedure. This is included as a component in GNA.

GNA is evaluated on a number of categorizations, including age, gender, level of education, socioeconomic status, Alzheimer's disease, ethnicity, and imaging facility. On some categorizations such as age and certain educational distinctions, it achieves over 90% accuracy classifying test images. On others, accuracy is lower, such as socioeconomic status (64% accuracy, compared to 50% for random guesses). The accuracy may reflect the level of neurological difference between each category. Full results are given in section 4.

2. Previous Work

Some previous work has been focused on the particular problem of automatic recognition of Alzheimer's Disease from MRI data. For example, Klöppel et al. (2008) consider each voxel to be a feature in a feature vector, and then use a support vector machine to

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.

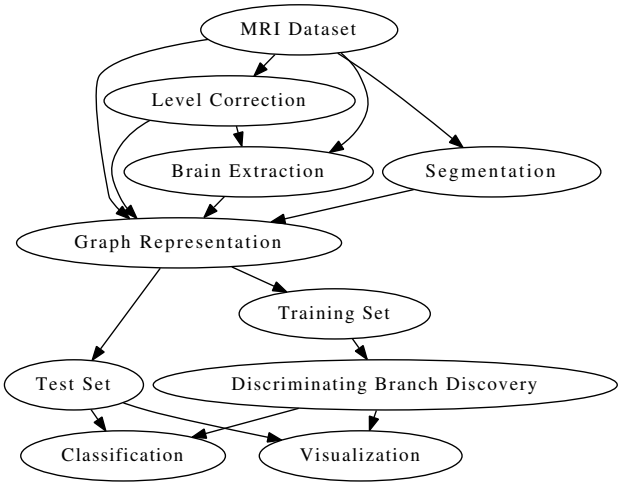


Figure 1. Overview of the correlation discovery process. Some components (Level Correction, Brain Extraction, Segmentation) can be used, bypassed, or combined. Classification accuracy can be tested, or a visualization of the differences between categories produced.

classify the resulting feature vectors. Cuingnet et al. (2010) discuss and compare 10 different methods using a large dataset from 509 participants. As such, the study of automatic detection of Alzheimer's Disease is well-studied. The method by Klöppel et al. (2008) differentiates discriminating vs. non-discriminating voxels, somewhat like the discriminating branches we propose below. However, a discriminating branch may represent a variable number of voxels depending on the length of the branch, and never represents as few as one voxel. Accuracy of the Alzheimer-specific methods evaluated in is higher than that of our method, although they were evaluated on a larger dataset which may affect accuracy.

Elsayed et al. (2010) have used graph-based shape representation to classify MR images using the 2D shape of the corpus callosum as it appears in a midsaggital section. Images were classified as either from a musician, or a non-musician, with up to 95% accuracy. Shape analysis was done by recursively subdividing the image into 4 quadrants to form a quad-tree, terminating a branch if the area to be subdivided was sufficiently uniform in color. These trees were then classified by a frequent sub-tree classification method. This current work also represents shape using a tree of subdivisions.

The primary difference between GNA and these previous systems is that GNA is not tailored to any particular classification, and is intended to address classifications which have not been well-studied.

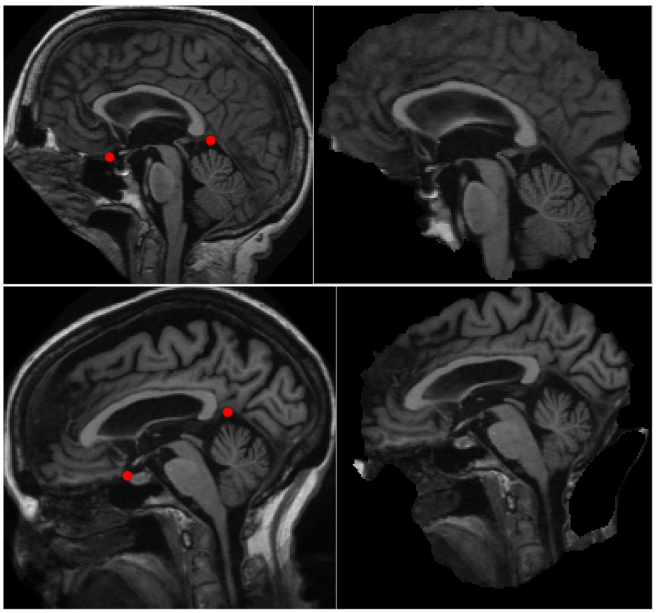


Figure 2. Left column: Images before application of BET. Right column: Post-BET images. Most images in the dataset process similar to the top image, however in some cases more of the brain stem is included, as on the bottom right. Also note that the bottom image retains an artifact on the right side after brain extraction. Dots in the pre-BET images indicate points for level correction.

3. Method

In general, classification is performed by forming a graph representation of each MRI, finding a set of subgraphs which characterize each class, forming feature vectors using these subgraphs, and using a support vector machine to classify the feature vectors. See figure 1 for an overview of the procedure. If instead of an accuracy test, the discriminating neural regions themselves are to be the result of the process, DNR discovery is performed in the training data, and each DNR is evaluated individually on the test data.

3.1. MRI Data

A Magnetic Resonance Image (MRI) is stored as a three dimensional greyscale image. For each voxel, an intensity value is provided. On T_1-weighted scans, as we are using for this work, a low value indicates a low fat content (and typically high water content, although area outside the skull also exhibits low values). The exact value range for brain tissue compared to other content such as cerebrospinal fluid (CSF) varies from one image to the next, which must be taken into account. Resolution is 150*256*256 in one dataset (IXI) used to test GNA and 160*256*256 in the other (OASIS). The datasets are described in more detail in 4.1.

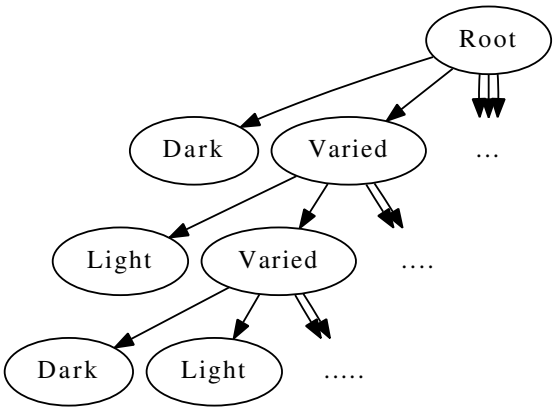


Figure 3. Example of how a portion of a tree representing a neural image may look.

3.2. Graph Shape Representation

Shape is represented as a graph by recursively subdividing the image into 8 equal boxes, forming a 2x2x2 grid at each subdivision. Subdivision is continued until each box is either sufficiently uniform in color, or the depth limit is reached. A tree is formed from this subdivision process, with each division forming a node, and each box which will not be further subdivided forming a leaf. The tree size can be tailored by limiting the maximum depth or adjusting the requirement for uniform color. Nodes are labeled to indicate the reason for termination, and edges are labeled to indicate which subdivision they represent. This allows area represented by any node to be located in 3D space. The result is a large tree, of which a portion may look like figure 3. In our tests, graphs range in size from 300 to 20,000 nodes depending on the generation method. This is similar to the representation in (Long & Holder, 2012a) except as required to support neural segmentation.

Empty space is cropped from around each image before beginning this procedure.

3.3. Brain Extraction

Brain Extraction Tool (BET) (Smith, 2002) from the FMRIB Software Library (FSL) (Woolrich et al., 2009; Smith et al., 2004) was used to remove the skull skull and face from the images, as show in figure 2. Removing the skull and facial features from the images is important for some categorizations such as gender, where the program will otherwise classify images primarily by skull shape. It also improved accuracy on some classifications.

Calculating a crop threshold correctly is more difficult when the brain is not surrounded by bone tissue which appears as high intensity in the images. In cases where not all skull tissue is removed, such as in figure 2, the image may not be cropped in a manner consistent with images where the brain was extracted more accurately, which introduces additional noise into the data.

3.4. Level Correction

As seen in figure 2, not all images are rotated identically. In order to increase rotational uniformity, we marked two locations in each image, then rotated all images to place those two points horizontally. Both points were marked in a midsaggital view, and correction was only made for rotation about a line running medial-lateral halfway between those two points. The image can be (optionally) cropped in a sphere centered between the two marked points. Cropping in this manner unifies the amount of neck in each image, which increases the consistency of what neural structure is represented by each branch.

The primary difficulty is in choosing two points which give a representative sample of the orientation of the brain. The brain is highly variable from one individual to another, with the result that a pair of completely satisfactory points may not be possible. The points marked in figure 2 were used for this work, and the position of the corpus callosum relative to the anterior mark is not highly consistent. This no doubt limits the effectiveness of the method. Nevertheless, it improves accuracy on many classifications.

3.5. Image Segmentation

Considering a few important neural structures independently rather than the brain as a whole minimizes the effect of rotational inconsistency, and allows consideration of borders between white and gray matter structures. The tool FIRST from the FSL package (Patenaude et al., 2011) can be used to separate an MRI into separate images of major structures. We use this tool as a preprocessing step, and generate a separate shape tree for each structure. The independent shape trees are linked to a common root. Level correction was not applied in combination with this technique.

Fifteen structures are used: Left and right thalamus, caudate, putamen, pallidum, hippocampus, amygdala, accumbens, plus the brain stem combined with the 4th ventricle.

Total runtime using this method has the potential to be long if it is not restricted from producing a tree 15 times larger than one representing the brain shape

as a whole. This can be reduced by restricting the maximum tree depth or changing branch termination conditions. In our tests, attempts to lower the average tree size below 20K nodes resulted in a substantial accuracy decrease.

In some cases, image segmentation increased classification accuracy. In all cases, runtime was increased substantially due to the increase in graph size, in addition to the time required for the segmentation process itself.

3.6. Tree Classification

Support Vector Machines are commonly used to classify items represented by feature vectors (Muller et al., 2001), where a feature vector consists of binary attributes representing useful information about each item. One way to apply this to graph classification is to generate the feature vectors from a set of subgraphs, where presence or absence of each subgraph in an item determines the value of one feature (Deshpande et al., 2005). This requires each graph be searched for each subgraph. Subgraph isomorphism is NP-Complete.

In this case, the use of trees allows a simple optimization. Feature vectors are constructed based on branches which are required to include the root node. All such branches in a tree can be enumerated in polynomial time, and a tree can be searched for such a branch in time proportional to the length of the branch. This does not allow for consideration of image features which do not have a fixed location relative to the bounds of the crop area. However, it does allow timely processing of trees containing thousands of nodes. The process is performed on a computing cluster. Details of the algorithm for this are given in (Long & Holder, 2012a).

The conditions under which a branch is selected to determine a part of the feature vector greatly affects accuracy. Branches which only occur in one image are ignored completely, as are branches which occur in equal numbers in both categories. Other branches are scored based on disparity of prevalence between categories, and prevalence in the category in which it is more often encountered. A branch receives a perfect score if it is found in every tree of one category and none of the other. A branch occurring in half the trees of one category, and none of the other, would receive a perfect score for disparity (only occurs in one category), and a 50% score for prevalence (occurs in 50% of the trees of the category it is commonly found in). The overall score is a weighted average of the two, and the weighting is adjustable. Only the best branches are picked to determine features, and GNA contains a parameter for the number of branches used. In our experience, 500 branches, and a equal weighting between prevalence and disparity will produce near-optimal accuracy on most benchmarks.

3.7. Optimal DNR Selection

When providing a set of relevant DNRs, each DNR should ideally be discriminating in general, and not a statistical anomaly present in only the training set. In order to test this, when displaying DNRs each DNR is validated individually on the test set. 10 trials are performed with 10 different test sets, as for an accuracy test. DNRs discovered and validated in multiple trials are displayed overlaid on neural images, as in figure 3.5.

Using the results from this process is not an accurate reflection of the actual classification hypothesis. Classification is performed based on DNRs found in the training data, using a Support Vector Machine trained only on the training data, and the only use of the test data is for evaluation. The DNR selection above is intended to provide a more accurate indication of the physiological patterns associated with a particular category of individuals. Classification is intended to test the system, and DNR display is intended to inform the user of the most useful patterns.

GNA supports coloration of branches either according to classification value, or according to meaning and status in the image they are displayed on. Examples of both styles can be found in section 4.

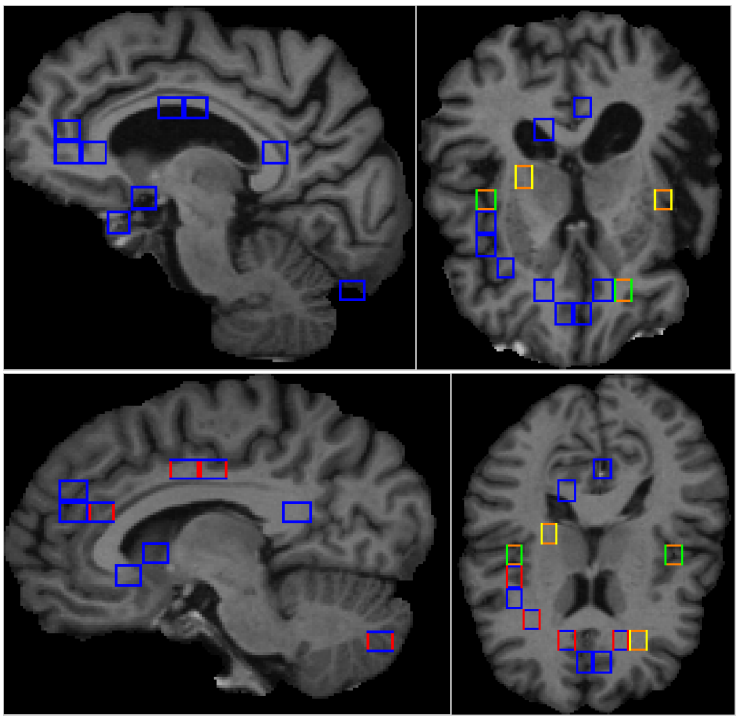
4. Results

4.1. Available MRI data

Data is available from the Open Access Structural Imaging Series (OASIS) project (Marcus et al., 2007). This is a dataset consisting of 416 structural MR images, some of individuals with varying levels of cognitive impairment. They are labeled according to the degree of cognitive impairment due to Alzheimer's disease. The data is in the Mayo Clinic Analyze 7.5 format. The Nipy library can be used to access this data from Python code (Millman & Brett, 2007).

Another dataset is available from the Information eXtraction from Images project [1], consisting of 590 images. We used the T1 images from this dataset, but other types are available as well. The IXI images were collected at three different facilities. This must affect the image, because GNA can classify the scan facil-

[1] Available from http://brain-development.org



Legend:
Positive/Found
Negative/Found
Vertical: Positive/Absent
Horizontal: Not-Brain
Vertical: Negative/Absent
Horizontal: Brain tissue
Horizontal: Depth cutoff

Figure 4. Images of an Alzheimer's Patient (top) and healthy individual (bottom). Both individuals are 65 years old. Saggital images (left) are offset slightly from midsaggital, to show more of the lateral ventricles. Branches are colored according to meaning and presence or absence in each image. Note changes in which branches are found in which image. Boxes represent the outline of 3D areas.

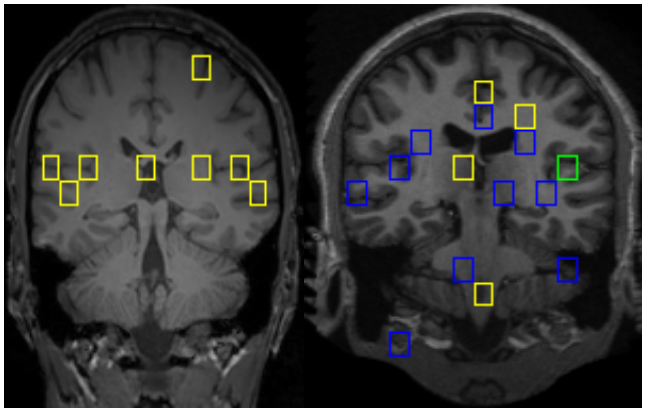


Figure 5. Age discriminating neural regions from IXI (left) and OASIS (right) datasets. Coloring is by score for each DNR. Green is highest, followed by blue, yellow, and orange. Boxes may extend either direction from the plane of the coronal sections shown.

ity with over 95% accuracy. The images are in NIFTI format, which can also be read by Nipy (Millman & Brett, 2007).

4.2. Age and Alzheimer's Disease

The OASIS dataset contains MRI scans from 100 individuals diagnosed with Alzheimer's Disease, graded into cognitive impairment levels 0.5 (70 scans), 1.0 (28 scans), and 2.0 (2 scans). Accuracy of clinical diagnosis of Alzheimer's Disease is not always perfect (Burns et al., 1990) which is a potential source of noise which we cannot eliminate. Accuracy on a 60-sample dataset constructed using only the CDR (Clinical Dementia Rating) 1.0 and 2.0 patients and 30 randomly selected healthy examples results in an accuracy of 88.3% using BET, and 80.0% without using BET. This is an increase from 79.3% reported in (Long & Holder, 2012a). However, controlling for age in the healthy examples such that no healthy scan is from an individual younger than the youngest example of Alzheimer's Disease results in a reduced accuracy of 70.0% without BET, and 68.5% using BET. This is one of only two tests on the OASIS data in which BET was found to reduce accuracy, the other being gender.

Alzheimer's Disease is known to increase ventricular size (Nestor et al., 2008). We also find a strong correlation between ventricular size and Alzheimer's Disease. Also, a curiosity of the OASIS dataset is that all participants with Alzheimer's Disease had at least some higher education. Finally, because only 30 individuals in the dataset exhibit CDR 1.0 or greater, the sample size for this test is small. An example of classification for Alzheimer's Disease is given in figure 3.5.

Many structures in the brain shrink with advanced age (Raz et al., 2005). This results in increased ventricular size and decreased overall brain mass. This overlaps with the changes due to Alzheimer's Disease. Classify-

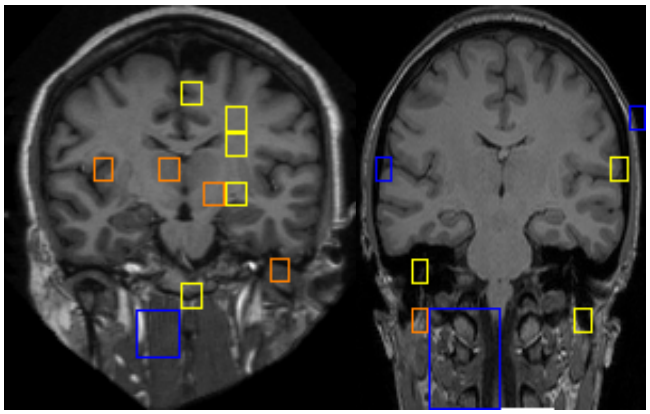


Figure 6. Discriminating neural regions from qualification (1 vs. 5, left) and education (right).

ing for age alone results 91.3% on the OASIS dataset, and 84.5% on the IXI dataset.

Because the bounding box is adjusted to the outside of the brain, overall size of the brain is not available to our model when using BET. Ventricular enlargement is detectable, as is sulci enlargement. Figure 5 shows a number of branches terminating in fissures and sulci, possibly due to overall brain shrinkage.

4.3. Education/Qualification

The OASIS data is annotated with number of years of higher education (if any) for each individual. 181 individuals are listed with no higher education, and 101 with 4 or more years. A balanced dataset was constructed using all individuals with 4 or more years, and an equal number of randomly selected individuals with no listed higher education. Accuracy on this dataset was 82.3% using BET, and 81.7% without using BET. Accuracy reported in (Long & Holder, 2012a) on this test is 77.9%.

There are a large number of older individuals in the OASIS dataset, and a greater percentage of them are highly educated compared to younger individuals. Also, education is known to have affects on the aging brain (Alley et al., 2007; Wilson et al., 2004). An example of an educated individual is given in figure 6.

The IXI dataset contains level of qualification, numbered 1 through 5. Meaning of each level is given in figure 7, as well as the best accuracy on the 2-class problem distinguishing each level from each other level. As expected more distinct levels (such as 1 vs. 5) are more easily distinguished than levels close together (such as 2 vs. 3).

Figure 7. Accuracy distinguishing each level from each other level. Meaning of each level:
1. No Qualifications (45 people)
2. O-levels, GCSEs, or CSEs (53 people)
3. A-levels (39 people)
4. Further education e.g. City and Guilds / NVQs (106)
5. University or Polytechnic Degree (307 people)

Class	2	3	4	5
1	78.9	75.7	77.8	95.6
2		68.0	63.5	86.7
3			67.1	87.1
4				76.8

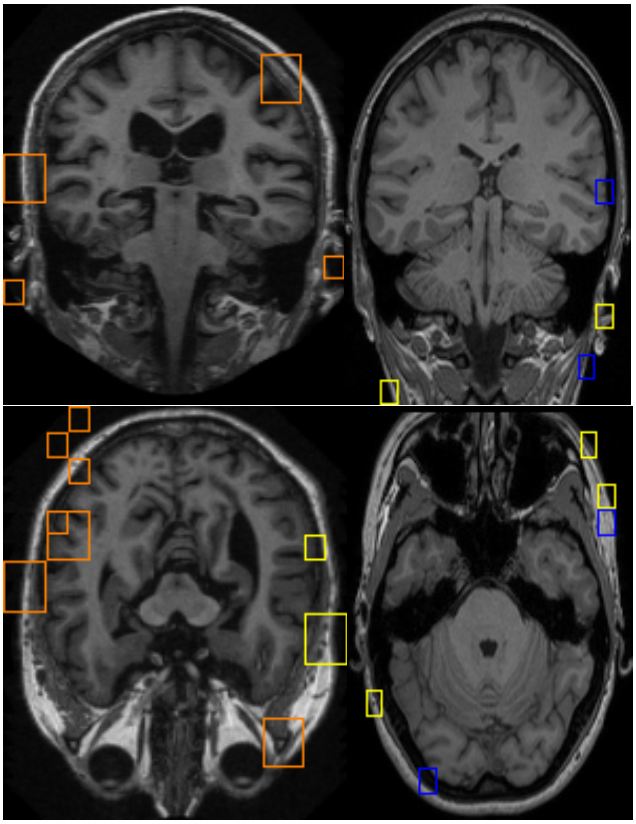


Figure 8. Images from OASIS (left) and IXI (right) from gender classification, showing discriminating regions based on skull shape rather than neural characteristics.

4.4. Gender

When not using neural segmentation or BET, the skull is part of the images to be classified. This allows GNA to select discriminating branches which represent facial features or skull shape, as show in figure 8. Accuracy based on these skull features is higher (81.2% on OASIS, 74.6% on IXI) than using methods which do not include the skull (72.1%, neural segmentation on IXI).

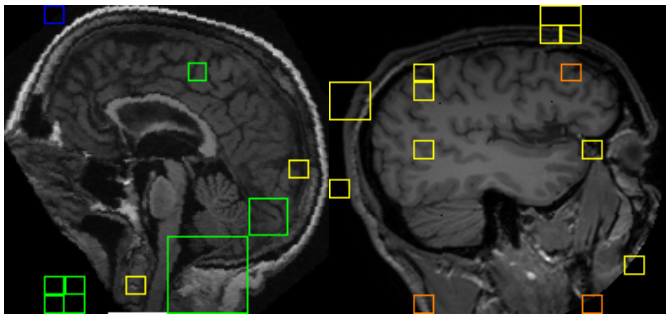


Figure 9. Midsaggital view showing some discriminating branches for classification by socioeconomic status (left), and ethnicity (right)

The brain is known to differ between genders (Goldstein et al., 2001), and so we anticipate that some future improvement to GNA may increase accuracy on this test.

4.5. Socioeconomic Status

The OASIS data is annotated with level of socioeconomic status, assessed by the Hollingshead Index of Social Position (Hollingshead, 1957). Levels 1 through 5 are given. In order to obtain a wide separation and reasonable sample size, level 1 is used as one category, and levels 4 and 5 as the other. This allows construction of a 100-sample balanced dataset. Maximum accuracy on this dataset is 64.0%. A midsaggital view showing some of the discriminating branches is shown in figure 9. Although the program is able to correctly categorize nearly $\frac{2}{3}$ of the examples, the interpretation of the discovered branches is not obvious.

4.6. Ethnicity

The IXI data is annotated by ethnicity, with the following possibilities and number of examples: White (451), Black (15), Asian (50), Chinese (14), Other (14).

The primary race given is "White", as is predominant in the United Kingdom where the data was collected. In order to test the performance of GNA on this task, a dataset consisting of the 14 "Chinese" and 50 "Asian" examples as one class, and 64 "White" examples was used. Accuracy on this dataset was 86.0%. Some branch locations found are given in figure 9.

4.7. Scan Location

The IXI data is annotated with scan location, and it is possible to classify scans on this attribute with high accuracy. Three scan locations are used, HH (180 scans), Guys (314 scans), and IOS (69 scans). This presents three two-class problems, with the following accuracy results: HH vs. Guys, 97.2%, IOS vs. Guys, 98.6%, HH vs. IOS: 100%.

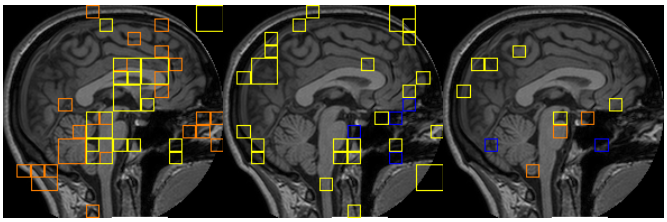


Figure 10. Midsaggital views showing discriminating branches for distinguishing collection facilities(from left to right) hh vs. guys, hh vs. ios, and ios vs. guys. All are overlaid on the same image.

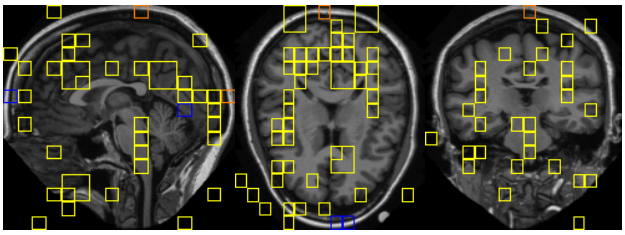


Figure 11. Saggital, horizontal, and coronal sections showing locations of discriminating branches for distinguishing one data collection facility (Guys) from the other two. There is no strong correlation between scan location and any other annotated feature in the data.

Different equipment is used at each of three facilities, both in manufacturer and power, which may account for the distinctions found by GNA. Classification of Guys vs. both other options results in 98.2% accuracy. This does not indicate any similarity between HH and IOS, except that Guys is distinguishable from both others by the set of discriminating branches in figure 11.

As seen in figure 10, discriminating branches are found over much of the image. The highest score achieved by any branch was 0.687.

4.8. Correlation Analysis

In some cases, a high degree of correlation between two categorizations may enable accurate classification on one categorization due to neural differences relating to the other. For example, if all highly qualified people were also elderly, age-based distinctions could result in apparently high accuracy on qualification.

The degree of correlation between each categorization tested is given in figure 12 for Oasis data.. None of

Figure 12. Table of correlations for OASIS data. For the IXI data, only one value was above 0.12, which was qualification level 1 vs, level 5 and age, with a value of .22.

Classification	**1**	**2**	**3**	**4**	**5**
Age (1)					
Gender (2)	0.10				
Education (3)	0.48	0.02			
SES (4)	0.00	0.01	0.0		
Alz w/gap (5)	0.27	0.02	0.10	0.05	
Alz w/o gap (6)	0.49	0.02	0.24	0.06	0.54

the categorizations have a high degree of correlation except for education and age on the OASIS data. On the IXI data, the highest correlation values are age vs some qualification level distinctions (1 and 5, 0.22, 2 and 5, .12, 4 and 5, .12). The table is omitted due to space constraints, and the remainder of values are all less than 0.12. We cannot rule out correlations with unlabeled factors.

4.9. Processing Time

Time complexity of the discriminating branch finder is exponential relative to the average number of nodes in the trees, and the total number of trees in the dataset. Thus these two parameters strongly affect the time to run the algorithm. Initial processing using BET takes a few hours using 3 out of 4 cores on an Intel Q6600 running at 2.9 gHz. Generating graphs takes approximately 12 hours using the same hardware.

In all cases, we have performed discriminating branch discovery using a computer cluster of either 296 or 1,968 processors. On this hardware, any of the accuracy values used in this paper can be reproduced in a few hours except for those using larger graphs from image segmentation, which take up to a few days due to the increased number of nodes. We do not have exclusive access to the cluster, and so the time depends to some extent on load from other users. Some tests were performed using a maximum tree depth of 6, with an average graph size of 12,880 nodes. The Alzheimer's dataset without CDR 0.5 requires nearly 24 hours to process on the 296-cluster compared to less than half an hour using a maximum tree depth of 5, and no accuracy increases were observed.

No attempt was made to complete processing on desktop hardware. This may be possible in the future given the regular increases seen in processing speed (Tuomi, 2002).

5. Conclusion

The Graph Neural Analyzer is capable of finding classification criteria for a variety of conditions with no assistance regarding the nature of the classification. For a classification where nothing is known about what structural changes are correlated with the condition, GNA is capable of automatically producing an overview. When applied to studied conditions such as Alzheimer's Disease and aging, it produces findings consistent with present knowledge. For less-studied distinctions such as socioeconomic status, it can perform classification and find correlations for which the meaning is not obvious.

Level correction improved accuracy on some classifications, showing promise. However, the points marked may not have been ideal. Given the results, it may be worth developing a more advanced level correction system for GNA.

FIRST and BET both improve accuracy in at least some classifications. BET in particular introduced a substantial amount of noise by not removing all skull tissue from some images. Still, when classifying by gender, without use of these components no DNRs could be discovered.

Acknowledgments

The Open Access Structural Image Series project is funded under the following grants: P50 AG05681, P01 AG03991, R01 AG021910, P20 MH071616, U24 RR021382.

References

Alley, D., Suthers, K., and Crimmins, E. Education and cognitive decline in older americans. *Research on aging*, 29(1):73, 2007.

Burns, A., Luthert, P., Levy, R., Jacoby, R., and Lantos, P. Accuracy of clinical diagnosis of alzheimer's disease. *British Medical Journal*, 301(6759):1026, 1990.

Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehéricy, S., and Habert, M.O. Automatic classification of patients with alzheimer's disease from structural mri: A comparison of ten methods using the adni database. *Neuroimage*, 2010.

Deshpande, Mukund, Kuramochi, Michihiro, Wale, Nikil, and Karypis, George. Frequent substructure-based approaches for classifying chemical compounds. *IEEE Transactions on Knowledge and Data*

Engineering, 17(8):1036–1050, 2005. ISSN 1041-4347.

Elsayed, A., Coenen, F., Jiang, C., García-Fiñana, M., and Sluming, V. Corpus callosum MR image classification. *Knowledge-Based Systems*, 23(4):330–336, 2010.

Goldstein, J.M., Seidman, L.J., Horton, N.J., Makris, N., Kennedy, D.N., Caviness, V.S., Faraone, S.V., and Tsuang, M.T. Normal sexual dimorphism of the adult human brain assessed by in vivo magnetic resonance imaging. *Cerebral Cortex*, 11(6):490, 2001.

Hollingshead, A.B. Two factor index of social position. 1957.

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

Long, S. and Holder, L.B. Graph-based shape analysis for mri classification. *International Journal of Knowledge Discovery in Bioinformatics (IJKDB)*, 2 (2):19–33, 2012a.

Long, S.S. and Holder, L.B. Graph based mri brain scan classification and correlation discovery. In *Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), 2012 IEEE Symposium on*, pp. 335–342. IEEE, 2012b.

Maguire, E.A., Gadian, D.G., Johnsrude, I.S., Good, C.D., Ashburner, J., Frackowiak, R.S.J., and Frith, C.D. Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences*, 97(8):4398, 2000.

Maguire, E.A., Spiers, H.J., Good, C.D., Hartley, T., Frackowiak, R.S.J., and Burgess, N. Navigation expertise and the human hippocampus: a structural brain imaging analysis. *Hippocampus*, 13(2):250–259, 2003.

Marcus, D.S., Wang, T.H., Parker, J., Csernansky, J.G., Morris, J.C., and Buckner, R.L. Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *Journal of Cognitive Neuroscience*, 19(9):1498–1507, 2007.

Millman, K.J. and Brett, M. Analysis of functional magnetic resonance imaging in Python. *Computing in Science & Engineering*, pp. 52–55, 2007. ISSN 1521-9615.

Muller, K.R., Mika, S., Ratsch, G., Tsuda, K., and Scholkopf, B. An introduction to kernel-based learning algorithms. *IEEE transactions on neural networks*, 12(2):181–201, 2001.

Nestor, S.M., Rupsingh, R., Borrie, M., Smith, M., Accomazzi, V., Wells, J.L., Fogarty, J., and Bartha, R. Ventricular enlargement as a possible measure of alzheimer's disease progression validated using the alzheimer's disease neuroimaging initiative database. *Brain*, 131(9):2443, 2008.

Patenaude, B., Smith, S., Kennedy, D., and Jenkinson, M. A bayesian model of shape and appearance for subcortical brain segmentation. *NeuroImage*, 2011.

Raz, N., Lindenberger, U., Rodrigue, K.M., Kennedy, K.M., Head, D., Williamson, A., Dahle, C., Gerstorf, D., and Acker, J.D. Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. *Cerebral Cortex*, 15(11): 1676–1689, 2005.

Smith, S.M. Fast robust automated brain extraction. *Human brain mapping*, 17(3):143–155, 2002.

Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., et al. Advances in functional and structural mr image analysis and implementation as fsl. *Neuroimage*, 23:S208–S219, 2004.

Tuomi, I. The lives and death of moore's law. *First Monday*, 7(11), 2002.

Wilson, R.S., Li, Y., Aggarwal, NT, Barnes, LL, McCann, JJ, Gilley, DW, and Evans, DA. Education and the course of cognitive decline in alzheimer disease. *Neurology*, 63(7):1198, 2004.

Woolrich, M.W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M., and Smith, S.M. Bayesian analysis of neuroimaging data in fsl. *NeuroImage*, 45(1):S173–S186, 2009.