Annals of the BMVA Vol. 2007, No. 1, pp 1–14 (2007)

# Classification of MRI Brain Scan Data Using Shape Criteria

Ashraf Elsayed[1], Frans Coenen[1], Marta García-Fiñana[2] and Vanessa Sluming[3]

[1]Department of Computer Science, [2]Department of Health Sciences, [3]Institute of Translational Medicine, The University of Liverpool, Liverpool, L69 3BX, UK

⟨{a.el-sayed,coenen,m.garciafinana,vanessa.sluming}@liv.ac.uk⟩

**Abstract**

Two mechanisms for classifying Magnetic Resonance Image (MRI) brain scans according to the nature of the corpus callosum are described. The first mechanism uses a hierarchical decomposition approach whereby each MRI scan is decomposed into a hierarchy of "tiles" which can then be represented as a tree structure (one tree per scan). A frequent sub-graph data mining mechanism is then applied so that sub-graphs that occur frequently across the image set are identified. These frequent sub-graphs can be viewed as describing a feature space; as such the input images can be translated, according to this feature space, into a set of feature vectors (one per image) to which standard classification techniques can be applied. The second approach uses a time series mechanism to represent the corpus callosum in each image. Using this representation a pre-labelled training set was used to define a Case Base (CB) to which Case Based Reasoning (CBR) techniques can be applied so as to classify new cases. Extremely accurate results were obtained with respect to datasets used for evaluation purposes.

## 1 Introduction

This paper describes and compares two approaches to classifying (catagorising) MRI brain scans according to the nature of the corpus callosum, a structure within the mammalian brain that connects the two hemispheres. The first approach is founded on the concept of graph mining and the second on time series analysis. Both approaches, although operating in very different manners, are essentially supervised learning mechanisms whereby a pre-labelled training set is used to build a "classifier" which can then be applied to unseen data. The first approach uses a hierarchical decomposition technique coupled with a tree based representation, one tree per image. A graph mining technique is then applied to identify frequently occurring sub-graphs (sub-trees) within this tree representation. The identified frequent subtrees can be viewed as defining a feature space which can be used to represent the image set. The image set is thus recast into this format so that each image is represented by a feature vector whose elements are some subset of the global set of identified frequent

Annals of the BMVA Vol. 2007, No. 1, pp 1–14 (2007)

sub-trees making up the feature space. Standard classifier generation techniques can then be applied to build a classifier that can be applied to unseen data. The second approach is founded on a time series representation coupled with a Case Based Reasoning (CBR) mechanism. In this approach the features of interest are represented as time series, one per image. These time series are then stored in a Case Base (CB) which can be used to categorise unseen data using a Case Based Reasoning (CBR) approach. The unseen data is compared with the categorisation in the CB using a Dynamic Time Warping (DTW) similarity checking mechanism, the categorisation associated with the most similar time series (case) in the CB is then adopted as the categorisation for the unseen data. The work described builds upon earlier work reported in [Elsayed et al., 2010a].

The rest of this paper is organised as follows. Section 2 describes the MRI application domain in the context of the corpus callosum. The start point for the two described techniques is a segmented Region Of Interest (ROI), the corpus callosum in this case. It should be noted that the objective of this paper is not to propose a new segmentation algorithm, indeed any appropriate ROI segmentation algorithm will suffice. However, for completeness, the segmentation algorithm used by the authors (a graph-based algorithm) is outlined in Section 3. The two proposed classification approaches are then described in Sections 4 and 5 respectively. The two approaches are then evaluated and compared in Section 6 and some conclusions are drawn in Section 7. The most noteworthy aspect of the work is the high accuracy obtained by both techniques.

## 2 Application Domain

The work described in this paper is directed at the classification of MRI brain scan data according to the corpus callosum. This is a highly visible structure in MR images whose function is to connect the left and right hemispheres of the brain, and to provide the communication conduit between these two hemispheres. In Figure 1, the left-hand image gives an example MRI scan; the corpus callosum is located in the center of the image, the corpus callosum has been highlighted in the right-hand image for ease of understanding [1]. A related structure, the *fornix* is also indicated. The fornix often "blurs" into the corpus callosum and thus presents a particular challenge in the context of the segmentation of these images.

The corpus callosum is of interest to medical researchers for a number of reasons. The size and shape of the corpus callosum have been shown to be correlated to sex, age, neurodegenerative diseases and various lateralized behaviour in people. It is also conjectured that the size and shape of the corpus callosum reflects certain human characteristics (such as mathematical or musical ability). Several medical studies indicate that the size and shape of the corpus callosum, in humans, are correlated to (for example) brain growth and degeneration [Hampel et al., 1998], handedness [Cowell et al., 1993] and epilepsy [Conlon and Trimble, 1988, Riley et al., 2010, Weber et al., 2007]. Although the work described in this paper is directed at representations (models) to support the application of classification processes, some work on modelling the corpus callosum with respect to other applications has been reported. For example [Stegmann et al., 2004, 2006] described a method for automatically analysing and segmenting the corpus callosum using Active Appearance Models (AAMs).

[1]The highlighting has been included simply to help readers identify the corpus callosum, it does not indicate the result of the application of some segmentation technique.

Annals of the BMVA Vol. 2007, No. 1, pp 1–14 (2007)

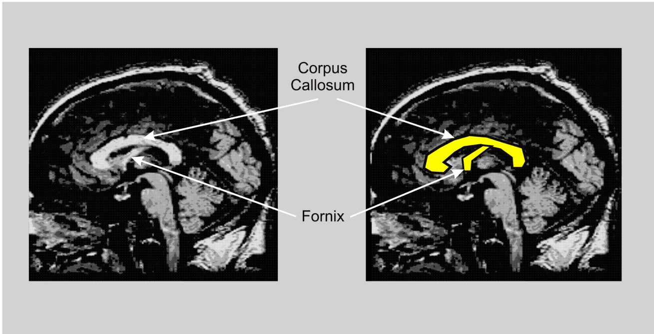


Figure 1: corpus callosum in a midsagittal brain MR image.

## 3 Registration and Segmentation

MRI brain scans comprised a sequence of "image slices", we refer to this as a *bundle*. The raw dataset used to evaluate the techniques described in this paper consisted of collections of MRI scan bundles. For the mechanisms described in this paper to operate we only required the middle slice from each bundle. This is referred to as the *midsagittal slice* and is the slice that separates the left and the right hemispheres of the brain. It should be noted that as a part of the collection process, all slices in all bundles were aligned so that each bundle was centered on the same axes. The alignment (registration) was conducted manually by trained physicians using the Brain Voyager QX software package [Goebel et al., 2006]. Figure 2 shows a typical MRI brain scan registered to a "standard" coordinate system using the Brain Voyager QX software package.

When attempting to categorise images according the nature of a ROI, regardless what technique is to be used, the first task is to identify and isolate the feature of interest. In the case of the corpus callosum we know, approximately, where it is located with respect to the boundaries of an MRI brain scan. Thus we can apply a segmentation algorithm to identify the corpus callosum pixels. As noted above the nature of the segmentation algorithm used is not the focus of this paper. This paper is directed at the evaluation of two mechanisms for classifying MRI brain scans according to a particular ROI (the corpus callosum in this case). Although different results may be obtained using different segmentation techniques, it is the relative performance of the two techniques that is of interest here. However, for completeness this section briefly describes the segmentation techniques used (a graph-based approach) and suggests some alternative segmentation techniques that can be usefully employed.

For the work described in this paper the Efficient Graph-based Segmentation (EGS) algorithm proposed in [Felzenszwalb and Huttenlocher, 2004] was used. This method is based on Minimum Spanning Trees (MST). All pixels of the original image are viewed as separate components. Two components are merged if the external variation between the components is small compared to the internal variation. Note that the segmentation can be problematic as a related tissue structure, the Fornix (also shown in the example given in Figure 1) is often included together with some other spurious pixel clusters. Some data cleaning must

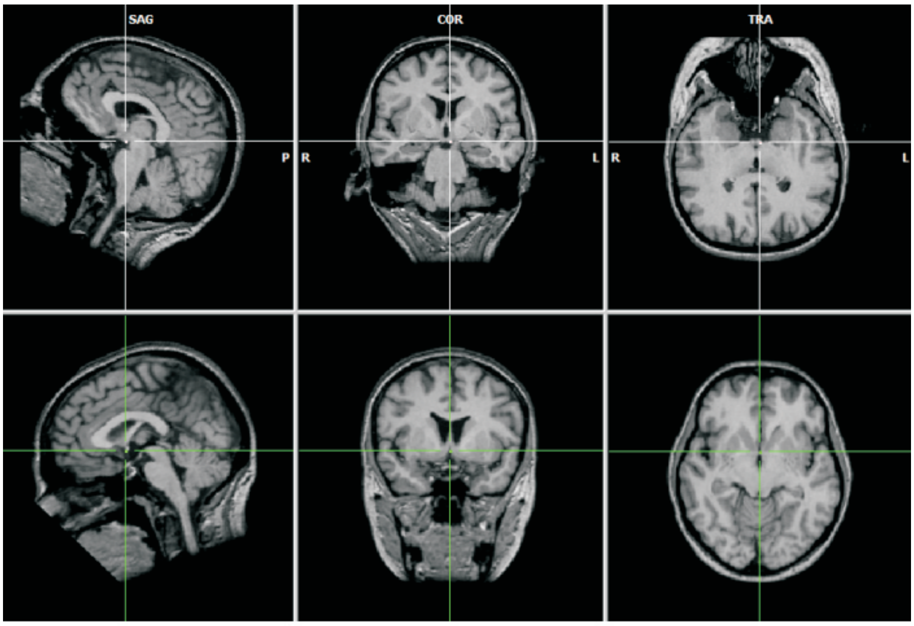


Figure 2: MRI brain scan registration.

therefore be undertaken. A smoothing technique was first applied to the MR images before the application of segmentation but so as to preserve the boundaries between regions. This smoothing operation is fully described in [Elsayed et al., 2010b]. In summary the smoothing was founded on the observation that the corpus callosum pixel intensity values follow the normal distribution with mean $\overline{X} = 160$ and standard deviation $s = 20$. With a threshold interval set at $\overline{X} \pm 3s$ it was found that the corpus callosum was clearly defined. The significance of this was that although the threshold values may differ depending on individual images, the high intensity property of the corpus callosum can be exploited to yield a segmentation algorithm that is both effective and efficient across the input image set. Therefore the interval $\overline{X} \pm 3s$ was used to exclude intensity values outside the interval. This strategy was incorporated into EGS segmentation algorithm and used to successfully extract the corpus callosum

Although with respect to this paper we have used the EGS algorithm, alternative segmentation techniques could have been applied such as the Normalized Cuts [Shi and Malik, 2000] or Multiscale Normalized Cuts [Cour et al., 2005] graph-based algorithms, or approaches popularised in computer vision systems such as the active contour or snake model [Kass et al., 1988].

## 4 Graph-Based Approach

The proposed graph based classification process commences with a segmentation phase, as described above, so as to isolate the corpus callosum in each image. The pixel represented corpus callosum is then *tesselated* into homogenous sub-regions. The tessellation process entails the recursive decomposition of the ROI, into quadrants. The tesselation continues until

Annals of the BMVA Vol. 2007, No. 1, pp 1–14 (2007)

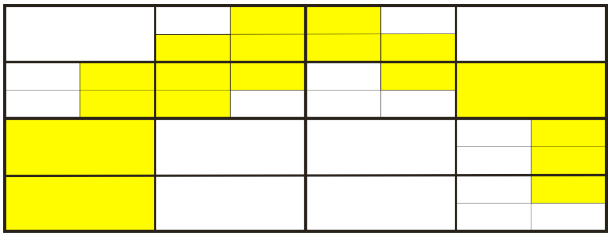

Figure 3: Hierarchical decomposition (tesselation) of Corpus Callosum.

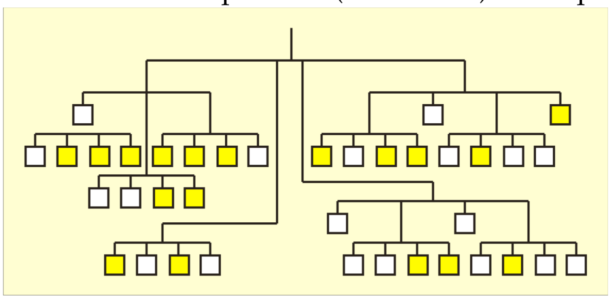

Figure 4: Tree representation of hierarchical decomposition.

either sufficiently homogenous quadrants are identified or some user specified level of granularity is reached. The result is then stored in a quadtree data structure such that each leaf node represents a *tile* in the tesselation. Nodes nearer the root of the tree represent larger tiles than nodes further away. Thus the tree is "unbalanced" in that some root nodes will cover larger areas of the ROI than others. It is argued that tiles covering small regions are of greater interest than those covering large regions because they indicate a greater level of detail (as expected these are located on the boundary of the ROI). The advantage of the representation is thus that it maintains information about the relative location and size of groups of pixels (i.e. the shape of the corpus callosum). The decomposition process is illustrated in Figure 3 and Figure 4. Figure 3 illustrates the decomposition (in this case down to a level of 3). Figure 4 illustrates the resulting quadtree.

A weighted frequent sub-graph mining technique was developed to identify commonly occuring sub-trees within the quadtree represented image set. Frequent sub-graph mining is a branch data mining concerned with the identification of sub-graphs that frequently occur across a graph represented data set. The input to a frequent sub-graph mining algorithm is a collection of graphs $G$ (in our case $G$ comprises a collection of trees each representing a corpus callosum). The sub-graph is considered to be frequent if its occurrence count, $s$ (referred to as its *support*) is greater than or equal to some user specified *support threshold* $\sigma$. The value $s$ for a specific candidate frequent sub-graph is the number of graphs in $G$ in which it occurs (a maximum count of one per graph). The value of $\sigma$ is then expressed as a percentage of the number of graphs in $G$, typically the value of $\sigma$ is low (1% or 2%) so that no significant sub-graphs are missed. Frequent sub-graph mining algorithms typically proceed in an "Apriori manner" starting with one edge candidate sub-graphs, and proceeding to two edge sub-graphs and so until there are no more sub-graphs to be discovered. At each iteration $k$, the $s$ values are determined for each $k$ sized candidate sub-graph and those graphs whose $s$ value is less than $\sigma$ are removed (pruned). On the next $k+1$ iteration knowledge of

Annals of the BMVA Vol. 2007, No. 1, pp 1–14 (2007)

the identified $k$ sub-graphs is used to generate the $k+1$ set of candidate sub-graphs.

Note that the lower the value of $\sigma$ the greater the number of frequent sub-graphs that will be identified. Because low values of $\sigma$ are typically used a great number of frequent sub-graphs may be identified. In many cases the number of sub-graphs is unmanageble. However, many of the discovered subgraphs are often found to be redundant (subsets of other graphs). To address this issue weighting schemes have been produced so that only significant frequent sub-graphs are discovered. In our case the weightings were calculated according to the reverse distance of individual nodes to the root node in each tree. This weighting concept was built into a variation of the well known gSpan algorithm [Yan and Han, 2002]. The algorithm operates in an Apriori manner, level by level, following the "generate, calculate support, prune" loop described above. A detailed description of this weighted sub-graph mining algorithm adopted with respect to the work described in this paper can be found in [Jiang and Coenen, 2008] and [Jiang et al., 2010]. Frequent sub-graph mining is a substantial topic within the domain of data mining and any more detailed discussion is beyond the scope of this paper. However a detailed review of the subject can be found in [Jiang et al., 2013].

The identified frequent sub-trees (graphs) each describing, in terms of size and shape, some part of a corpus callosum that occurs regularly across the data set, are then used to form the fundamental elements of a *feature space*. In this context a feature space is an $N$ dimensional space where $N$ is equivalent to the number of features and each feature is a numerically valued attribute. In our case each feature is a frequently occurring sub-graph with the values 0 and 1 associated with it (0 if it is absent in a particular image, and 1 if it is present), we say that the attributes are "binary valued". Using this feature space each image (corpus callosum) can be described in terms of a feature vector of length $N$, with each element corresponding a particular feature (sub-graph) and having either the value 0 or 1 (thus the image set can be described in terms of a set of binary valued vectors) .

As noted above the graph mining process typically identifies a great many frequent sub-graphs; more than that required for the desired classification. Therefore a feature selection strategy is applied to the feature space so that only those sub-graphs that serve as good discriminators between classes are retained. A straightforward wrapper method was adopted whereby a decision tree generator was applied to the feature space. Features included as "choice points" in the decision tree were then selected[2], while all remaining features were discarded. For the work described here, the well established C4.5 algorithm [Quinlan, 1993] was adopted, although any other decision tree generator will suffice. On completion of the feature selection process each image is described in terms of a reduced binary-valued feature vector indicating the selected features (sub-graphs) that appear in the image. Once the image set has been represented in this manner any appropriate classifier generator may be applied. For additional information regarding the graph based approach, including the tesselation process, interested readers are referred to [Elsayed et al., 2010b]. With respect to the work described here, the C4.5 decision tree generator was again used to produce the desired classifier. Readers wishing to gain an additional insight to decision tree classifiers are referred to [Rokach and Maimon, 2008].

[2]Decision trees are a type of decision support tool that use a tree model of decisions and their possible outcomes. The nodes in the decision tree are referred to as "choice points".

Annals of the BMVA Vol. 2007, No. 1, pp 1–14 (2007)

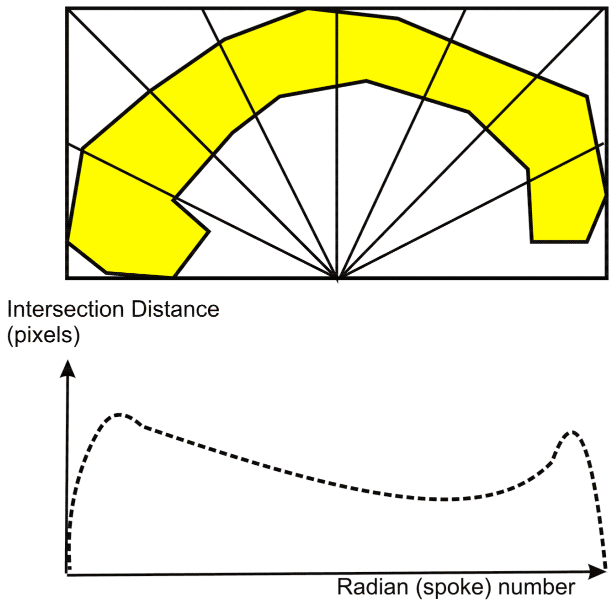


Figure 5: Conversion of corpus callosum into time series.

## 5 Time Series Based Approach

As in the case of the graph based approach, the time series based approach commences with the segmentation and registration of the input images. Note that in the context of this paper the precise nature of the ROI segmentation technique is not significant, although for the work described we used graph based image segmentation technique proposed by Felzenszwalb and Huttenlocher [Felzenszwalb and Huttenlocher, 2004]. Once the ROI have been segmented and identified the next step is to derive a time series according to the boundary line circumscribing the corpus callosum. Note the phrase "time series" is used with respect to the adopted representation because the proposed corpus callosum classification technique is founded on work in time series analysis, not because the representation includes some temporal dimension.

Using the proposed technique the time series is generated using an ordered sequence of $M$ "spokes" radiating out from a single reference point. The derived time series is then expressed as a series of values (one for each spoke) describing the size (length) of intersection of the vector with the ROI. The representation thus maintains the structural information (shape and size) of the corpus callosum. It should also be noted that the value of $M$ may vary due to the differences of the shape and size of the individual ROI within the image data set.

With respect to the corpus callosum application the time series generation procedure is illustrated in Figure 5. The midpoint of the lower edge of the Minimum Bounding Rectangle (MBR) was selected as the reference point. This was chosen as this would ensure that there was only one intersection per spoke. The vectors were derived by rotating an arc about the reference point pixel. The interval between spokes was one pixel measured along the edge

Annals of the BMVA Vol. 2007, No. 1, pp 1–14 (2007)

Table 1: Details of datasets used.

| Data Set | TE (ms) | TR (ms) | Flip Ang.$^\circ$ | FOV ($mm^2$) | # Slices | Voxel Size (mm) |
|---|---|---|---|---|---|---|
| Musicians & Epilepsy | 9.0 | 34 | 30 | 200 | 192 | $0.781 \times 0.781 \times 1.6$ |
| Handedness | 5.57 | 2040 | 8 | 256 | 176 | $1 \times 1 \times 1$ |

of the MBR. For each spoke, the distance $D_i$ (where $i$ is the spoke identification number) was measured over which the spoke intersected with the corpus callosum pixels. The result was a time series with the spoke number $i$ representing time and the value $D_i$, for each spoke, the magnitude. By plotting $D_i$ against $i$ a time series was derived as shown in Figure 5.

Each time series is then conceptualised as a *proto-type* or case contained in a Case Base (CB), to which a Case Based Reasoning (CBR) mechanism can be applied. CBR is a branch of Artificial Intelligence (AI) founded on the idea that humans solve problems according to their experience, i.e. CBR conjectures that humans solve problems by attempting to match previous successfully addressed problems to the current problem. As such a CBR system comprises a Case Base (CB) and some matching strategy to align a new problem (case) with with previously solved problems (cases) in the CB. Typically it will not be possible to find an exact match and thus the matching strategy will have to find the most relevant case or cases. The CBR community has proposed many techniques to identify the desired best match, and derivation of optimum matching strategies remains a topic of research with the domain of CBR. Case Based Reasoning (CBR) has a well established body of literature associated with it. Recommended reference works include [Leake, 1996] and [Kolodner, 1993]. For a review of the application of CBR in medical domains see [Bichindaritz and Marling, 2006] or [Holt et al., 2005].

CBR can be used for classification purposes [Pal et al., 2011] where, given an unseen record (case), the record can be classified according to the "best match" discovered in the CB. With respect to the corpus callosum application, the CB comprises a set of pre-labelled (classified) time series, each describing a corpus callosum record. A time series matching strategy was then adopted to identify a best match with a new ("unseen") corpus callosum time series. More specifically a Dynamic Type Warping (DTW) time series analysis technique for comparing curves [Berndt and Clifford, 1994] has been adopted. The advantage offered by DTW is that it is able to find the optimal alignment between two time series $Q$ and $C$, of length $n$ and $m$ respectively where $n$ does not necessarily have to be equal to $m$.

DTW operates as follows. Given a *query sequence* $Q = \{q_1, q_2, \ldots, q_i, \ldots, q_n\}$, which we wish to compare with a *comparitor sequence* $C = \{c_1, c_2, \ldots, c_j, \ldots, c_m\}$, with the aim of (say) classifying Q. These two sequences can be compared by first constructing a $n \times m$ grid (matrix) such that the value for element $< i, j >$ is the squared euclidean distance from point $c_j$ on curve $C$ to point $q_i$ on curve $Q$. If $Q$ and $C$ are identical the values at grid points $< i, j >$, where $i = j$, will be zero. The best match between the two sequences $Q$ and $C$ is the warping path that minimises the total cumulative distance (grid values) from $< 0, 0 >$ to $< n, m >$. A warping path is any contiguous set of matrix elements from $< 0, 0 >$ to $< n, m >$. The warping cost associated with a particular path is its cumulative distance. Given two identical series the warping path will be zero.

# 6 Evaluation

To evaluate and compare the two proposed approaches three scenarios were considered: distinguishing between musicians and non-musicians, left-handedness and right-handedness, and epilepsy patients and healthy subjects. For the musicians study, a data set comprising 106 MR images was used, 53 representing musicians and 53 non-musicians (i.e. two equal classes). The scans were obtained using a Siemens 1.5 Tesla scanner. The study was of interest because of the conjecture that the size and shape of the corpus callosum reflects human abilities (such as a mathematical or musical ability). There is significant evidence, amongst the medical community, that traits such as musical ability, influence the shape and size of the corpus callosum. It should be noted that a visual inspection of the MR images does not indicate any discernible distinction between the two categories. For the handedness study, a data set comprising 82 MR images was used, 42 representing right-handed and 40 left-handed. The data was obtained using a Siemens Trio 3 Tesla whole body MRI system. The study was of interest because of the conjecture that the size and shape of the corpus callosum reflects certain human characteristics (such as handedness). For the epilepsy study, a data set comprising 212 MR images was used. The data set comprised the 106 MR images used for the musicians study, augmented with 106 epilepsy cases. The latter were also obtained using a Siemens 1.5 Tesla scanner. The objective was to seek support for the conjecture that the shape and size of the corpus callosm is influenced by conditions such as epilepsy ([Conlon and Trimble, 1988, Riley et al., 2010, Weber et al., 2007]). In all cases the data sets were balanced in terms of age, sex etc. To the best knowledge of the authors the musicians study did not include any epilepsy patients. Some further background details concerning the data sets is given in Table 1.

Table 2: TCV Classification accuracy (%) for musicians study using GB and TSB approaches.

| Test set ID | GB | TSB |
|---|---|---|
| 1 | 100 | 91 |
| 2 | 100 | 100 |
| 3 | 91 | 91 |
| 4 | 91 | 100 |
| 5 | 100 | 100 |
| 6 | 90 | 100 |
| 7 | 100 | 100 |
| 8 | 90 | 100 |
| 9 | 91 | 100 |
| 10 | 100 | 100 |
| Average | 95.3 | 98.2 |
| SD ($\sigma$) | 4.97 | 3.79 |

Table 3: TCV Classification accuracy (%) for handedness study using GB and TSB approaches.

| Test set ID | GB | TSB |
|---|---|---|
| 1 | 100 | 100 |
| 2 | 88 | 100 |
| 3 | 89 | 100 |
| 4 | 100 | 89 |
| 5 | 88 | 88 |
| 6 | 88 | 88 |
| 7 | 100 | 100 |
| 8 | 88 | 100 |
| 9 | 100 | 100 |
| 10 | 100 | 100 |
| Average | 94.1 | 96.5 |
| SD ($\sigma$) | 6.23 | 5.64 |

Ten-fold Cross Validation (TCV) was used through out the evaluation. TCV is a well established statistical evaluation technique on the lines of "leave one out". Given a data set we divide it into tenths and then run the evaluation 10 times, testing on a different 1/10th each time, and training on the remaining 9/10ths. Thus, in the case of the musicians data

Table 4: TCV Classification accuracy (%) for epilepsy study using GB and TSB approaches.

| Test set ID | GB | TSB |
|---|---|---|
| 1 | 91 | 82 |
| 2 | 86 | 77 |
| 3 | 90 | 81 |
| 4 | 86 | 76 |
| 5 | 95 | 86 |
| 6 | 81 | 71 |
| 7 | 90 | 81 |
| 8 | 86 | 71 |
| 9 | 77 | 71 |
| 10 | 81 | 76 |
| Average | 86.3 | 77.2 |
| SD ($\sigma$) | 5.46 | 5.26 |

Table 5: TCV Classification accuracy (%) using the graph based technique with different levels of decomposition (musicians study).

| | Support Threshold (%) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Levels | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 |
| 4 | 71 | 70 | 69 | **72** | 69 | 62 | 53 | 51 |
| 5 | **91** | 84 | 80 | 86 | 80 | 81 | 80 | 71 |
| 6 | 86 | **95** | 85 | 84 | 91 | 84 | 77 | 75 |
| 7 | 84 | 86 | **90** | 87 | 88 | 75 | 76 | 78 |

set, the test set will comprise 10 or 11 records and the training set the remainder. The idea is that TCV will smooth out any irregularities in the ordering of the data.

Table 2 presents TCV classification results for the musicians study obtained using the proposed techniques. The columns labelled GB (Graph Based) and TSB (Time Series Based) indicate the classification accuracy obtained for each tenth. With respect to the GB approach a quad tree of depth 6 (decomposition level), coupled with a 30% support threshold for the graph mining, produced the best classification accuracy. Note that with respect to Table 2 the test set comprised either 10 or 11 records, thus for each test run all or all but one of the test cases were classified correctly. Table 3 shows the TCV classification results with respect to the handedness data and Table 4 the results obtained with respect to the epilepsy data.

Table 6: TCV Classification accuracy (%) using graph based technique with different levels of decomposition (handedness study).

| | Support Threshold (%) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Levels | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 |
| 4 | 67 | 68 | 70 | **71** | 67 | 60 | 51 | 49 |
| 5 | 78 | 83 | **89** | 84 | 79 | 78 | 78 | 69 |
| 6 | 84 | **94** | 89 | 84 | 83 | 79 | 76 | 74 |
| 7 | 83 | 84 | **88** | 87 | 85 | 77 | 74 | 72 |

Annals of the BMVA Vol. 2007, No. 1, pp 1–14 (2007)

Table 7: TCV Classification accuracy (%) using graph based technique with different levels of decomposition (epilepsy study).

| | Support Threshold (%) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Levels | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 |
| 4 | 59 | 61 | **63** | 60 | 60 | 52 | 44 | 42 |
| 5 | 77 | **82** | 77 | 75 | 72 | 68 | 62 | 59 |
| 6 | 82 | **86** | 80 | 76 | 73 | 67 | 65 | 62 |
| 7 | 75 | 77 | 79 | **80** | 78 | 70 | 63 | 60 |

From the tables (2, 3 and 4) it can be seen that excellent results were obtained throughout. The time series based approach produced the best classification results, in terms of accuracy, with respect to the musicians and handedness studies, while the graph based approach produced the best results with respect to the epilepsy study. Best overall classification accuracy results were obtained using the musicians study (Table 2); for the majority of the TCV runs a 100% accuracy was obtained using the time series based approach. Good results were also obtained with respect to the handedness study (Table 3) with some TCV runs producing 100% accuracies (again using the time series based approach). The techniques did not perform as well for the epilepsy study (Table 4) although the 86% overall classification accuracy obtained using the graph based approach was still reasonable (significantly better than chance). The suspicion here is that results reflect the fact that although the nature of the corpus callosum may play a part in the identification of epilepsy there are also other factors involved. Based on the data there is not sufficient statistical evidence to conclusively suggest that the TSB approach provides better accuracy than the GB approach for the musicians and handedness data sets (P-values>0.05). On the other hand, statistical comparison indicates that the GB approach provides better accuracy than than the TSB approach for the epilepsy data set (P<0.01).

Tables 5, 6 and 7 give some further evaluation results, using the graph based technique, with respect to the musicians, handedness and epilepsy studies. The tables present the TCV accuracy results obtained using a variety of quad-tree depths and support thresholds. From the table it can be seen that a decomposition level of 6 coupled with a support threshold of 30% seem to be the most appropriate values in the context of classification accuracy. These were also the values used for the experiments reported in Tables 2, 3 and 4).

# 7 Conclusions

Two approaches to the classification of MRI brain scans according to the nature of the corpus callosum, founded on graph mining and time series analysis respectively, have been described. The most noteworthy element of the work is the high classification accuracy obtained for both approaches; in terms of accuracy the time series approach out performs the graph based approach in the case of the musicians and handedness data sets, and the graph based approach produced the best result in the case of the epilepsy data set. The results were of particular interest because visual inspection of the segmented images indicated that there was no discernible distinction between the images. The research team is currently working on mechanisms whereby "explanations" can be generated to describe the reasons for particular classifications in terms of the nature of the corpus callosum. The intention is to generate explanations so that clinicians can determine why a certain classification arose, not to invent reasons to justify results. For example we might want to highlight that a particular classification occurred because of some feature on the time series curve or because of the presence of some protuberance on the corpus callosum indicated by a particular type of sub-graph. Alternative mechanisms to those described in this paper, whereby the quantitative aspects of the structure of image objects can be described, are reported in the literature; one example is Geometric Texton Theory (GTT) [Griffin et al., 2004, Griffin, 2005]. These alternative mechanisms may provide further fruitful means where by the nature of MRI brain scan objects can be captured for the purpose of input to classification algorithm, and thus will also merit further investigation.

## 8 Acknowledgements

The authors would like to thank Ms Joanne Powell of the Department of Eye and Vision Science at the University of Liverpool for her support with respect to the collation and preperation of the handedness data set used to support the evaluation of the work described in this paper.

## References

D. Berndt and J. Clifford. Using dynamic time warping to find patterns in time series. In *Proc. AAAI-94 workshop on Knowledge Discovery in Databases*, pages 359–370, 1994.

I. Bichindaritz and C. Marling. Case-based reasoning in the health sciences: What's next? *Artificial Intelligence in Medicine*, 36(2):127–135, 2006.

P. Conlon and M. Trimble. A study of the corpus callosum in epilepsy using magnetic resonance imaging. *Epilepsy Res*, 43:122–126, 1988.

T. Cour, F. Benezit, and J. Shi. Spectral segmentation with multiscale graph decomposition. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 1124 – 1131 vol. 2, 2005.

P. Cowell, A. Kertesz, and V. Denenberg. Multiple dimensions of handedness and the human corpus callosum. *Neurology*, 43:2353–2357, 1993.

A. Elsayed, F. Coenen, M. García-Fiñana, and V. Sluming. MRI brain scan classification according to the nature of the corpus callosum. In *Proc. Medical Image Understanding and Analysis (MIUA'10)*, pages 19–23, 2010a.

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

P. Felzenszwalb and D. Huttenlocher. Efficient graph-based image segmentation. *Int. journal of Computer Vision*, 59(2):167–181, 2004.

R. Goebel, F. Esposito, and E. Formisano. Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: From single-subject to cortically aligned group general

Annals of the BMVA Vol. 2007, No. 1, pp 1–14 (2007)

linear model analysis and self-organizing group independent component analysis. *Human Brain Mapping*, 27(5):392–401, 2006.

L.D. Griffin. Geometric texton theory: The 1-d, 2nd-order jet. *Perception*, 34:246–247, 2005.

L.D. Griffin, M. Lillholm, and M. Nielsen. Natural image profiles are most likely to be step edges. *Vision Research*, 44(4):407–421, 2004.

H. Hampel, S. Teipel, G. Alexander, B. Horwitz, D. Teichberg, M. Schapiro, and S. Rapoport. Corpus callosum atrophy is a possible indicator of region and cell type-specific neuronal degeneration in Alzheimer disease. *Archives of Neurology*, 55:193–198, 1998.

A. Holt, I. Bichindaritz, R. Schmidt, and P. Perner. Medical applications in case-based reasoning. *The Knowledge Engineering Review*, 20:289–292, 2005.

C. Jiang and F. Coenen. Graph-based image classification by weighting scheme. In *Proc. AI'2008, Springer*, pages 63–76, 2008.

C. Jiang, F. Coenen, and M. Zito. Frequent sub-graph mining on edge weighted graphs. In *12th Int. Conf. on Data Warehousing and Knowledge Discovery*, pages 77–88. Springer, LNCS 6263, 2010.

C. Jiang, F. Coenen, and M. Zito. *A Survey of Frequent Subgraph Mining Algorithms*. In press Knowledge Engineering Review, 2013.

M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int. Jo. Of Computer Vision*, 1(4):321–331, 1988.

J.L. Kolodner. *Case-based Reasoning*. Morgan Kaufmann Series in Representation and Reasoning, 1993.

D.B. Leake. *Case-based Reasoning: Experiences, Lessons and Future Directions*. AAAI Press Co-Publications, 1996.

S. Pal, D. Aha, and K. Gupta. *Case-Based Reasoning in Knowledge Discovery and Data Mining*. Wiley-Blackwell, 2011.

R. Quinlan. *C4.5: Programs for machine learning*. Morgan Kaufmann, 1993.

J. Riley, D. Franklin, V. Choi, R. Kim, D. Binder, S. Cramer, and J. Lin. Altered white matter integrity in temporal lobe epilepsy: Association with cognitive and clinical profiles. *Epilepsia*, 51:536–545, 2010.

L. Rokach and O. Maimon. *Data Mining With Decision Trees: Theory And Applications*. World Scientific Publishing, 2008.

Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22:888–905, 2000.

M.B. Stegmann, R.H. Davies, and C. Ryberg. Corpus callosum analysis using mdl-based sequential models of shape and appearance. In *Proc. International Symposium on Medical Imaging (SPIE'04)*, pages 612–619, 2004.

Annals of the BMVA Vol. 2007, No. 1, pp 1–14 (2007)

M.B. Stegmann, K. Sjöstrand, and R. Larsen. Sparse modeling of landmark and texture variability using the orthomax criterion. In *IProc. International Symposium on Medical Imaging (SPIE'06)*, pages 6–12, 2006.

B. Weber, E. Luders, J. Faber, S. Richter, C. Quesada, H. Urbach, P. Thompson, A. Toga, C. Elger, and C. Helmstaedter. Distinct regional atrophy in the corpus callosum of patients with temporal lobe epilepsy. *Brain*, 130:3149–3154, 2007.

X. Yan and J. Han. gspan: Graph-based substructure pattern mining. In *Proc. ICDM'02: 2nd IEEE Conf. Data Mining*, pages 721–724, 2002.