A Proposed Stratification Approach for MRI Images

Madhulika Bhatia¹*, Abhay Bansal¹, Divakar Yadav² and Priya Gupta³

¹Department of Computer Science and Engineering, Amity University, Noida - 201301, Uttar Pradesh, India; mbhadauria@amity.edu, abansal1@amity.edu
²Department of Computer Science and Engineering, Jaypee University, Noida - 201307, Uttar Pradesh, India; divakar.yadav0@gmail.com
³J S Hospital, Sector 132, Noida – 201304, Uttar Pradesh, India; drpriyagupta@gmail.com

Abstract

Background/Objectives: To classify brain MRI Images utilizing neural system. The methodology helps the specialist to robotize the examination of MRI Images and to group them into healthy and unhealthy one. Methods/Statistical Analysis: A methodology is proposed to recognize grey matter so that unhealthy MRI images can be ordered from healthy MRI Images. Classification of Brain MRI Images on the premise of Gray matter and white matter is done utilizing Neural Network pattern recognition. The novel methodology utilizes preprocessing strategies and various filters available in matlab for viable extraction and investigation of Brain MRI Images taken from BRAINIX and Neuroimaging data repository. Findings: The work has been tested on 227 variables on fifteen patients. Out of 15 samples, 70% of samples are taken for Training, 15% for validation and 15% for testing. Application/Improvements: This methodology is a simple alternative to be utilized as a part of doctor’s facilities and medicinal places to guide doctors as well as researchers.

Keywords: Image Processing, MRI Images, Neural Network, Pattern Recognition

1. Introduction

The programmed extraction procedure of tumor from MRIs is a difficult task¹. Numerous methodologies are there that perform tumor division. Voxel-based MRI studies have indicated that mind regions show changes when it is affected with tumor². Utilizing bisection techniques and contour choice calculations grey and white matter is extricated from mind regions³. It indicates that we can segregate tissues like grey and white matter among normal and abnormal MRI of brain. In this paper we have presented an approach based on a blend of various bisection strategies in Matlab. This determines a quantitative similarity based matrix from classified MRI dicom images taken from MRI databases like Neuroimaging data repository in the form of columns and rows. We have recovered a matrix of 217x120 from every subject of MRI and have applied the same technique on 15 subjects including normal and abnormal both Utilizing Segmentation methods and Feature based calculations we have isolated cerebrum zones having grey and white matter as 0’s and 1’s respectively. Grey matter is recognized as one and white matter is recognized as zero. With the help of Neural Network Pattern Recognition we have selected data followed by creation and training of a new network. Consequently we evaluated its performance using mean square error and confusion matrices. Utilizing neural network and pattern recognition classification of 15 subjects as inputs into a set of target categories has been performed. This provides an isolation of the subjects with normal MRI and abnormal MRI. Thus we have proposed a methodology for demarcating healthy and unhealthy MRI subjects on the considered fifteen subjects.

1.1 Segmentation and Classification

In numerous therapeutic explores and applications, division and arrangement of brain tissue in MRI is a troublesome undertaking⁴. Besides, manual division and order of MRI brain images is a lengthy procedure. MRI cerebrum images have their applications in the fields such
as examination, restorative examination, doctoring and curing and so on. Due to the presence of non-consistency and undesirable elements, pictures get defiled thereby leading to a complicated study of the brain structure. Robotized and Soft-processing procedures help in an efficient disposal of this issue. The proposed technique aims at eliminating the noise from MRI Images. Later the bisection technique and contour extraction is merged to extricate white and grey matter from the MRI images. Feature extraction then undergoes regulated grouping i.e. supervised learning and assumes a significant role in investigation of medicinal images. Locale of interest depicts specific images by taking into account the numerical elements from the element vectors. The separated grey and white matter recognized by string of zeros and ones is examined utilizing proposed the methodology.

In proposed a novel automated strategy can be used for characterizing tissues of brain using MRI. It uses pruning procedure and redid set for the purpose of training. Utilizing likelihood maps, set of tests are produced, examples assigned as erroneous are initially lessened utilizing mind based correlations and least spanning. KNN (k-Nearest Neighbour) classifier is utilized on right specimens to characterize them. The methodology is non parametric when changeability in picture quality is considered. Dissemination in tissue illumination is not mulled over. Acceptance tests are directed both quantitatively and subjectively on reproduced MRI information of 43 subjects.

Low level operations characterize brain tissue in MRI utilizing weighted (T1). Numerous operations such as sifting by dispersion, distinguishing the borders and performing morphology numerically are applied. Histograms of neighborhood are created utilizing incomplete volume tissue estimation model to gauge focuses. For every zone and mean tissue power and commotion difference quality are figured from worldwide picture. Voxel construct order is done in light of force standardized picture and utilizing posteriori classifier. Spatial properties of the brain are displayed utilizing incomplete volume tissue estimation model with a Gibbs former.

According to segmentation is done on the premise of contours determined utilizing Least Squares Support Vector Machines (LS-SVM). Radial Basis Function (RBF) pieces, straight and non-straight were contrasted and all current classifiers like SVM, Multi Layer Perceptron (MLP) classifier and KNN classifier. Results demonstrated that LSSVM classifier performed remarkably than every single other classifier tried.

In, sample and information preparing procedure consolidated with probabilistic neural system was proposed. For substantial measure of information arrangement techniques controlled by a single administrator were not legitimate to utilize. Genuine mistakes in arrangement were acquired as clamor made by administrator execution. The mix of neural systems and fuzzy rationale furthermore with probabilistic neural system were utilized. Elements were extricated in two stages utilizing the Principal component analysis and the PNN (Probabilistic Neural Network). The executions were assessed with probabilistic neural system which gave precise results for classification.

DAUB-4 wavelet system is utilized to characterize non tumor and tumor pictures. DAUB-4 gives better complexity to a picture which enhances effectively hanging signs of a picture and decreases the overhead. Principal component analysis is utilized for grouping best elements acquired. The components received from PCA investigation are given as information to support vector machines. Exploratory results demonstrate that among straight kernels and radial premise kernel, radial premise piece gages better exactness.

A therapeutic picture has numerous layers and perspectives to reduce the pictures to it becomesimportant to perceive sicknesses. Hence the pictures need to separated and denoised. Automated characterization of fit or unfit individual is proposed by watershed calculation with Euclidean separation classifier for quick reckoning. This method is utilized in collaboration with preprocessing and post handling technique applied on database comprising of both healthy and unhealthy examples of cerebrum pictures in dicomm design. Categorizing is a phase where objects are sorted into classes in the middle of normal and abnormal.

2. Proposed Methodology and Material

2.1 Subjects
Subjects with healthy and unhealthy MRI are taken from BRAINIX. Dicomm MRI of healthy subjects was taken from Neuroimaging data repository.

2.2 Methodology
We proposed a novel way to deal with MRI. We have taken
fifteen subjects from BRAINIX and Neuroimaging data repository out of which seven were healthy and eight were unhealthy. This area clarifies the proposed methodology and system. The proposed technique comprises of number of stages, for example, DICOMM change, grey transformation, use of filters, separating grey and white matter as 0 and 1’s. Later features in the form of numeric values are imported in neural network pattern recognition tool and of Neural Network Pattern Recognition then selected data is followed by creation and training of a new network. Then evaluation of its performance using mean square error and confusion matrices are done. The complete procedure is demonstrated in Figure 1. Correlation filter in matlab is utilized for separating grey and white matters. Sizes of edges are extricated utilizing relationship channel. We figured Dx and Dy for every pixel in picture. The size of edge is $\sqrt{(Dx^2 + Dy^2)}$. The range is represented as a component vector as 0 and 1.

Figure 1. Classification of MRI Images.

### 2.2.1 Load MRI Image of Subjects

MRI image in dicomm format is loaded. The data set consists of fifteen brains MRI in which seven images are healthy and eight images are unhealthy. The size of each image is $256 \times 256$. The dataset of subjects Brainix Neuro imaging data repository.

### 2.2.2 Preprocessing

The objective of preprocessing is that it enhances information of picture so that undesirable elements can be deleted and helps in improving some vital picture elements so that further computing can be done. By applying these operations picture elements can’t be expanded. Excess in pictures are evacuated by the preprocessing strategies.

### 2.2.3 Displaying Gray Scale Image

To show a grayscale picture in Matlabim show method is used. Exhibiting images is done by scaling the power values which is used into a grayscale shading. The pixel value is shown 0.0 (dark) if intensity is double if intensity is uint8 then pixel value 1.0 is displayed as white, and if the pixel qualities are in the middle of, then it is displayed as gray.

### 2.2.4 Creating 2-D Special Filters

In creation of two dimensional filter method called Imfilter is utilized for this. The closest fringe pixel is utilized for repeating the quality for deciding any pixel esteem outside the range. There are numerous alternatives like symmetric filters. The grey matter and white matter from the image is detected as 0 and 1’s. Unsharp contrast upgrading is done by three by three unsharp filter. To control the state of laplacian and to make laplacian filter negative parameter alpha is used. The range of parameter alpha is in between 0 to 1. We can also give default value of two. A method in matlab known as fspecial is used to create laplacian filter and can be represented as mathematical statement as shown in 1.1 and 1.2.

$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$

$\alpha / 4 \quad (1 - \alpha) / 4 \quad \alpha / 4$

$\alpha / 4 \quad (1 - \alpha) / 4 \quad -1 \quad (1 - \alpha) / 4$

$\nabla^2 = 4/(\alpha + 1) \quad \alpha / 4 \quad (1 - \alpha) / 4 \quad \alpha / 4$

### 2.2.5 Image Cropping and Segmentation

Unrelated parts of image are removed by focusing on region of interest. The two-dimensional convolution operation is utilized to examine pictures. On the premise of weighted normal of pixel in k x k neighborhood of the focal pixel another quality is attributed to a given pixel. Weights supplied in a square lattice are utilized to represent in convolution kernel. Image components can be shown as $f(x, y)$ is brought with components $h(\alpha, \beta)$ of the convolution portion and the output picture component $g(x, y)$ are shown to by the mathematical statement 1.3.

$g(x, y) = \sum_{i=-k}^{k} \sum_{j=-k}^{k} f_i(x, j) k(x - i, y - j)$

*Vol 8 (22) | September 2015 | www.indjst.org*
2.2.6 Segmentation

It’s the procedure of isolating image into numerous fragments. The target of division is to break down the image representation which is straightforward. In the division process representation of the items in images is exceptionally crucial that we can separate between the objects of interest and remaining part. The remaining piece of gathering is additionally alluded as the foundation.

2.2.7 Gray Matter and White Matter Extraction

Gray matter and white matter is extracted in the form of zeros’ and ones’ and shown in the form of matrix and saved in excel file for each subject.

2.2.8 Counting Number of Zeros and Ones

Matrix representation for zeros and ones is opted which was extracted in the form of Gray matter and white matter and stored in excel file for each subject. For counting number of zero’s and ones are row for each subject is considered.

2.2.9 Subtraction

Total number of one’s counted is subtracted from total number of zeros counted. Same is repeated for fifteen subjects.

2.2.10 Classification

The aim is to classify between healthy and unhealthy subjects on the basis of grey and white matter. Neural Network pattern recognition is used for classification in Matlab environment. Feed forward neural network is trained accordingly to implement Pattern recognition. Output patterns are associated with input patterns during training. After counting number of zero’s and ones we subtracted total number of one’s from total number of zero’s and taken as set of input vectors in the form of column arranged. A set of input vectors as columns is written in a matrix. Then another set of target vectors so that they indicate the classes to which the input vectors are assigned. We have used the approach when there are only two classes we set each scalar target value to either 1 or 0. The healthy subjects fall into target class containing 1 and unhealthy subjects fall into target class 0.

- This data set consists of fifty elements as input vectors.
- Inputs is a 227x15 matrix, representing 15 samples of 227 elements.
- Targets is a 1x15 matrix, representing 15 samples of 1 element.
- There are two elements in each target vector, because there are two categories (healthy or unhealthy) associated with each input vector.

3. Proposed Algorithm

Steps

Start

for each subsets Si(i=1 to n) /* n=number of subsets */
    Generates matrix P_i[m][n] with values 0 and 1
for each row j= 1 to m
    Count number of zero’s say P_j,x
    Count number of 1’s say P_j,y
    Subtract z = x-y
    a[i][j] = z
End of for loop
End of outer for loop

Apply statistical analysis

End

4. Neural Network Pattern Recognition

In the field of restorative and clinical work on imaging is an obligatory. Different image computing systems that procedure are available in clinical field. To dissect therapeutic pictures factually is customary in cutting edge medicinal research. Pattern recognition can be implemented by using a feed-forward as shown in Figure 2. neural network that has been trained accordingly. Association of outputs with input patterns are done to train the network during training. It’s the network when it is used, it recognize the input pattern and tries to output the conjoin output pattern.

Figure 2. Feed Forward Neural Network.
4.1 Pattern Recognition
In pattern recognition problems, we want a neural network to classify inputs into a set of target categories. The Neural Network Pattern Recognition tool helps you select data, create and train a network, and evaluate its performance using mean square error and confusion matrices.

A two-layer feed-forward network, with sigmoid hidden and output neurons helps to classify vectors randomly by taking enough neurons in its hidden layer\(^{11}\). The network will be trained with scaled conjugate gradient back propagation. The weight adjustment is done in the steepest descent direction in basic back propagation algorithm i.e. negative of gradient and according to this direction the performance function decreased rapidly but does not yield faster convergence.

The derivative functions of weight, net input and transfer functions are responsible for training the network in existing scaled conjugate gradient back propagation\(^{32}\). The training stops when any of these state occurs:

- The maximum number of epochs (repetitions) is reached.
- The maximum amount of time is exceeded.
- Performance is minimized to the goal.
- The performance gradient falls below min_grad.
- Validation performance has increased more than max_fail times since the last time it decreased (when using validation).

4.2 Validation and Testing Data
In preparation of multilayer network, we divide total samples into three subsets. The first is the training sample which is responsible for computing gradient and updating weights of network and bias and also help in discovering potentially predictive relationships. This contains a set of data that has pre-classified target and predictor variables. These are presented to the network during training, and the network is adjusted according to its error. Second is the validation which are used to measure network generalization, and to halt training when generalization stops improving. Third is the testing samples which are responsible for evaluating model with data outside the training set. These have no effect on training and so provide an independent measure of network performance during and after training.

We have divided 15 samples of 227 elements between Training, Validation, Testing as shown in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Training</th>
<th>Validation</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>70%</td>
<td>15%</td>
<td>15%</td>
</tr>
<tr>
<td>Samples</td>
<td>11</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

4.3 Hidden Neurons
To create an overall neural network architecture number of neurons in the hidden layers is important. These layers do not directly interact with the external environment but they have an influence on the final output\(^{33}\). Both the number of hidden layers and the number of neurons in each of these hidden layers is chosen. Choice of few neurons in hidden layer result in underfitting and choice of too many neurons result in overfitting. We have 20 hidden neurons and Neural Network created with 227 inputs is shown in Figure 3.

![Figure 3. Neural Network.](image)

4.4 Training Network
Training is done using scaled conjugate gradient back propagation. In basic back propagation algorithm adjustment of weights is done in the negative of the gradient. According to gradient performance function decreased rapidly but it does not produce the fastest convergence. As compared to the conjugate gradient algorithms a search is performed along conjugate directions, which in turn generally has faster convergence than steepest descent directions\(^{34}\).

In this section, we present four different variations of conjugate gradient algorithms. When generalization stops improving training is stopped automatically which can be observed in increase in the mean square error of the validation samples. Mean Squared Error (MSE) is the average squared difference between target and output. Zero value means no error and Percent Error (%E) shows number of samples that are classified wrong. Zero shows no classification and hundred means maximum classification.
5. Results and Discussions

Grey matter and white matters from seven brain MRI were used; out of which, seven images represents unhealthy and eight images from healthy group. The measure of every picture is 256×256 having dicomm file format. Two dimensional filters are created and defined in matlab, using these filters images are cropped and segmented. One represent grey matter and zero represent white matter. Counting of number of zeros and number of one’s is performed. Aggregate number of one’s subtracted by the quantity of zeroes in row. This step is repeated for fifteen patients and the data retrieved after subtraction is imported in Matlab as shown in Table 2.

**Table 2.** Preprocessing and extraction of grey and white matter in the form of 0 and 1

<table>
<thead>
<tr>
<th>Healthy Subjects</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient 1 Brain_007. dcm210Rx177C</td>
<td><img src="image1.png" alt="Image" /></td>
</tr>
<tr>
<td>Patient 2 Brain_015. dcm212R X 162C</td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
<tr>
<td>Patient 3 Brain_006. dcm 207R X 177C</td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>Patient 4 Brain_19. dcm 200R X 158C</td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
<tr>
<td>Patient 5 Brain_004. dcm 199R X 177C</td>
<td><img src="image5.png" alt="Image" /></td>
</tr>
</tbody>
</table>
Patient 6 Brain_017.
dcm 208R X 169C

Patient 7 Brain_018.
dcm 205R X 164C

Unhealthy Subjects

Patient 8 IM_0035.
dcm 210R X 152C

Patient 9 IM_0035.
dcm 190R X 174C

Patient 10 IM_0032.
dcm 210R X 153C

Patient 11 IM_0078.
dcm 195R X 165C

Patient 12 IM_0047.
dcm 229R X 173C
5.1 Result Analysis

Then 227 elements of fifteen samples are taken as input and 1 element of 15 samples are taken as output. Fifteen samples are divided into Validation, testing and training. The network is trained with 20 neurons in hidden layer as shown in Figure 3. We have generated a confusion matrix shown in Figure 7. It's a table layout in the form of matrix called error matrix. It shows performance of an algorithm and it is called learning in a supervised environment. It shows prediction and information done by classification system about actual and predicted one.

5.1.1 Training Confusion Matrix

We have taken fifteen MRI images of subjects. To perform training on fifteen subjects the samples are randomly divided and eleven subjects are taken for training purpose which in turn showing seven falling under healthy category and 4 falling under unhealthy category as shown in Figure 4.

5.1.2 Validation Confusion Matrix

Then the validation confusion matrix is generated with 2 samples belong to output class 2(unhealthy) as shown in Figure 5.

5.1.3 Testing Confusion Matrix

Figure 4. Training confusion matrix.

Figure 5. Validation confusion matrix.
Testing confusion matrix shows one sample belong to output class1(healthy) and one sample belong to output class2(unhealthy) as shown in Figure 6.

The division of 15 samples into training, validation, testing is done randomly by Neural Network pattern recognition tool. Overall accuracy is shown in Figure 7. showing that seven healthy samples are lying in category 1 (Healthy) and eight are lying in category 2 (Unhealthy). The overall accuracy of classification is shown as 100% in extreme right corner of confusion matrix. Mean square error is also obtained with error percentage and the overall result is shown in Figure 7.

Figure 6. Testing confusion matrix.

Figure 7. Overall accuracy.

5.1.4 Mean Square Error(MSE) and %E(Percentage Error)

Mean square error is also obtained with error percentage as shown in Table 3 and performance plot is shown in Figure 8 plotting MSE (Mean Square Error) on Y axis and epoch (15) on X axis. Best validation performance is 1.7286e-007. Percentage Error (%E) shows number of samples classified wrongly.

<table>
<thead>
<tr>
<th>Results</th>
<th>Samples</th>
<th>MSE (Mean Square Error)</th>
<th>% E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>11</td>
<td>5.66825e-8</td>
<td>0</td>
</tr>
<tr>
<td>Validation</td>
<td>2</td>
<td>1.16528e-3</td>
<td>0</td>
</tr>
<tr>
<td>Testing</td>
<td>2</td>
<td>1.27667e-6</td>
<td>0</td>
</tr>
</tbody>
</table>

5.1.5 Performance Plot

Performance plot is also obtained as shown in Figure 8.

Figure 8. Performance plot.

6. Conclusion

This paper shows a completely programmed and novel way to deal with analysis of brain MRI images. We have used neural network pattern recognition toolbox for classification among healthy and unhealthy subjects. The work is demonstrated 100% precision on 15 subjects.

7. References

6. Cocosco CA, Zijdenbos AP, Alan CE. A fully automatic and
A Proposed Stratification Approach for MRI Images


