International Journal of Engineering & Technology, 7 (2.31) (2018) 199-202

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Predicting malignant and benign brain tumor using image processing

Rishabh Saxena[1*], Aakriti Johri[2], Vikas Deep[3], Purushottam Sharma[4]

[1]*Amity School of Engineering and Technology, Amity University Uttar Pradesh, India.*
[2]*Amity School of Engineering and Technology, Amity University Uttar Pradesh, India.*
[3]*Amity School of Engineering and Technology, Amity University Uttar Pradesh, India.*
[4]*Amity School of Engineering and Technology, Amity University Uttar Pradesh, India.*
**Corresponding author E-mail:rishabhrhea95@gmail.com*

Abstract

Brain is the most important and versatile organ of the human body. One of the most deadly diseases that damage the brain is the accumulation of unwanted and deadly cells near the curvature of brain known as brain tumor. There are two types of brain tumor namely malignant and benign. Malignant is a cancerous tumor and benign is a non cancerous tumor. Primarily brain tumor grows in the brain tissue. The project uses MATLAB to develop a prediction system which uses original hospital brain MRI to predict the brain tumor. Project uses digital image processing to predict the brain tumor. The use of certain image mining algorithms helps in predicting the correct spot and area of brain tumor by image segmentation. The procedure starts with uploading MRI image of human brain, forward by the pre-processing of the image.

Keywords: *Digital image processing, malignant benign, MATLAB, discrete wavelet transformation, K means, support vector machine, principal component analysis.*

1. Introduction

As seen in the past, it has been observed that the rate of brain tumor is increasing in India. Therefore, there was a need of a prediction system that would help detect brain tumor at an early stage in patients and help doctors to cure it as soon as possible. For building this prediction system Digital image processing have been used.Image is taken as an input and the result is the digitally produced image using certain algorithms and methods. Adobe Photoshop can be the best example of DIP. DIP allows the user a better range of algorithms to be used as an input data and can reduce constraints like signal and noise misrepresentation during the process. Since images are usually found over 2D, digital image processing may be used in its best form of multidimensional systems. The process of digitally improving the image is done by following three main levels of processing. They are namely low level, mid-level and high level process.

Low level process takes an image as input and generates image as an output. The main aim of low level process is to remove the noise from the image and sharpening the quality. This is mainly done to enhance the image before the analysis. Mid level process takes an image as input and generates different attributes as its output.

The main aim of mid level process is to recognize the objects present in the image and segmentation of the image. It basically makes clusters of similar data points and calculates its distance from the centroids. High level process takes attributes as the input and generates the final understanding as the final result of analysis. The main aim of this process is to accumulate the final result of the processing. It also provides with autonomous navigation. Today the digital image processing is being used in all kinds of sphere and area. Various applications of Image processing are: Image enhancement/restoration, Artistic effects, Medical visualization, Industrial inspection, Law enforcement and Human computer interfaces.

The fundamental stages in image processing are Image Acquisition, Image Enhancement, Image Restoration, Colour Image Processing, Wavelets and Multi-Resolution Processing, Compression, Morphological Processing, Segmentation, Representation and Description, Object recognition and Knowledge Base.

Brain tumor one of the most dangerous and vulnerable problems related to brain is growing significantly around the nation. There are two types of brain tumor namely malignant and benign. Malignant is a cancerous tumor and benign is a non-canceroustumor. Primarily brain tumor grows in the brain tissue. Doctors have grouped brain tumor into groups grade I, II, III and IV which is the most dangerous one. The grade is determined by the size of brain cell under the microscope. From grade I to II the brain tumor is non-cancerous and from grade III to IV the brain tumor is said to be cancerous.

Higher is the grade of tumor larger would be the size of the brain cell and higher would be the aggressiveness of the tumor. Diagnosis of a brain tumor is executed using a neurologic exam, CT (computer tomography scan) and a magnetic resonance imaging (MRI), and other tests like an angiogram, spinal tap and biopsy. Your diagnosis helps predict the treatment. Tocure brain tumor various techniques used are surgery, radiation or chemotherapy.

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

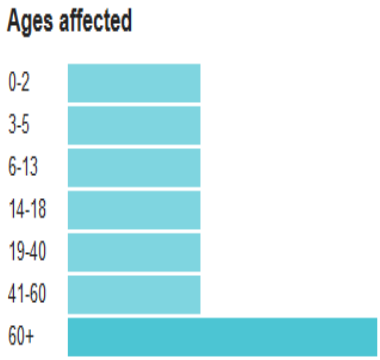

Fig. 1: Affect of brain tumor at different age

This prediction system have been made using MATLAB. Matrix Laboratory was developed by Math Works. It is high-level and an interactive programming language which is being used by the engineers for data analysis and data visualization. The various environments on which MATLAB can be implemented are UNIX, Macintosh and Windows. The various operations which can be done are: matrix manipulations, function and data plotting, algorithm development and implementation, creating user interface, data analysis, creating applications and models and also interfacing with programs that are written in C, C++, Java and FORTRAN languages.

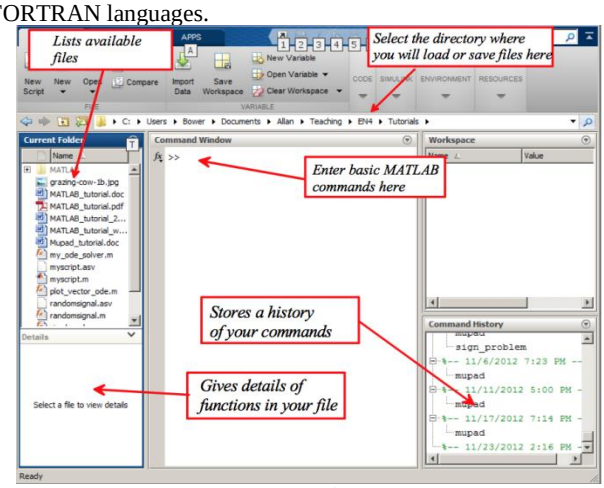

Fig. 2: Layout of MATLAB

2. Materials and methods

Image Processing allows the user a better range of algorithms to be used as an input data can reduce constraints like signal and noise misrepresentation during the process. Since images are usually found over 2D, digital image processing may be used in its best form of multidimensional systems. Image processing mainly focuses on producing a system that will be performing on an image. Image is taken as an input and the result is the digitally produced image using certain algorithms and methods. The steps involved during image processing are:

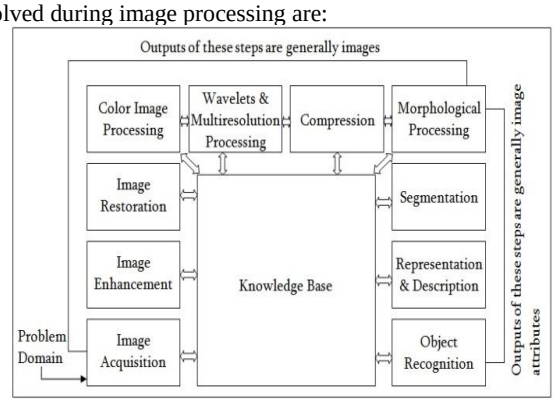

Fig. 3: Stages of image processing

Discrete Wavelet Transformation (DWT)

Image compression is considered to be the most important steps of image processing. As the amount size and quality of images are increasing, the demand for a method to compress these images plays a very important and significant role. Discrete Wavelet Transformation (DWT) is efficiently used for signal and image processing. This transformation is done in the second step of image processing, that is in pre-processing step. DWT is basically used for reducing the redundancy of the images so that it is easier to store and transmit image data in an efficient way. This transformation helps the user to analyze the finer bands of signals as well the images. The analysis here is called multi-resolution analysis, which helps the user to detect various patterns which cannot be seen or visible in the raw data.

K-Means clustering

K means clustering algorithm helps to partition n objects into k clusters in which each and every object has a corresponding cluster having nearest mean value. The main aim of the algorithm is to reduce the total intra-cluster variance.

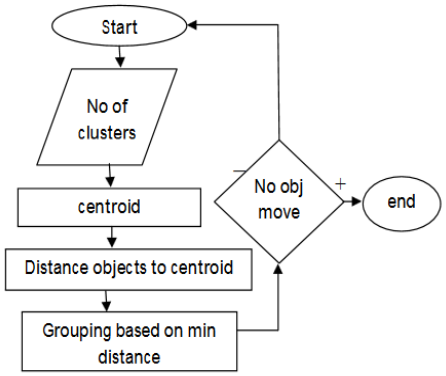

Fig. 4: Algorithm of K-means

Algorithm of K-means Clustering is as follows:

1. Cluster the data into k groups.
2. Select *k* point as random to show the center of cluster.
3. Allocate the objects to their neighboring cluster by using the Euclidean distance formula.
4. Find the mean of each of the cluster separately.
5. Repeat steps from 2 to 4 so as to get the value of each point separately.

Principal Component Analysis (PCA)

Principal component analysis (PCA) is a statistical technique which helps us to perform linear transformation. PCA was invented by Karl Pearsonin 1901. Principal component analysis works as a commanding means for both data analysis as well as pattern recognition which is further used for signal and image processing as a method to compress, reduce dimension and de-correlate the given image.

PCA is basically used for feature extraction in image processing. The main aim of Principal component analysis is to find various similar patterns in the data image. It helps us to find correlation between the image variables. It attempts to reduce the dimensionality of the image data whenever there is strong correlation between the variables. No information is lost darning this whole process of data extraction

Support vector machine

Support Vector Machine (SVM) or support vector network is a supervised machine learning algorithm which is effectively used for both classification and regression. In this algorithm, a graph is plotted for each data point in n-dimensional graph (where n represents the number of features). In this graph, the value of each feature is considered as the value of a particular coordinate. After

plotting the graph, classification is done where we find Hyper-plane which helps to differentiate the two classes.

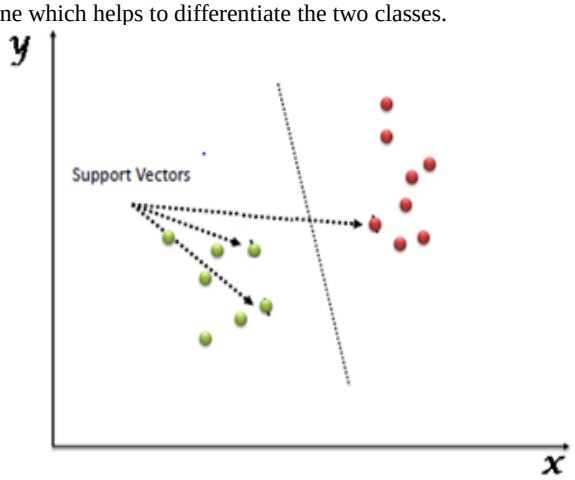

Fig. 5: Concept of vector in SVM

A hyperplane is an imaginary line that linearly classifies and distinguishes a set of data. More the distance from the hyperplane the data points lies; more confidence is there that the data is correctly and accurately classified. Data points are to be as far away from the hyperplane as possible so as to correctly classify it in all regards.

3. Result and discussion

The MRI image containing brain tumor is first loaded on to MATLAB graphical user interface. The image is loaded on to the axis. The axis is responsible for aligning the image properly. It helps in fitting the image accurately.

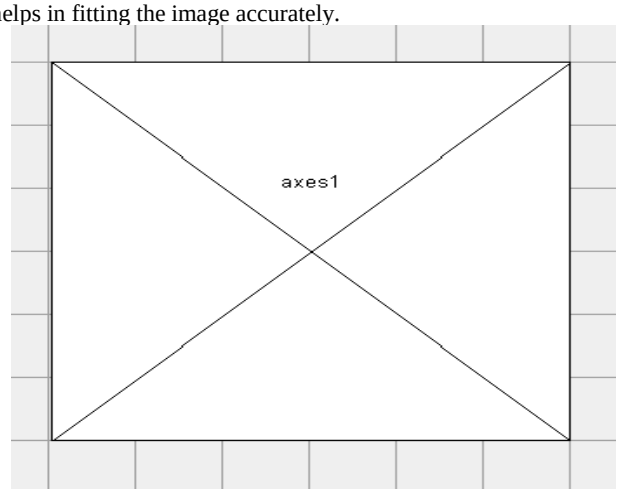

Fig. 6: Axes to load MRI image

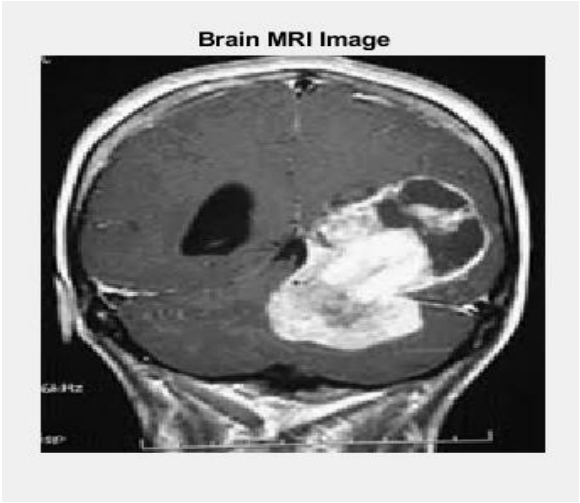

Fig. 7: MRI image loaded in MATLAB

After the image is uploaded on the axis, the next step is to perform the image processing. The first stage of digital image processing is to pre-process the image, so as to remove noise and reduce the space of the image. One such method used in this project is DWT (Discrete Wavelet Transformation. DWT helps in removing the unwanted white noise from the image. This transformation is done in the second step of image processing, that is in pre-processing step. DWT is basically used for reducing the redundancy of the images so that it is easier to store and transmit image data in an efficient way. This transformation helps the user to analyze the finer bands of signals as well the images. The analysis here is called multi-resolution analysis, which helps the user to detect various patterns which cannot be seen or visible in the raw data.

In this prediction system, Single level 2-Dimensional DWT has been used. Here, the image can be viewed in three orientations and they are vertical, horizontal and diagonal as compared to Single level 1-Dimensional where only two orientations are there and they are vertical and horizontal.

The next step is to do segmentation of the image. For segmentation of the image K means algorithm has mean used. As K means algorithm is mostly used for clustering. It is the best algorithm to segment the images. K means clustering algorithm helps to partition n objects into k clusters in which each and every object has a corresponding cluster having nearest mean value. The main aim of the algorithm is to reduce the total intra-cluster variance.

Now the next step is select the required feature that would be give best accuracy. Feature selection is done by PCA (Principle Component Analysis). The main aim of Principal component analysis is to find various similar patterns in the data image. It helps us to find correlation between the image variables. It attempts to reduce the dimensionality of the image data whenever there is strong correlation between the variables. No information is lost darning this whole process of data extraction. PCA ignores the pre-defined class labels and consider whole datasets for feature extraction. It helps us to represent multi-dimensional data having large number of variable with less number of variables without losing the main features in the data. The variables that are lost are considered as not so important feature for the analysis. Cross validation is performed in the project by creating 10 folds. 10 folds are created for training

Next is to calculate the accuracy of linear, RBF, polygonal and quadratic kernel. Every accuracy has their own uses. These accuracies are all calculated by the SVM algorithm. The images are classified using SVM algorithm. Support Vector Machine (SVM) or support vector network is a supervised machine learning algorithm which is effectively used for both classification and regression. In this algorithm, a graph is plotted for each data point in n-dimensional graph (where n represents the number of features). In this graph, the value of each feature is considered as the value of a particular coordinate. After plotting the graph, classification is done where we find Hyper-plane which helps to differentiate the two classes.

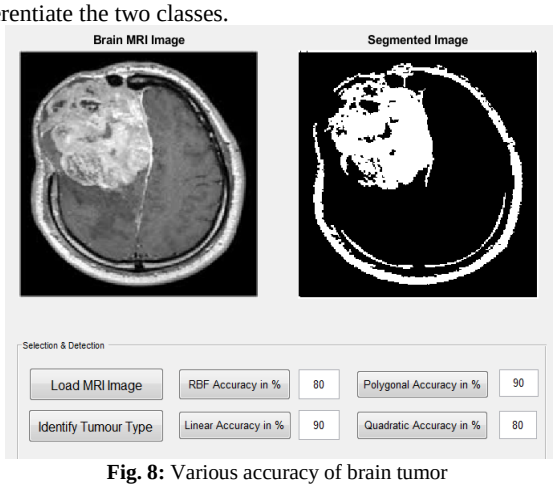

Fig. 8: Various accuracy of brain tumor

We were able to observe various features (values) of the MRI image like mean, standard deviation, variance, correlation, entropy, smoothness, energy, homogeneity, etc.

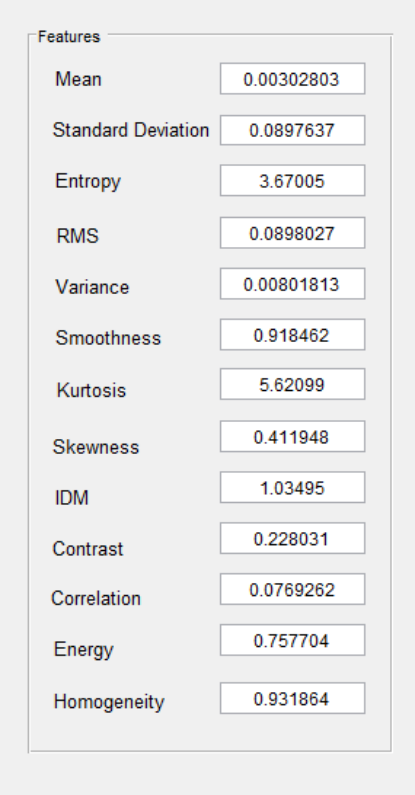

Fig. 9: Various features of the brain tumor selected

4. Conclusion

The project is used to automatically and efficiently predict the type of brain tumor in humans. In the project two types of brain tumor have been considered: malignant and benign. Malignant is a cancerous tumor and benign is a non-canceroustumor. The malignant brain tumor is more dangerous than the benign as it grows faster than it. The cell of malignant brain tumor grows in a much faster pace. They are a life threatening brain tumor and can mostly leads to death if treatment is not started at an early stage. Malignant brain tumor is sometimes also known as brain cancer, but it is not actually a brain cancer. Cancer has a tendency to spread from one body organ to another which does not take place in malignant brain tumor. On the other hand benign brain tumor grows at a slower pace. They mostly have a well defined border and can be easily removed by a surgery. Benign brain tumor can also be considered as life threatening if the tumor is located at the critical locations of the brain like the cells which control the breathing operation of the human body. This prediction system has been developed using MATLAB and using the concept of digital image processing. Digital image processing mainly focuses on producing a system that will be performing on an image. Image is taken as an input and the result is the digitally produced image using certain algorithms and methods. Today digital image processing is being used in all kinds of field and area. Various applications of Image processing are: Image enhancement/restoration, Artistic effects, Medical visualization, Industrial inspection, Law enforcement and Human computer interfaces. Braintumor is one of the most dangerous and difficult to cure disease. This system can be used in medical industry to predict the occurrence and locate the tumor in brain using image segmentation. Image segmentation in the process is done using k means algorithm. The main aim of developing such a system is to test the MRI images with test data sets and get accurate results. Here, Discrete Wavelets Transformation (DWT) has been used for reducing the redundancy of the images so that it is easier to store and transmit image data in an efficient way. This transformation helps the user to analyze the finer bands of signals as well the images. Principal Component Analysis (PCA) has been used for feature extraction in image processing. The main aim of Principal component analysis is to find various similar patterns in the data image. It helps us to represent multi-dimensional data having large number of variable with less number of variables without losing the main features in the data. K-means clustering has been used for image segmentation. This clustering algorithm helps to partition n objects into k clusters in which each and every object has a corresponding cluster having nearest mean value. The main aim of the algorithm is to reduce the total intra-cluster variance. For classification, Support Vector Machine (SVM) or support vector network is a supervised machine learning algorithm which is effectively used for both classification and regression. In this algorithm, a graph is plotted for each data point in n-dimensional graph (where n represents the number of features). In this graph, the value of each feature is considered as the value of a particular coordinate. After plotting the graph, classification is done where we find Hyper-plane which helps to differentiate the two classes. After using the above four concepts we were able to predict type of brain tumor in the patients. We were able to observe various features (values) of the MRI image like mean, standard deviation, variance, correlation, entropy, smoothness, energy, homogeneity, etc.

References

[1] Zang Q, "Wavelet Network in Nonparametric Estimation", *IEEE Trans. Neural Networks*, Vol.8, No.2, (1997), pp.227-236.
[2] Tajunisha S & Saravanan V, "Performance analysis of k-means with different initialization methods for high dimensional data", *International Journal of Artificial Intelligence & Applications (IJAIA)*, Vol.1, No.4, (2010), pp.44-52.
[3] Ashraf R & Akbar M, "Absolutely lossless compression of medical images", *IEEE 27th Annual Conference Proceedings of the Engineering in Medicine and Biology*, (2005), pp.4006-4009.
[4] Murali Mohan S & Sathyanarayana P, "FPGA Implementation of Hybrid Architecture for Image Compression Optimized for Low Power and High Speed applications", *International Journal of Scientific & Engineering Research*, Vol.4, No.5, (2013).
[5] Ye Z, Mohamadian H & Ye Y, "Information Measures for Biometric Identification via 2D Discrete Wavelet Transform", *3rd Annual IEEE Conference on Automation Science and Engineering*, (2007), pp.835-840.
[6] Chris D & Xiaofeng H, "k-means Clustering via Principal component Analysis", *21st international conference on Machine Learning*, (2004).
[7] Jiawei Han MK, *Data mining Concepts and Techniques, morgan Kaufmann publishers*, An imprint of Elsevier, (2006).
[8] Jolliffe IT, *Principal Component Analysis*, Springer, Second edition, (2002).
[9] Papageorgiou EI & Spyridonos P, "Brain Tumor characterization using the soft computing technique of fuzzy cognitive maps", *Applied soft computing*, Vol.8, (2008), pp.820-828.
[10] Gladis Pushparathi VP & Palani, S, "Linear Discriminant analysis for brain tumor classification using Feature Selection", *Int. J. Communication and Engineering*, Vol.5, No.4, (2013), pp.1179-1185.
[11] Matthew CC & Lawrence OH, "Automatic tumor segmentation using Knowledge based techniques", *IEEE transactions on medical imaging*, Vol.17, No.2, (1998).
[12] Shafaf I, Noor EAK, "Image Mosaicing for evaluation of MRI Brain Tissue abnormalities segmentation study", *Int.J.Biology and Biomedical Engineering*, Vol.5, No.4, (2011), pp.181-189.
[13] Iftekharuddin KM, "On techniques in fractal analysis and their applications in brian MRI", *Medical imaging systems: technology and applications, Analysis and Computational Methods*, Vol.1, (2005).
[14] Bezdek JC, Hall LO, Clarke LP, "Review of MR image segmentation techniques using pattern recognition", *Med. Phys.*, Vol.20, No.4, (1993), pp.1033-1048.