# SUPERVISED LEARNING (CLASSIFICATION) ALGORITHMS FOR BRAIN TUMOUR DETECTION

**Sarvesh Mahajan1, Ruchi Rautela2**

1PG Student, Vivekanand Education Society's Institute of Technology, Dept. of MCA, Mumbai, India
2Assistant Professor, Vivekanand Education Society's Institute of Technology, Dept. of MCA, Mumbai, India

---------------------------------------------***-------------------------------------------------
-

## 1 ] ABSTRACT :

Several medical diagnostic applications use Automated defect detection. Automated tumour diagnosis in magnetic resonance imaging (MRI) is extremely important since it provides details about aberrant tissues that are required for treatment planning. Human inspection is the standard procedure for flaw detection in magnetic resonance brain pictures. For vast amounts of data, this strategy is impractical. Therefore, automated tumour identification techniques are created in order to free up radiologist time. Due to the intricacy and variety of malignancies, MRI brain tumour detection is a challenging task . In this study, machine learning techniques are used to find tumours in brain MRIs. The proposed study is broken down into three sections: applying preprocessing techniques to brain MRI images, extracting texture features using a **Gray Level Co-occurrence Matrix (GLCM)**, and classifying the results using a machine learning method.

Key Words : Machine Learning , MRI , Feature Extraction , Grey Level Co occurrence Matrix

## 2 ] INTRODUCTION :

Numerous different cell types make up the human body.
Every cell serves a distinct purpose. The body's cells develop, divide, and create new cells in an organised fashion. These new cells contribute to the maintenance of the human body's functionality and health. Some cells develop chaotically when they lose the ability to regulate their own growth. Tumors are tissue masses made of excess cells and are the result of this process. Both benign and malignant tumours are possible. **While**

**benign tumours do not cause cancer, malignant tumours do.** 40,000 patients were diagnosed with benign and malignant brain tumours in 2002, according to a report released by the central brain tumour registry of the United States (CBTRUS). It shows that there are **15 primary brain tumours per 100,000 people,** whether they are malignant or benign. Medical image data from various biomedical equipment that use imaging techniques like X-ray, CT scan, and MRI are a crucial component in making a diagnosis. A patient's body's water molecules include hydrogen atoms, and magnetic resonance imaging (MRI) relies on the detection of magnetic field vectors produced following the right combination of strong magnetic fields and radio frequency pulses.
Since the MRI scan uses no radiation, it is far more effective for diagnosis than the CT scan. Using MRI, radiologists may assess the brain.
The MRI method can identify whether a brain tumour is present. Currently tumour detection in MR Images is done through human inspection . This processis very time consuming . **For vast amounts of data, it is inappropriate. Additionally**, **noise introduced by operator interaction in the MRI can result in incorrect classification.** Automated systems are required since they are more cost-effective because there is a large volume of MRI data to evaluate. As great precision is required when dealing with human life, automated tumour detection in MRI images is required.

Artificial neural networks, support vector machines, and unsupervised methods like self-organization maps (SOM) and fuzzy c-means when coupled with feature extraction approaches are used to classify MR images of human brains. Other supervised classification methods, including k-nearest neighbours (kNN), similarly classify pixels according to how similar each feature is to each other.

Both supervised and unsupervised algorithms can be used to categorise MR pictures as either normal or abnormal.
**In this paper, machine learning techniques are used to suggest an effective automated categorization method for brain MRI. The classification of brain MR images using the supervised machine learning technique.**

Consequently, we discovered via a thorough literature review that the majority of the existing brain tumour detection method uses **texture, symmetry, and intensity as features**. As texture perception plays a vital role in the human visual system of recognition and interpretation, texture features are an important characteristic of the brain. Here, we suggest extracting texture features like homogeneity, energy, contrast, and correlation. Co-occurrence of Gray Level The extraction of texture features uses a matrix.

Additionally, we suggest using ML models to overcome the drawbacks of traditional classification methods. In this paper, we analyze and compare the performance of two machine learning algorithms, MLP and Naive Bayes. As these ML algorithms are found to perform accurately in most of the pattern recognization. Given its capacity to learn intricate input-to-output mappings, neural networks are useful. Categorization exercises.Neural networks are precisel as well as helpful as they can learn complex mappings between input and output and solve challenging classification problems .They are capable of solving much more complicated classification problems. However, Naïve Bayes performs well in cases of categorical input data compared to numerical data .

## 3 ]RELATED WORK :

1] Based on feed-forward neural network classifiers and rough set theory, Rajesh and Malar proposed classifying brain MR images. Using rough set theory, the characteristics are retrieved from MRI scans. With an accuracy of roughly 90%, the chosen features are sent into a feed-forward neural network classifier, which distinguishes between normal and disordered brain.

2] Based on Support Vector Machines, Jafari and Shafaghi suggested a hybrid method for brain tumour detection in MR images (SVM).

## 4 ]PROPOSED SYSTEM :

According to various research sources, automated system for brain tumour identification is required . As human lifes are involved in this process, **accuracy** is very important factor . Using machine learning algorithms for feature extraction and classification, automated tumour detection in MR images is possible. A approach to automatically identify tumours in MR images is proposed in this paper, as seen in figure

1.

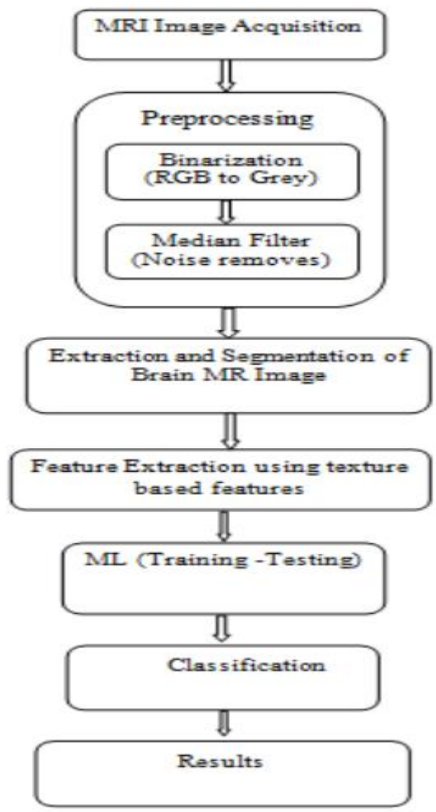


**Image Acquisition**: MRI images of the brain are obtained and are given as input to the pre-processing stage, sample brain MRI images are shown in Figure 2.A

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 07 | July - 2022 Impact Factor: 7.185 ISSN: 2582-3930

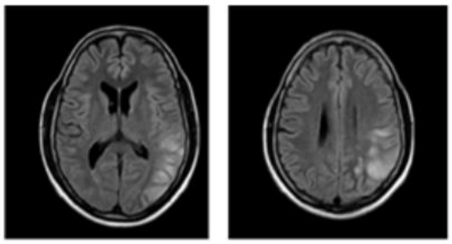

Fig 2 .A

**Preprocessing**: Preprocessing is essential because it enhances some of the image features that are crucial for further processing by enhancing the image data.The pre-processing steps that are applied to MR image are as follows :

I. The RGB MR image is converted to gray scale image and then median filter is applied for noise removal from brain MR images as shown in figure 3(b). Noise elimination is necessary , Since higher accuracy is required in further process .

II. Then edges are detected from filtered image using canny edge detection as shown in figure 3(c). The edge detected image is required for image segmentation technique .

III. Then watershed segmentation is carried out to identify the tumor's position in the brain image as shown in
figure 3(d). Image segmentation is a method in which a digital image is broken down into various subgroups called Image segments which helps in reducing the complexity of the image to make further processing or analysis of the image simpler. The purpose of the Image segmentation is to transform the image representation into something that is easier to analyze.

Label image is the result of the watershed segmentation.All the different objects that are recognized will have different pixel values.Each pixel of the first object in the label image will have the value 1, the pixels of the second object will have the value 2, and so on.

Figure 3 displays the various preprocessing techniques used on a brain MR picture.

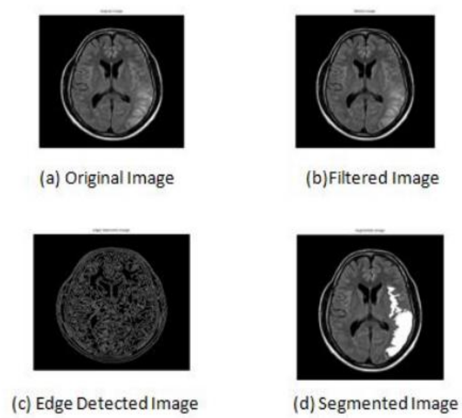


Fig 3.

**Feature Extraction :**

In machine learning and pattern recognition, a feature vector is an n-dimensional vector of numerical features that represent some object . When an algorithm receives a very big and redundant input, it is transformed into reduced representative set of features called **feature vector**.Feature extraction is the process of transforming a set of input data to features .In this step, the important features needed for image classification are extracted.The segmented brain MRI images are used and texture features are extracted from this segmented images which shows the texture property of the images. The Gray Level Co-occurrence Matrix (GLCM), a robust and effective approach, is used to extract these properties. The GLCM texture feature extraction approach is extremely competitive since it minimises the size of the GLCM, which lowers the computational cost of the algorithm while maintaining good classification rates. To identify between a normal and abnormal brain texture, the GLCM features are used. Important details regarding the surface structural arrangement are present in texture. In image classification, the textural features based on gray-tone spatial relationships are generally applicable.

The **Gray Level Co-occurrence Matrix** texture features that are extracted for developing model are as follows**:**

**(1) Energy :** Energy gives a measure of textural homogeneity, that is, measure of pixel pair repetitions.

**Energy feature**

$$Energy = \sum_{i,j=0}^{N-1} \left(P_{ij}\right)^2$$

$$E = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p(i,j)^2 \quad \text{.........(1)}$$

**Range=[0,1]**

(2) **Contrast**: Contrast gives a measure of intensity contrast between two pixel in whole image.

$$Con = \sum_{n=0}^{N_g-1} n^2 \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p(i,j)^2 \quad \text{.....}$$

Range=[0,1]

$$Contrast = \sum_{i,j=0}^{N-1} P_{ij} (i-j)^2$$

(3) **Correlation**: Correlation gives a measure of how correlated two pixels are in whole image.

$$Correlation = \sum_{i,j=0}^{N-1} P_{ij} \frac{(i-\mu)(j-\mu)}{\sigma^2}$$

$$C = \frac{1}{\sigma_x \sigma_y} \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i,j) p(i,j)^2 - \mu_x \mu_y \quad \text{..(3}$$

Range=[-1,1]

(4) **Homogeneity**: Homogeneity indicates how closely an element's distribution to the GLCM diagonal.

$$Homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{ij}}{1+(i-j)^2}$$

$$H = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{p(i,j)}{(1+mod(i,j))} \quad .....($$

Range=[0,1]

Thus to categorise brain as normal or abnormal , the MRI brain images are gathered, preprocessing steps and segmentation techniques are then applied. The segmented image is used for feature extraction, and GLCM is used to extract texture features. Matlab 2012a is used for the preprocessing and feature extraction processes.

**Classification:**

The Machine learning algorithms are used for the classification of MRI brain images either as normal or abnormal. The major aim of ML algorithms is to automatically learn and make smart choices. The feature set created by above method was applied to **Multi-Layer Perceptron (MLP) and Naive Bayes** for classification of MR images. MLP is a feed forward artificial neural network model that maps sets of input data into a set of appropriate output. It is known as feed forward neural network, because it does not contain any cycles and network output depends only on the current input instance. In Multilayer perceptron, each node is a neuron with a nonlinear activation function . It is based on supervised learning method. Learning takes place by changing connection weights after each piece of data is processed, depending on the amount of error in the target output as compared to the expected result. The goal of the learning procedure is to reduce error by enhancing the current values of the weight associated with each edge. The weights are changed backwards during this process, which is why the model is known as back-propagation.

Naive bayes is a supervised learning method for classification. It is simple Bayes theorem-based probabilistic classifier. It assumes that the value of a particular feature is independent of the presence or absence of any other feature. The posterior probability is calculated by calculating the prior probability and likelihood .For estimating the parameters , the method of maximum posterior probability is used. With small amount of training data parameters required for classification can be estimated .Hence time taken for classification as well as training reduces .

## 5] EXPERIMENTAL RESULTS :

The experiment was carried out on 460 brain MR images.Texture-based features are extracted using the weka tool. These features are used for classification . The texture based features such as contrast, correlation, energy and homogeniety are extracted using GLCM. The Multi-Layer Perceptron (MLP) and Naïve bayes with 66% percentage split is used for classification. In 66% percentage split, 66% of the instances are used for training and remaining instances are used for testing.

**Table 1. Experiment result analysis**

| ML Algorithm | Total samples | Model Build Time | Classification Rate (%) |
|---|---|---|---|
| MultiLayerPerceptron | 460 | 68.6 | 98.6 |
| Naive bayes | 460 | 0.08 | 97.6 |

From the Table 1(Above Table) , we can find the classification rate of brain MR images using MLP and Naive bayes. The accuracy of about 98.6% and 91.6% is obtained respectively.

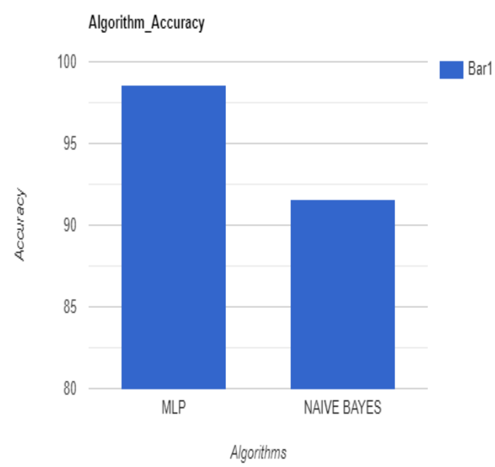


Fig 4. Graphical representation of accuracy rate .

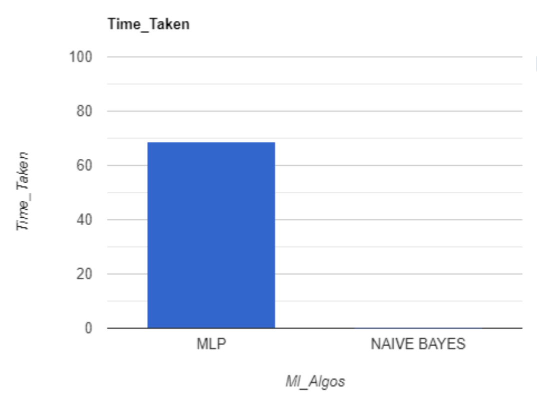


Fig 5. Graphical representation of time taken to build algorithm.

The MLP provides more accuracy and takes more time to build the model , while the time taken by Naïve Bayes is less and is accuracy is also less .Due to complex appearances and complexity of tumors, the proposed method gives the satisfied accuracy. Great accuracy is preferred since human life is involved .

## 6 ] CONCLUSION

This paper proposes a work on brain tumor detection system based with classification algorithms. The texture based features are extracted using Gray Level Co occurrence Matrix (GLCM). The texture features of the image considered in this proposed work include correlation, homogeneity, energy ,contrast. For the classification purpose, Multi-Layer Perceptron and Naïve bayes machine learning algorithm is used and the maximum accuracy 98.6% and 91.6% is obtained by considering 460 samples . This accuracy can be increased by considering a large data set and extracting intensity based features in addition to the texture based features.

## 7] REFERENCES :

[1] T. Rajesh, Rough “Set Theory and Feed Forward Neural Network Based Brain Tumor Detection in Magnetic Resonance Images" ,IEEE International on Advanced Nanomaterials & Emerging Engineering Technologies, 20 13.

[2] J. Selvakumar, " Brain Tumor Segmentation and Its Area Calculation in Brain MR Images using K-Mean Clustering and Fuzzy C-Mean Algorithm" , IEEEInternational Conference On Advances In Engineering, Science And Management, ICAESM, 2012.

[3] Amir Shahzad, " Enhanced Watershed Image Processing Segmentation " , Journal of Information & Communication Technology, Vol. 2, No. 1, 2009.

[4] Neelam Marshkole, "Texture and Shape based Classification of Brain Tumors using Linear Vector Quantization" , International Journal of Computer Applications (0975 – 8887) Volume 30– No.11, September 2011.

[5]

https://www.researchgate.net/publication/351321003_Brain_Tumor_Classification_Using_Texture_Feature_Extraction .

[6]
https://www.sciencedirect.com/science/article/abs/pii/S0045790622002774

[7] Mehdi Jafari, "A Hybrid Approach for Automatic Tumor Detection of Brain MRI using Support Vector Machine and Genetic Algorithm", Global Journal of Science Engineering and Technology, Issue-3, 2012