Image Analyzing for MRI Based Brain Tumor Detection System

Lalit Roshankhede[1], Ujwal Kurhade[2], Priyanka Waghmare[3], Nilkanth Bhagat[4]

[1] *Lalit Roshankhede Student, Computer Science and Engineering, Priyadarshini J.L. College of Engineering, Maharashtra, India*
[2] *Ujwal Kurhade Student, Computer Science and Engineering, Priyadarshini J.L. College of Engineering, Maharashtra, India*

[3] *Priyanka Waghmare Student, Computer Science and Engineering, Priyadarshini J.L. College of Engineering, Maharashtra, India*

[4] *Nilkanth Bhagat Student, Computer Science and Engineering, Priyadarshini J.L. College of Engineering, Maharashtra, India*

ABSTRACT

Brain tumor at early stage is very difficult task for doctors to identify. MRI images are more prone to noise and other environmental interference. So it becomes difficult for doctors to identify tumor and their causes. So here we come up with the system, where system will detect brain tumor from images. Here we convert image into grayscale image. We apply filter to image to remove noise and other environmental interference from image. User has to select the image. System will process the image by applying image processing steps. We applied a unique algorithm to detect tumor from brain image. But edges of the image are not sharp in early stage of brain tumor. So we apply image segmentation on image to detect edges of the images. In this method we applied image segmentation to detect tumor. Here we proposed image segmentation process and many image filtering techniques for accuracy.

Keyword : - *Segmentation, Pre-Processing , RGB Conversation ,SVM , Linear Regression .*

1. INTRODUCTION

Brain tumor is that the most ordinarily occurring disease among citizenry, so study of brain tumour is vital. it's the main cause for the rise in mortality among children and adults. Tumor is defined because the abnormal growth of the tissues. brain tumour are generally classified into two type Benign and Malignant tumors.
MRI (Magnetic Resonance Imaging) Images plays a crucial role in brain tumour for analysis, diagnosis and treatment planning. The detection of brain tumour using MRI images is challenging tasks due to the complex structure of the brain.

MRI Images provide better result than CT scan, Ultrasound and X-ray because it's a complicated medical imaging technique and it uses powerful magnet to supply high resonance images of all a part of body. This MRI Image can processed and one can detect the brain tumour using image processing techniques by forming an automatic detection process using various algorithms because the manual detection of tumor from MRI Image

may give human error.

In this project our focus are going to be mainly to enhance the prevailing approaches of image processing or to style a far better approach for the detection of the tumor

2. LITRATURE REVIEW

In recent years, interest in designing tools for diagnosing brain tumors has been increasing. The work of Gopal and Karnan [1] uses image processing clustering algorithms toclassify images into a group that has a brain tumor and another group which does not. The dataset used in this work is composed of 42 MRI images obtained from the KG hospital database.

In the preprocessing phase, the authors remove the film artifacts (labels and X-ray marks). They also use the filter Median to remove high frequency components in the MRI image. The authors then use an algorithm called Fuzzy C Means (FCM) as an image clustering algorithm, in addition to using a Genetic Algorithm (GA) as an intelligent optimization tool.

The results of the experiments showed that, the classification algorithm FCM achieved a classification accuracy of 74.6% with less than 0.4% error rate. To enhance the accuracy, the authors used an optimization technique called Particle Swarm Optimization (PSO). They managed to reach an accuracy level of 92%.
Four different types of classification algorithms wereused as follows.

1. SVM (Support Vector Machine)
2. Random Forest
3. Linear Regression

Natarajan et al. proposed method using Median filter for the preprocessing of the Brian MRI image, for segmentation applied threshold segmentation and morphological operations and then the image subtraction technique is used to get the region of interest in the brain MRI image.

Uchita and Lalit proposed unsupervised MRI brainclassification neural network learning technique. The MRI brain images are first pre-processed including theremoval of the noise edge detection and segmentation of the tumor. Self Organizing Maps (SOM) are used to classify the brain as normal or abnormal, i.e. whether or not it contains tumor.

Muhammad Sajjad, et.al proposed a new CNN-based multi-grade brain tumor classification mechanism in this paper. A deep learning technique was utilized initially forsegmenting the tumor regions from MR image. Further, the proposed system was trained efficiently using theextensive data augmentation in the second step. However, when multi-grade brain tumor classification is to be done using MRI, the issue of lack of data was avoided. Towards the end, the brain tumor gradeclassification was performed by using augmented data and applying pre-trained VGG-19 CNN model on it. Evaluations were performed on original as well as augmented data and as a result performance metrics i.e. sensitivity, specificity and accuracy got improved.

Komal Sharma Automated defect detection in medical imaging has become the emergent field in several medical diagnostic applications. Automated detection of tumor in Magnetic Resonance Imaging (MRI) is very crucial as it provides information about abnormal tissues which is necessary for planning treatment. The conventional method for defect detection in magnetic resonance brain images is human inspection. This method is impractical for large amount of data. So, automated tumor detection methods are developed as it would save radiologist time. The MRI brain tumor detection is complicated task due to complexity and variance of tumors. In this paper, tumor is detected in brain MRI using machine learning algorithms. The proposed work is divided into three parts: preprocessing steps are applied on brain MRI images, texture features are extracted using Gray Level Co- occurrence Matrix (GLCM) and thenclassification is done using machine learning algorithm.

3. METHODOLOGY

1. **Dataset Input (MRI Images) :-** The primary stage i.e. the image acquisition stage which starts with taking a set of images from the database. Images are store in Pycharm are going to be displayed as gray scale image. As our MRI images dataset was in (.jpg and .png) file format, we've extracted the pictures using Pycharm code.

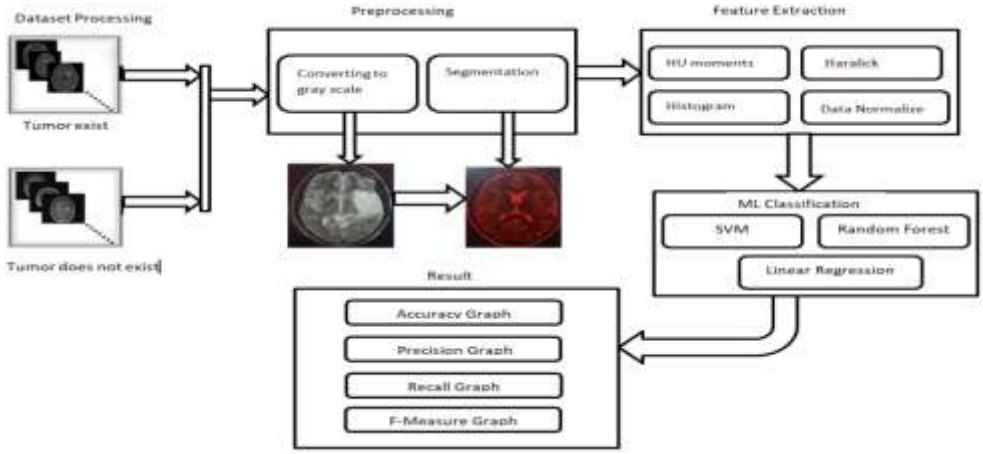

Fig 1. System Architecture

2. **Image Pre-Processing :-** This is often a second stage of image processing where the pictures are wont to enhance the probabilities of detecting suspicious reason. Further images to enhance the enhancement accuracy of the pictures. Its aim to enhance the image data by suppressing the undesired distortions and enhanced a number of image feature which will be helpful in further processing the goal of pre-processing is provide Contrast enhancement to enhance the image quality. The features perform by pre-processing is
   - **Gray Scale Conversion**
   - **Contrast Enhancement**

3. **Image Segmentation :-** The third stage is image segmentation. Image segmentation may be a process of partitioning digital image into multiple segments. Image segmentation while be typically wont to locate object and limits in images and its help to form the analysis easier. it's the method of assigning a label to each pixel in a picture. Such pixels with an equivalent label share the certain visual characteristics. There are several algorithm for segmentation. How ever the choice of segmentation method depends on the kinds of the features to be preserved and extracted.

   - **Thresholding**
   - **Enhancement**
   - **Edge Detection**

4. **Feature Extraction**

   I. HU Moments :- - Hu Moments are wont to characterize the outline or "silhouette" of an object in a picture. Normally, we obtain this shape after applying some kind of segmentation (i.e. setting

the background pixels to black and therefore the foreground pixels to white). Thresholding is that the commonest approach to get our segmentation. After we've performed thresholding we've the silhouette of the thing within the image.

II. **Haralick:-** Haralick texture features are calculated from a Gray Level Co-occurrence Matrix, (GLCM), a matrix that counts the co-occurrence of neighboring gray levels within the image. The GLCM may be a matrix that has the dimension of the amount of gray levels N within the region of interest (ROI). These are texture features, supported the adjacency matrix (the adjacency matrix stores in position (i,j) the amount of times that a pixel takes the worth i next to a pixel with the worth.

III. **Histogram :-** - A histogram is essentially wont to represent data provided during a sort of some groups. it's accurate method for the graphical representation of numerical data distribution. it's a kind of bar plot where Xaxis represents the bin ranges while Y-axis gives information about frequency. to make a histogram the primary step is to make bin of the ranges, then distribute the entire range of the values into a series of intervals, and therefore the count the values which fall under each of the intervals. Bins are clearly identified as consecutive, non-overlapping intervals of variables.

IV. **Data Normalization :-** Normalization refers to rescaling real-valued numeric attributes into a 00 to 11 range. Data normalization is employed in machine learning to form model training less sensitive to the size of features. this enables our model to converge to raised weights and, in turn, results in a more accurate mode.

5. ML Classification

I. SVM (Support Vector Machine) :-

SVM is one among the classification technique applied on different fields like face recognition, text categorization, cancer diagnosis, glaucoma diagnosis, microarray organic phenomenon data analysis [15]. SVM utilizes binary classification of brain MR image as normal or tumor affected. SVM divides the given data into decision surface, (i.e. a hyperplane) which divides the info into two classes. The prime objective of SVM is to maximise the margins between two classes of the hyper-plane [16]. Dimensionality reduction and precise feature set given as input to the SVM on the duration of coaching part also as during the testing part. SVM is predicated on binary classifier which employs supervised learning to supply better results

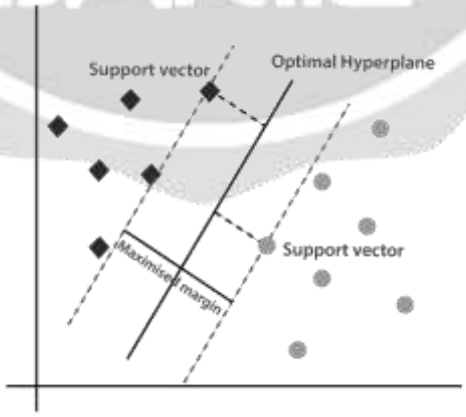

Fig 2. Best line or Decision boundary

II. Random Forest

Random Forest may be a popular machine learning algorithm that belongs to the supervised learning technique.

It are often used for both Classification and Regression problems in ML. Random Forest works in two-phase first is to make the random forest by combining N decision tree, and second is to form predictions for every tree created within the first phase. The Working process are often explained within the below steps and diagram: Step-1: Select random K data points from the training set. Step-2: Build the choice trees related to the chosen data points (Subsets). Step-3: Choose the amount N for decision trees that you simply want to create. Step-4: Repeat Step 1 & 2. Step-5: for brand spanking new data points, find the predictions of every decision tree, and assign the new data points to the category that wins the bulk votes. Now we'll implement the Random Forest Algorithm tree using Python. For this, we'll use an equivalent dataset "user_data.csv", which we've utilized in previous classification models. By using an equivalent dataset, we will compare the Random Forest classifier with other classification models like Decision tree Classifier, SVM, rectilinear regression, etc.

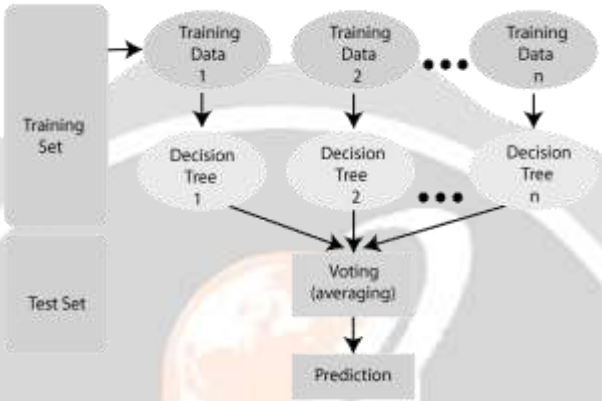

Fig 3. Random forest classification

III. Linear Regression:- - rectilinear regression is one among the simplest and hottest Machine Learning algorithms. it's a statistical procedure that's used for predictive analysis. rectilinear regression algorithm shows a linear relationship between a dependent (y) and one or more independent (y) variables, hence called as rectilinear regression. Since rectilinear regression shows the linear relationship, which suggests it finds how the worth of the variable is changing consistent with the worth of the experimental variable. The rectilinear regression model provides a sloped line representing the connection between the variables. Consider the below image.

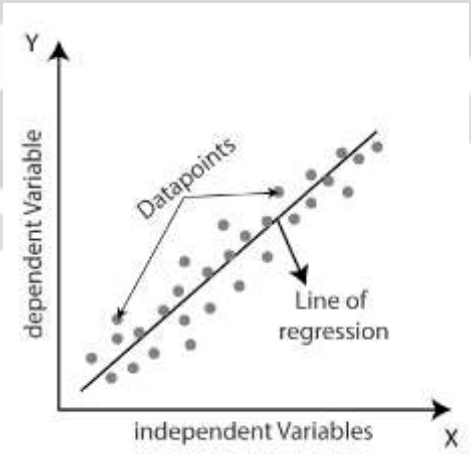

Fig 4.Linear regression relationship variables.

Mathematically, we can represent a linear regression as:

$y= a_0+a_1x+ \varepsilon$

Y= Dependent Variable (Target Variable)
X= Independent Variable (predictor Variable)
a0= intercept of the line (Gives an additional degree of freedom)
a1 = Linear regression coefficient (scale factor to each input value).
ε = random error

The values for x and y variables are training datasets for Linear Regression model representation. A linear line showing the relationship between the dependent and independent variables is called a **regression line**.

4. RESULT ANAYSIS

Fig 5. Result Analysis

Accuracy Comparison Graph : One of the more obvious metrics, it is the measure of all the correctly identified cases. It is most used when all the classes are equally important.

Calculation formula :

$$\text{Accuracy} = \frac{\text{True positive+True Negative}}{\text{(True positive+False Negative+True Negative+False Positive)}}$$

Precision Comparison Graph : It is implied as the measure of the correctly identified positive cases from all the predicted positive cases. Thus, it is useful when the costs of False Positives is high.

Calculation formula :

$$\text{Precision} = \frac{\text{True positive}}{(\text{True positive} + \text{False Positive})}$$

Recall Comparison Graph : It is the measure of the correctly identified positive cases from all the actual positive cases. It is important when the cost of False Negatives is high.

Calculation formula :

$$\text{Recall} = \frac{\text{True Positive}}{(\text{True positive} + \text{False Negative})}$$

F-Measure Comparison Graph : F-measure is a single measure of classification procedure's usefulness. F-measure consider both the precision and the recall of the procedure to compute the score. The higher the f-measure the better the predicative power of the classification procedure. A measure of 1 means the classification procedure is perfect. The lowest position f-measure is 0.

0<=F <=1

Calculation formula :

$$\text{F} - \text{measure} = \frac{2 * \text{Precision} * \text{Recall}}{(\text{Precision} + \text{Recall})}$$

5. CONCLUSIONS

Algorithms for analyzing and classifying medical images have gained a great level of attention recently. The experiments we present in this work show that after preprocessing MRI images, neural network classification algorithm was the best. SVM (Support Vector Machine) did very well and came in second. Random Forest and Linear Regression came in last. A much higher accuracy can be achieved by gaining a better dataset with high-resolution images taken directly from the MRI scanner. Moreover, classifier boosting techniques can be used to raise the accuracy even higher and reach a level that will allow this tool to be a significant asset to any medical facility dealing with brain tumors

6. FUTURE SCOPE

Encouraged by these results, future work will involve the improvement of classification result and overall accuracy. The number of output classes can also be increased if more data is available. With a more extensive and diverse dataset, the overall classification accuracy can be dramatically increased. Another approach to

improve the result would be to increase the number of hidden layers of the KNN (K-Nearest neighbors). By increasing the number of hidden layers, the weights will be better adjusted and thus increase the classification.

7. REFFRENCES

[1] N. N. Gopal and M. Karnan. Diagnose brain Komal Sharma, AkwinderKaur, ShrutiGujral," Brain Tumor Detection based on Machine Learning Algorithms",International Journal of Computer Applications (0975 – 8887)Volume 103– No.1, October 2014 http://iriset.org/images/extraimages/Paper%20ID%201411003B%20.pdf

[2] Tumor through MRI using image processing clustering algorithms such as Fuzzy C Means along with intelligent optimization techniques.In IEEE International Conference on Computational Intelligence and ComputingResearch (ICCIC), pages 1–4, Dec 2010. https://www.researchgate.net/publication/224216219_Diagnose_brain_tumor_through_MRI_using_image_processing_clustering_algorithms_such as_Fuzzy_C_Means_along_with_intelligent_optimization_techniques

[3] a tarajan P, Krishnan.N, Natasha SandeepKenkre, Shraiya Nancy, Bhuvanesh PratapSingh,"Tumor Detectionusing threshold operation in MRI Brain Images" , IEEEInternational Conference on Computational Intelligence and Computing Research, 2012. https://ieeexplore.ieee.org/document/6510299

[4] SuchitaGoswami, Lalit Kumar P. Bhaiya, " Brain Tumor Detection Using Unsupervised Learning based Neural Network" , IEEE International Conference on Communication Systems and Network Technologies,2013. https://www.scribd.com/document/396914931/9db73a242801430602c108786cbe6796b7f2

[5] Mahmoud Al-Ayyoub , "Machine learning approach for brain tumor detection " , Conference: Proceedings of the 3rd International Conference on Information and Communication Systems, April 2012. https://www.researchgate.net/publication/262163807Machine_learning_approach_for_brain_tumor_detectionhttp://sersc.org/journals/index.php/IJGDC/article/view/13930

[6] Dr. U. Srilakshmi et al. " Segmentation of MRI Image for the Detection of Brain Tumour Segmentation UsingClustering Methods and Algorithms ",April 2020. http://sersc.org/journals/index.php/IJGDC/article/view/13930