International Journal of Scientific Research & Engineering Trends
Volume 9, Issue 1, Jan-Feb-2023, ISSN (Online): 2395-566X

Machine Learning Based Approach for Brain Tumor Detection

Dr.E.Shanmugapriya
College Of Engineering Guindy,
Chennai, India

O.Rajasekar
College Of Engineering Guindy,
Chennai, India

***Abstract* – 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 objective of this project is to analysis the use of pattern classification methods for distinguishing different types of brain tumors, such as primary gliomas from metastases, and also for grading of gliomas. The availability of an automated computer analysis tool that is more objective than human readers can potentially lead to more reliable and reproducible brain tumor diagnostic procedures. A computer-assisted classification method combining conventional MRI and perfusion MRI is developed and used for differential diagnosis. The proposed scheme consists of several steps including ROI definition, feature extraction, feature selection and classification. The extracted features include tumor shape and intensity characteristics as well as rotation invariant texture features. Feature subset selection is performed using Support Vector Machines (SVMs) with recursive feature elimination. The Convolution neural network 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 convolution neural network algorithm. The proposed work is divided into three parts: preprocessing Segmentation and classification steps are applied on brain MRI images, texture features are extracted using Gray Level Co-occurrence Matrix (GLCM),DWT and then classification is done using svm algorithm.**

***Keywords* – Brain tumor CNN, Classification, GLCM, Matlab, MRI, Segmentation SVM.**

I. INTRODUCTION

Medical image encompasses different image modalities and processes to image the human body for treatment and diagnostic purposes and hence plays a paramount and decisive role in taking actions for the betterment of the health of the people. Image segmentation is a crucial and essential step in Image processing which determines the success of a higher level of image processing the primary goal of image segmentation in medical image processing is mainly tumor or lesion detection, efficient machine vision and attaining satisfactory result for further diagnosis.

Improving the sensitivity and specificity of tumor or lesion has become a core problem in medical images with the help of Image calcification Brain and other nervous system cancer is the 8th leading cause of death, and the three-year survival rate for people with a cancerous brain is 34% for men and 36% for women. Moreover, the World Health Organization (WHO) states that around 400,000 people in the world are affected by the brain tumor and quality of images. The Image Segmentation is the process of extracting or splitting a digital image into multiple regions. The main purpose of segmentation is to change the representation of image into easier analyses and it would be meaningful. The segmented image has the meaningful objects and it analyse the segmented image. It is to locate the object and boundaries. Medical Image segmentation plays a significant role in diagnosis of clinical. The major problems in medical imaging have poor contrast, different types of noises, missing of boundaries. MRI scan gives comfort to analyse the data compared to Computed Tomography (CT). MRI scan do not affect the human body because it is based on magnetic field and radio waves whereas CT scan is harmful and it is based on radiation. Segmenting the tumour from the images have been in various forms to analyse them Many segmentation techniques are there which are broadly used as artificial neural network, we have more learning techniques also. From these we have to use the machine learning. CNN technique is used to segment the tumors.

II.RELATED WORK

identified the tumor by combining spatial fuzzy c-mean (SFCM) and region growing. The region growing methods required to initialize the seed points. Mostly, their

segmentations were not enough accurate when the initializations were unsuitable. These poor initial conditions limit their success for a larger number of data sets. Therefore, more sophisticated methods are needed. Various convolutional neural networks (CNNs) have been designed for brain tumor segmentation. The idea behind these networks is to learn the complex hierarchical features directly from domain data[11]

presented Nexus architectures for the segmentation of gliomas. In the Nexus model, the output of a basic CNN is treated as an additional input for a subsequent CNN. All of these CNNs are based on a single-label prediction scheme. Single-label prediction architecture takes image patches (local areas of the image) as input and categorizes the patch's central voxel into a tumor or non-tumor class. The segmented image is achieved by predicting the central voxel class of the input image patches. Therefore, single-label prediction networks are very slow in the conclusion stage. Subsequent advances in CNNs were in the introduction of dense prediction networks. These are superior to the traditional CNNs in predicting the labels of voxels within the input patch simultaneously. Fully convolutional network (FCN), a typical example of dense prediction network, was first used to segment brain tumors by Shen et al[12] 3D networks can directly process 3D MRI data. However, they require high GPU memory and increase computational cost. The segmentation algorithm may mistakenly detect small areas of the tumor as healthy tissue, and vice versa. Hence a post-processing step has been used in some studies to refine the result of segmentation.

For example, Pereira et al.[13] designed two-step post-processing: initially, pseudo glioma areas were removed using the region merging algorithm based on MDL criteria. Then, the improved distance regularized level set evolution method was used to eliminate small holes and correct tumor boundaries. In addition to these methods, simple operators such as mathematical morphology[14]. The machine learning domain offers the opportunity to fabricate or innovate a fully automated and non-invasive Glioma classification systems or applications which is based on using the different types of features like Intensity, texture, morphological based which are retrieved or obtained from the Medical imaging modalities like MRI, CT scans. In turn these derived features could be used to train various machine learning classifiers for accurately classifying Glioma brain tumor into low or high grade. When it comes to Glioma Classification mostly MRI sequences are used by the radiologist so does by the researchers who are working in this field of Glioma classification with the aid of machine learning [15, 16].

This paper simply presents the efficacies of different machine learning based approaches developed over the years for the purpose of Glioma classification in this section. A brief comparison is presented among all such approaches in this section. The approach to grade Glioma based on the Histogram signatures measured using cerebral blood volume (CBV) values taken from the Tumor ROIs in the DSC-MRI. These histogram signatures are used to train the SVM classifier used in this approach with different parameters values (number of support vector) and delivers result in the form of True positive rate (TPR) of 0.76 and True negative rate(TNR) of 0.88. Then Zacharaki et al. [28] come with an automated approach based on support vector machine for the binary classification of first classification of metastases from Gliomas and then classification of low grade Glioma against high grade with an accuracy of 88%. The shape based, texture and intensity based features are derived from the region of interest in this approach which is further used to train the SVM classifier.

III. METHODOLOGY

Detection and classification of MRI brain tumor images involve four modules: Pre-processing, Segmentation, Feature Extraction and Classification the steps followed for the implementation are depicted in Figure-1. Noise removal is used for pre-processing, and the pre-processed image is given as input for the segmentation process. Segmentation algorithms are used to extract the tumor regions: Convolution neural network. Statistical textural features are obtained by using GLCM. The obtained features are classified using SVM four type of tumors. That experiment has been implemented using MATLAB R2020a.

1.Data set

Dataset of MRI brain tumor images these images T1 and T2 weighted sequence images. the MRI images provided by Aarthi hospital, Chennai. These images are labeled by the expert radiologist.

IV. BRAIN TUMOR CLASSIFICATION

1. Preprocessing

MRI brain images cannot be fed directly as the input for the segmentation technique. Therefore noise removal is performed. The median filter is a nonlinear digital filtering technique, often used to remove noise. Such noise reduction is a typical pre-processing step to improve the results of later processing Median filtering is very widely used in digital image processing because, under certain conditions, it preserves edges while removing noise.

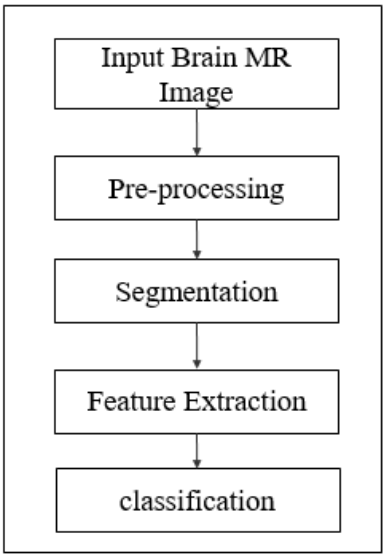


Fig.1. Block diagram

The main idea of the median filter is to run through the signal entry by entry, replacing each entry with the median of neighboring entries. The pattern of neighbors is called the "window", which slides, entry by entry, over the entire signal For 1D signals, the most obvious window is just the first few preceding and following entries, whereas for 2D (or higher dimensional) signals such as images, more complex window patterns are possible . Note that if the window has an odd number of entries, then the median is simple to define: it is just the middle value after all the entries in the window are sorted numerically. For an even number of entries, there is more than one possible median. This filter enhance the quality of the MRI image. Filter image shown blow Figure 4.

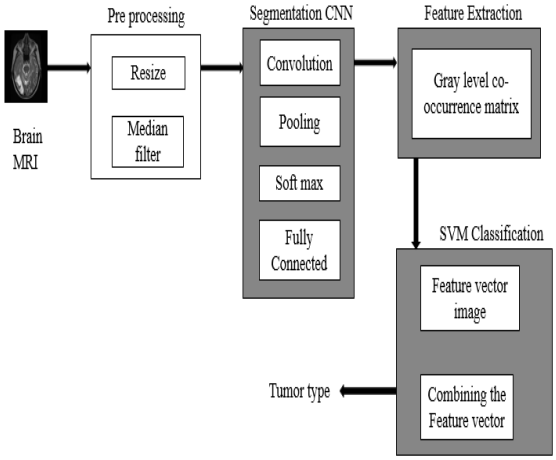


Fig. 2. Block diagram for classification of tumor

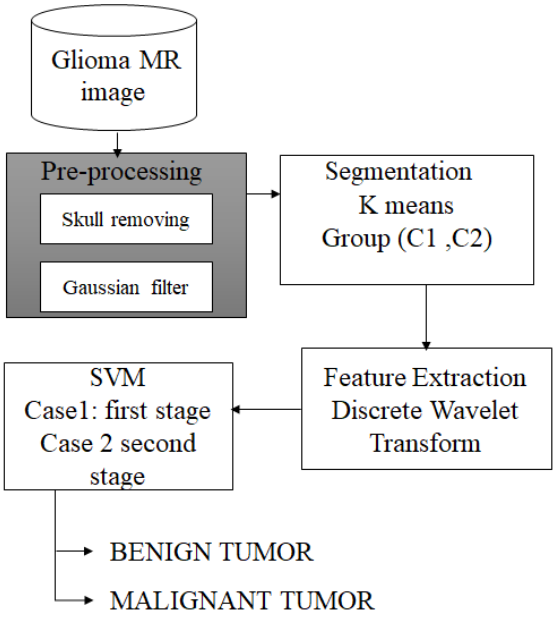


Fig. 3. Block diagram for Grading of glioma tumor

2. Segmentation

2.1Convolutional Neural Network Design:

1. The construction of a convolutional neural network is a multi-layered feed-forward neural network, made by assembling many unseen layers on top of each other in a particular order.
2. It is the sequential design that give permission to CNN to learn hierarchical attributes.
3. In CNN, some of them followed by grouping layers and hidden layers are typically convolutional layers followed by activation layers.
4. The pre-processing needed in a ConvNet is kindred to that of the related pattern of neurons in the human brain and was motivated by the organization of the Visual Cortex. After filtering the MRI images fed to segmentation. It's based on Convolution neural network

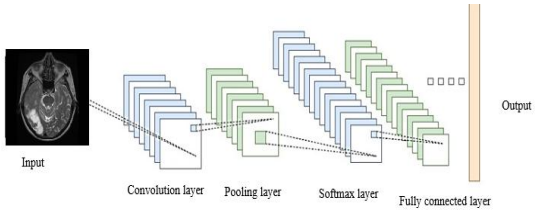


Fig. 4. CNN Layers

Convolution layer generates feature maps from MRI image next Reduce the Dimensional of feature map without any chances in an image This function is to progressively reduce the spatial size of the representation to reduce the amount of parameters and computation in the network .pooling layer operates on each feature map independently. Segment the normal and up normal cells in

image fully connected layer merging the all soft max value finally get the segmented MRI image

2.2. Feature Extraction

Features are said to be properties that describe the whole image. It can also be referred as an important piece of information relevant for solving the computational task related to specific application. From the segmented image, features are computed to encode the useful diagnostic information.

2.3Gray Level Co-Occurrence Matrix

A statistical method of examining texture that considers the spatial relationship of pixels is the GLCM, also known as the gray-level spatial dependence matrix. The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting the statistical measures from this matrix. GLCM calculates the co-occurrence matrix of an image by computing how often a pixel with a certain intensity "i" occurs in relation with other pixel, j" at a certain distance "d" and orientation. The gray-level co-occurrence matrix can reveal certain properties about the spatial distribution of the gray levels in the texture image. In order to extract information. Concerning spatial distribution and its orientation, four co-occurrence sub matrices can be calculated across four scanning directions θ (0, 45, 90,135) degrees. The features extracted using GLCM method are: contrast, correlation, energy, homogeneity, Entropy which are shown below :

3.1 Contrast: It define the intensity of pixels in the MRI images and their neighboring or adjacent pixels.

$$C(k,n) = \sum_i \sum_j (i-j)^k P_d[i,j]^n$$

3.2 Homogeneity: A homogeneous image will result in a co-occurrence matrix with a combination of high and low P[i,j]'s

$$C_h = \sum_i \sum_j \frac{P_d[i,j]}{1+|i-j|}$$

3.3 Correlation: Correlation is a measure of image linearity

$$C_c = \frac{\sum_i \sum_j [ijP_d[i,j]] - \mu_i \mu_j}{\sigma_i \sigma_j}$$

***3.3* Entropy:** Entropy is a measure of information content. It measures the randomness of intensity distribution.

$$C_e = -\sum_i \sum_j P_d[i,j] \ln P_d[i,j]$$

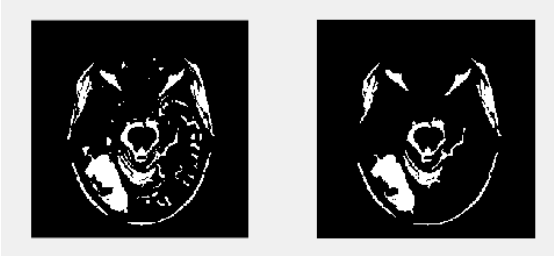


Fig.5. The GLCM calculated values in table GLCM values are calculated 0 degree scanning direction

TABLE 1 GLCM Calculated Values

Contrast	Homogeneity	Correlation	Entropy
0.2219 T1 weighted	0.8834 T1 Weighted	0.8957 T1 weighted	0.89 T1 weighted
0.2453 T2 Weighted	0.8898 T2 weighted	0.8452 T2 weighted	0.8945 T2 weighted

V. CLASSIFICATION

Algorithm
Start training
-class name = Glioma, Meningioma, Metastasis and Astrocytoma
For i= 1 => Tdata
If T data is tumor

Class 1 = T data 1=> Glioma
/ Feature of MRI from GLCM is labeled as class 1
else if T data meningioma

Class 2 = T data 1=> Meningioma
/ Feature of MRI from GLCM is labeled as class 2
else if T data glioma

Class 3 = T data 1=> Metastasis
/ Feature of MRI from GLCM is labeled as class 3
else if T data Astrocytoma

Class 1 = T data 1=> Astrocytoma / Feature of MRI from GLCM is labeled as class 4 else if T data Metastas is After extracting features using GLCM, they are directly given to the support vector machine (SVM) for the classification. The process involves two phases: Training phase and testing phase. In training phase the patterns in terms of features and class labels of Glioma, Meningioma, Metastasis and Astrocytoma tumor are fed to the classifier for training. In testing phase test pattern is fed and knowledge gained during training phase will classify the unknown pattern. SVM performs the robust non-linear classification with the kernel trick .It finds the separating

hyper plane in some feature space inducted by the kernel function while all the computations are done in the original space itself. For the given training set, the decision function is found by solving the convex optimization. The training and test images taken as table 2

VI. RESULT AND DISCUSSION

This section provides the experimental results of the present work.

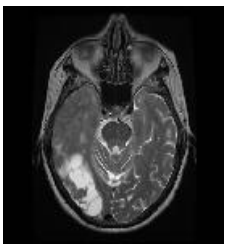


Fig. 6
Input mage

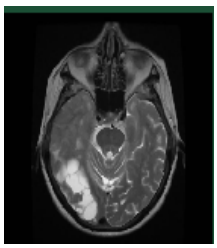
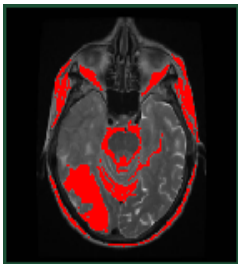


Fig. 7
Filtered image



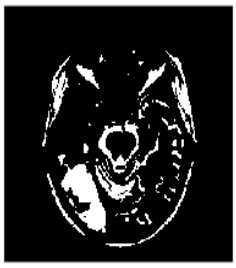
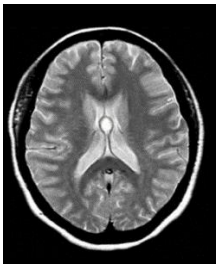


Fig. 8: Segmented MR Image

Fig. 9
Extracted Image

Fig. 10
classified image (Metastasis)



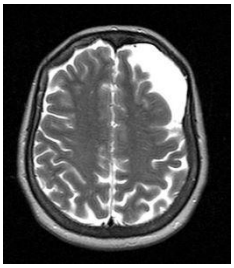
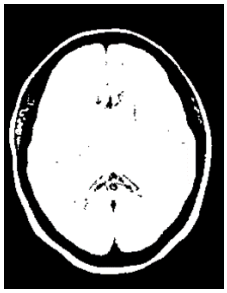


Fig. 11: Glioma MRI



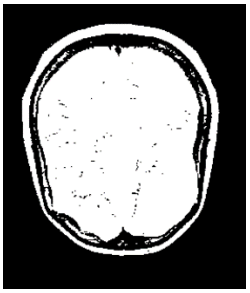
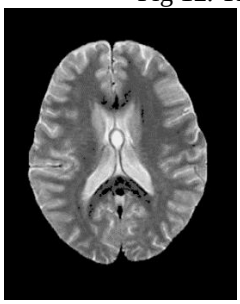


Fig 12: Threshold image



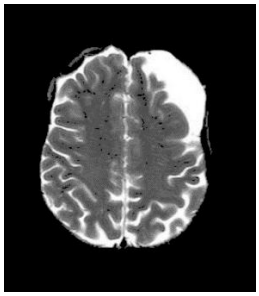


Fig 13: preprocessed image

VII. GLIOMA SEGMENTATION

Clustering refers to the process of grouping samples so that the samples are similar within each group k-means clustering treats each object as having a location in space. It finds partitions such that objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible. K-means clustering requires the number of clusters to be partitioned and a distance metric to quantify how close two objects are to each other. In image analysis, clustering can be used to find groups of pixels with similar gray levels, colors or local textures in order to discover the various regions in the image.

1. Kmeans

K-means clustering algorithm, which is an unsupervised method, to provides a segmentation of the image. K-means clustering was used since it is simple and has relatively low computational complexity. In addition, it was suitable for biomedical image segmentation as the number of clusters (k) is known for images of particular regions figure15 show the segmentation . The procedure followed a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. The main idea was to define k centroids, one for each cluster.

These centroids should be placed in a cunning way because of different location causes different result. So, the better choice is to place them as much as possible far away from each other. The next step was to take each point belonging to a given data set and associate it to the nearest centroid. When no point is pending, the first step is completed and an early group age was done. At this point

k new centroid were calculated as barycenter's of the clusters resulting from the previous step. After having these k new centroid, a new binding has to be done between the same data set points and the nearest new centroid. A loop has been generated. The result of this loop was that the k centroids have changed their location step by step until no more changes were done. In other words centroids did not move any more. Finally, this algorithm aims at minimizing an objective function, in this case a squared error function.

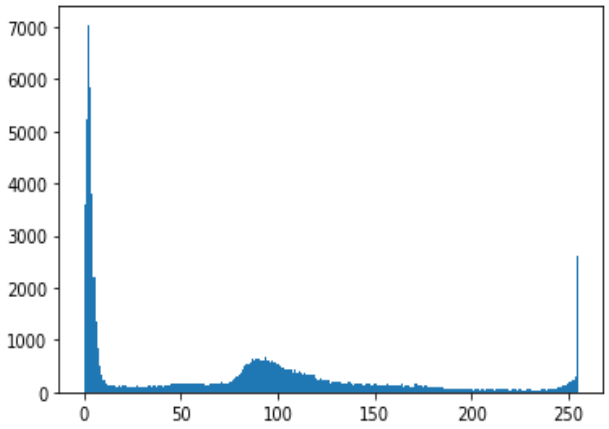


Fig.14. histogram of the intensities.

As a part of this project, an iterative version of this algorithm was implemented. The algorithm takes a 2 dimensional image as input. Various steps in the algorithm are as follows

Algorithm

1. CCompute the intensity distribution also called the histogram of the intensities (figure 14)
2. Initialize the centroids with k random intensities
3. Repeat the following steps until the cluster a label of the image does not change anymore.
4. Cluster the points based on distance of their intensities from the centroid intensities.
5. Compute the new centroid for each of the clusters.
6.

2. Feature Extraction

Discrete Wavelet Transform (DWT), used to convert time domain data into a discrete wavelet domain. It provides time and frequency analysis of an MRI image and is used for feature extraction, with the help of DWT MRI images are easily analyzed at different levels of resolution. Further PCA technique is used for feature reduction, to reduce the complexity of the model as well as unused features of input MRI images. The extracted feature values show in table 4

$$\begin{cases} d_{j,k} = \sum x(n) h_j^*(n-2jk), \\ a_{j,k} = \sum x(n) g_j^*(n-2jk). \end{cases}$$

Where

d j,k =detail component of signal x(n)

a j,k= approximation component of signal x(n)

h(n)= coefficients of the high-pass filter

g(n)= coefficients of the low-pass filter

j, k = wavelet scale & translation factor

Ts 2 Extracted Features from Glioma MRI Image

Features	Image T1	Image T2
Mean	0.2408	0.0053
Entropy	0.001	1.063
Smoothness	0.743	0.954
RMS	0.8359	0.089
Standard deviation	0.843	0.08734
Variance	0.839	0.087
Kurtosis	24.84	32.1
homogeneity	0.9402	0.9573
Contrast	0.560	0.4032
Correlation	0.7628	0.1786
Energy	0.855	0.87954

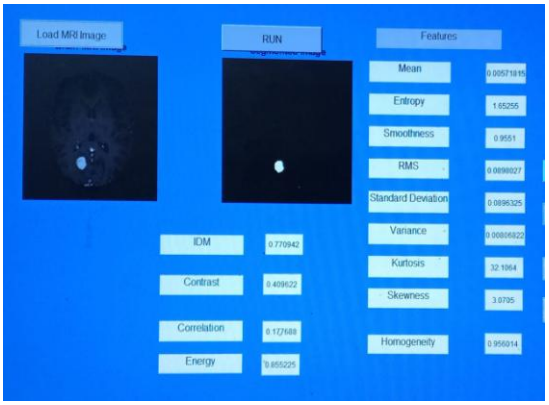


Fig.15: segmentation

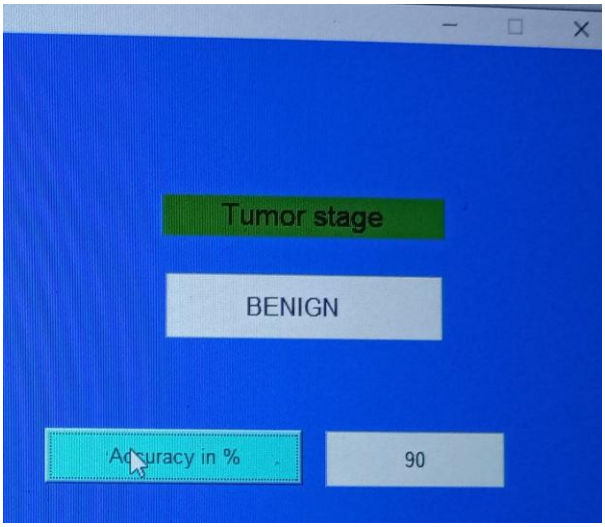


Fig 16 classification

Classification of Glioma

After implementing features extractions and selecting the best performed ones, we began the classification which contains two phases. In the learning phase, the data clearly

represents the nature of ROIs and teach the classifier whether it's benign or malignant tumor. figure 16 show the Classification sample result

Algorithm

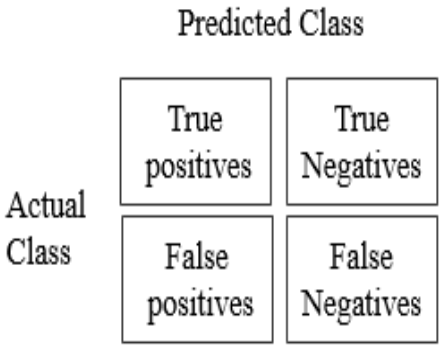
Start train
SVM initialization
Class name benign, malignant
For i =>B Data
If B -data (is subset of benign)
=> is ladled grad 1/level 1
Else –data (is subset of malignant)
=> is ladled grad 2/level 2
End if
End for
Svm structure =FIT B(B data, grad, leaner, class name)
Classification model
Current features = features of ROI
Result =run (svm structure, current feature)
If result 1
Grad =benign
Else result 2
Grad= malignant
End if

Performance Analysis classification of brain tumor:

Evaluation of the proposed MRI brain images classification method is carried out by calculating the performance in terms of true positive, false positive, true negative and false negative. By using these matrices sensitivity, specificity and accuracy can be obtained.

Table 3 RATIO of images used for training and testing

Class	Number of images	Training set (70%)	Testing set (30%)
Glioma	247	173	74
Meningioma	43	30	13
Metastasis	72	50	22
Astrocytoma	58	41	17



TP→ Number of pixels in the true tumor area
FP→ Number of pixels in the false tumor area
TN→ Number of pixels in the true non- tumor area
FN→ Number of pixels in the false non- tumor area

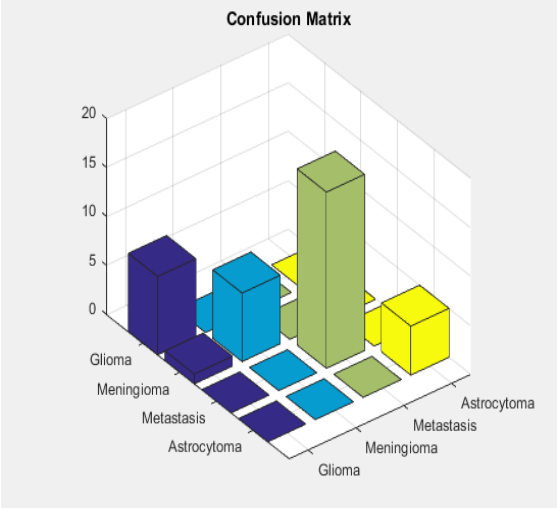


Fig.17. Confusion matrix.

Calculation of the accuracy, sensitivity and specificity is done according to these equations:

- SENSITIVITY = TP/ (TP+FN)
- SPECIFICITY = TN / (TN +FN)
- ACCURACY = TP+FN/ (TP+TN +FP+FN)

As the data base contains 420 MRI brain images of four cases (Glioma, Meningioma, Metastasis and Astrocytoma), the dataset is split. In order to split the feature subset into the training and testing sets, the Holdout method is selected, which is part of the model evaluation used to segregate training and test datasets, where two-third (70%) of the samples from all the classes are allocated to the training set and the remaining onethird (30 %) of the samples from all the classes are allocated to the testing set.

Table 4 Performance Evaluation of Multi class SVM

	Accuracy	Sensitivity	Specificity
Glioma	97%	100 %	96.8%
Meningioma	97.1%	100%	90.5%
Metastasis	97.9%	100%	87%
Astrocytoma	97.2	100%	89.3%

Table 5: Performance Evaluation of Glioma grad

	Benign	Malignant
Accuracy	90%	93.2 %
Polygonal Accuracy	84.6%	98%
Quadratic Accuracy	90.3%	93.01%
RBF Accuracy	95.7	97.5%

VIII. CONCLUSIONS

An improvement in the SVM for classification of brain tumor is present in the work based on the segmentation of tumor region with Convolution neural network technique. The application of improved segmentation technique is very crucial and important step that is used to analyze and segment the extract tumor region .this is the most important step in the diagnosis and provides the improved brain tumor classification results in this work four tumor categories are found in which features are selected and is used for the input of the SVM differentiate the types of tumor. After classification the Glioma tumor is graded in to benign tumor and malignant tumor . The Sensitivity Accuracy and specificity values are 100%, 97..2%and 90% respectively .In future we can enhance the extraction of features by using various approaches .

REFERENCES

[1].Brain tumor image segmentation via asymmetric/ symmetric UNet based on two pathway -residual blocks, Motahareh Aghalari, Ali Aghagolzadeh , Mehdi Ezoji -2020 Journal home page, Elsevier

[2]. Brain MRI Image Classification for Cancer Detection Using Deep Wavelet Auto encoder-Based Deep Neural Network,

[3] Pradeep Kumar Mallick, Seuc Ho Ryu (2019), IEEE Segmentation of Brain Tumour and Its Area Calculation in Brain MR Images using K-Mean Clustering and Fuzzy C Mean Algorithm, Mr.Rohit S. Kabade (2013),IEEE

[4] BRAMSIT: A Database for Brain Tumor Diagnosis and Detection , R.Tamilselvi , A.Nagaraj

[5]. Evaluation of Three Methods for MRI Brain Tumor Segmentation , R. B. Dubey, M. Hanmandlu (2011),IEEE

[6]. Performance Analysis of fuzzy c means algorithm in automated detection of brain tumor, R.Preetha,G.R.Suresh , (2014)

[7]. Brain tumor extraction from MRI images using matlab, Rajesh C. Patil, Dr. A. S. Bhalchandra, IJECSE

[8] Brain MRI Image Classification using Image Mining Algorithms, Vaibhavi Solanki, Ms. Vibha Patel,(2018)

[9]. Detection of Brain Tumor in MRI Images, using Combination of Fuzzy C-Means and SVM Parveen, Amritpal singh, (2015)

[10]. A new hybrid clustering algorithm based on K-means and ant colony algorithm, Jue Lu, Rongqiang Hu (2013)

[11]. M. Havaei, et al., Brain tumor segmentation with deep neural networks, Med. Image Anal. 35 (2017) 18–31.

[12]. H. Shen, J. Zhang, W. Zheng, Efficient symmetry-driven fully convolutional network for multimodal brain tumor segmentation. in 2017 IEEE International Conference on Image Processing (ICIP). IEEE. 2017.

[13].S. Pereira, et al., Brain tumor segmentation using convolutional neural networks in MRI images, IEEE Trans. Med. Imaging 35 (5) (2016) 1240–1251

[14]. S. Hussain, S.M. Anwar, M. Majid, Segmentation of glioma tumors in brain using deep convolutional neural network, Neurocomputing 282 (2018) 248–261.

[15]. Gordillo N, Montseny E, Sobrevilla P (2013) State of the art survey on MRI brain tumor segmentation. Magn Reson Imaging 31(8):1426–1438

[16]. Sajjad M, Khan S, Muhammad K, Wu W, Ullah A, Baik SW (2019) Multi-grade brain tumor classification using deep CNN with extensive data augmentation. J Comput Sci 30:174–182

[17]. Zacharaki EI, Wang S, Chawla S, Soo Yoo D, Wolf R, Melhem ER et al (2009) Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. Magn Reson Med 62(6):1609–1618.