Brain Tumor analysis Based on Shape Features of MRI using Machine Learning

Bhagyashri H. Asodekar
Dept. of Computer Engineering
Pimpri Chinchwad College of Engg.
Pune, India
bhagyashriasodekar@gmail.com

Sonal A. Gore
Dept. of Computer Engineering
Pimpri Chinchwad College of Engg.
Pune, India
Sonalgore@gmail.com

A. D. Thakare
Dept. of Computer Engineering
Pimpri Chinchwad College of Engg.
Pune, India
Adthakare2014@gmail.com

Abstract—**Nowadays there is need to develop Computer Aided Diagnosis (CAD) systems for diagnosis of brain tumor. Brain tumor detection at early stage has become very important. In experimentation, brain tumor magnetic resonance images (MRI) are used to detect and classify the malignant and benign brain tumors. MRI images are extracted from MICCAI BraTS 2015 dataset. Brain tumor are segmented by using image processing techniques. For feature extraction shape-based features are used. Extracted shape-based features are fed to machine learning algorithms as support vector machine and random forest algorithm to classify benign and malignant brain tumors. It achieved the accuracy for random forest is 86.66%.**

Keywords—MRI, brain tumor, Computer Aided Diagnostic system, image processing, shape features, classification

I. Introduction

According to the Central Brain Tumor Registry of the United States, rates of incidence happened in US use the 2000 united states standard population for 2011-2015 as per the 100,000 population. In 2015, United states has 166,039 people were suffered with brain tumors. Brain tumor is mass of uncontrolled or abnormal growth of cells within skull. Many different types of brain tumor exist. Some brain tumors are cancerous (malignant) and, some brain tumors are non-cancerous (benign). The growth rate of brain tumor and location of brain tumor decides the how it will affect the functions of your nervous system. In recent years, Medical image processing plays a vital role in brain tumor diagnosis system. Detection of brain tumor at early stage is necessary to prevent death and increase survival rate. Benign tumors are mostly non-cancerous and they don't have much growth but malignant tumors are cancerous and their growth rate is very fast and its dangerous to the health [1].

For medical imaging so many different techniques are used such as computed tomography (CT), X-rays, Magnetic resonance images (MRI). Among all these MRI is advanced uses the magnetic field and it will gives the broad information about the soft tissues inside the body which is helpful to plan the surgery and diagnosis the disease [2][13]. To develop CAD system with machine learning based required feature extraction of brain tumor image. Most of the texture feature based research are done for brain tumor. Very few research works done on shape feature based brain tumor detection and classification. In this study, brain tumor classification based on shape feature with machine learning using MRI images is proposed [3]. Benign and malignant tumors are classified using machine learning. For segmentation of region of interest of brain tumor image processing techniques are used as discussed below. Features are extracted by using shape features and these shape features are fed to classifier for classification of brain tumor as benign and malignant brain tumor.

II. Related work

Nilakshi Devi et al. [4] work focused on automatic brain tumor detection and classification of grades of astrocytoma. GLCM is used for the feature extraction technique and artificial neural network is used to detection of tumor. The next step is classification of astrocytoma using RBFN. Grade I,II, III, IV type of images are used.

Mustafa, Ikhlas qader et al [5] paper focused on brain tumor classification using statistical features and back propagation of neural network. For feature extraction DWT and Gabor filter used. Classification done for three types of tumor Meningioma, glioma, pituitary brain tumors. For classification back propagation neural network used.

Lichi Zhang, Han Zhang et al. [6] worked on malignant brain tumor classification using random forest method. Subject wise means age, gender and size of tumor and patch wise means intensity histogram and haar -like features are extracted. This extracted feature used for classification using random forest classifier with 59.88% accuracy achieved.

Sehgal, Aastha et al. [7] work represented for automatic brain tumor segmentation and extraction in MR image. For detection of brain tumor given method is fully automatic. Experiment performed with the BraTS dataset. The worked done with various image processing, method consists of various stages, viz., Image Acquisition, and in Preprocessing stage includes the noise removal and enhance the MRI brain images using mathematical operations or altering techniques. In first step, the MRI images are N3 corrected by using the N3 algorithm to remove bias field [18][15]. Fuzzy C Means technique is used for segmentation tumor extraction and evaluation. For tumor region extraction method Area and Circularity is used and Dice coefficient is calculated [17].

Sunit Sivaraj, Dr. B. Surendiran, Hariharan. et al [8] the shape-based features are categorized on the basis of the shape and margin of the mass. This is reason to the growth of tumor masses arise from one spot and grow circumferentially. The mass of tumor is circular in shape and it may be of irregular in shapes. Shape of the brain tumor region of interest is extracted from brain tumor image for classifying the tumor. Shape and

978-1-7281-4042-1/19/$31.00 ©2019 IEEE

texture features such as circularity, convex area, thinness ratio, equivalent diameter, entropy, shape index, standard deviation of edge, etc. are calculated for the classification.

Gupta, Sheifali, et al. [9] By using image processing techniques segmentation of brain magnetic resonance images. Classification of astrocytoma as low grade or high-grade with the k-NN classifier. Classification results are analyzed on the basis of performance of three evaluation parameters i.e., accuracy, sensitivity, and specificity. Dataset is collected from BraTs and it contains magnetic resonance images (MRI). For feature extraction process, extracted shape-based features. And clustering classification KNN classifier is used. Classified brain tumor images in two grades i.e., low grade or high grade.

Kimia Rezaei, Hamed Agahi et al. [10] In first step preprocessing of images noise removed by using median and wiener filter. They used DICOM images for their work. Extraction of texture feature for brain tumor. Malignant and benign tumors classified using support vector machine and K nearest neighbor classifier.

Anjali Joshi, V. Charan, Shanti et al. [11] paper focus on classification of malignant and benign tumor. Two stage segmentation method for MRI brain tumor is used. Gabor filter and contour level set segmentation are used. For detection of different regions of image contour level set method used. By using this method extraction of brain tumor part is achieved, which is useful for further feature extraction methods like size, shape, and density of the tumor[16]. Evaluation of work is based on the pixel values of tumor tissues. And the pixel values are higher than the normal tissue. For the shape features calculation area and shape matrix of tumor is considered. the metric value is 1 then it is benign tumor. And higher than 1 shape matrix then malignant tumor.

As per the related work very few researches done on the shape features of brain tumor classification compare to texture and intensity features. Shape features are plays major role in classification of benign and malignant tumor.

III. Proposed Method

In this study, two approaches are adapted, viz., image segmentation and image classification is given in Fig. 1.

A. Segmentation (Region of Interest)

In Preprocessing, ITK-SNAP gives semi-automatic segmentation using active contour methods, as well as manual delineation and image navigation. It will provide the 2 directional image for further image processing techniques for segmentation part. Region of Interest of brain image is done using image processing techniques and those segmented brain tumor images are used to extract the shape feature like center of gravity, circularity ratio, rectangularity, convexity, solidity of tumor region. These extracted brain tumors contains only brain tumor where all other objects are removed. These region of interest of tumor region are extracted using erosion, open, close and filtering of morphological operations. By using above methods we get expected ROI of brain tumor region which is helpful for further work. Explanation of each operation is given below:

- After extracting brain tumor MRI images from dataset, binarization of image is applied on it. The result of binary image is shown in Fig.2b. Image after conversion of binary image. All pixels having intensity value 0 and 1. So only few objects are visualized.
- After applying binarization, morphological erosion is applied on it. The result after erosion is shown in Fig.2c. The size of some objects are get reduced. And unnecessary objects are get removed. Morphological filtering and other morphological operations like opening and filling the object region is used given in Fig. 2d.
- In below fig.2e shows the sample of brain tumor segmentation result obtained using previous segmentation method. This only region of interest of tumor part from brain tumor image. After segmentation of brain tumor, brain mass is ready for feature extraction part. Shape features are extracted by using region of interest of brain tumor.

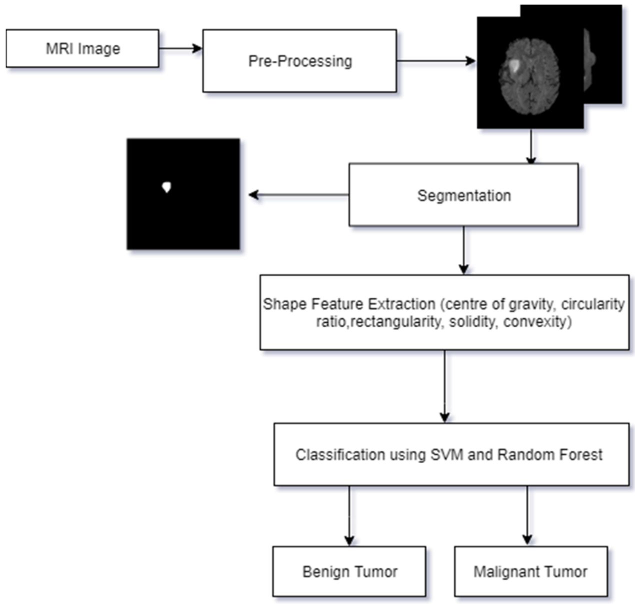

Fig. 1. Proposed Architecture

B. Shape features

In image processing, shape feature is important for image classification task. Most of the image retrieval systems are based on shape features. It gives the description of physical structure of the object. Here, different shape features are extracted from the segmented brain tumor to classify the tumor as benign and malignant brain tumor. Shape features are evaluated on the basis of how accurately those shape features allow one to retrieve the similar shape from database.

Centre of Gravity

Centre of gravity is defined as an object is the point at which weight is evenly scatter in different directions and all sides are in balance. For 2D image it is called centroid.

$$f(x,y) = \begin{cases} 1 \; if \; (x,y) \in M \\ 0 \quad otherwise \end{cases}$$

Where, M is the domain of binary shape image.
And centroid (gx, gy) is:

$$g_x = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad (1)$$

$$g_y = \frac{1}{N}\sum_{i=1}^{N} y_i \qquad (2)$$

Where N is the number of points in the shape of x and y directions,

$(x_i, y_i) \in \{(x_i, y_i) \mid f(x_i, y_i)=1\}$

Circularity Ratio

According to definition circle is the most circular object and it is most compact that it encloses most of the area for given perimeter. Circularity is used extensively in image processing to sort or identification of object.

$$\textit{Circularity ratio} = \frac{A_c}{p^2} \quad (3)$$

Where Ac is the 4*pi* area of shape,

p is the perimeter of shape.

Rectangularity

Rectangularity is defined as the property of being shaped like a rectangular. Standard approach to measure the rectangularity of shape is to ratio of area of given shape to the area of minimum bounding rectangle of the shape.

$$\textit{Rectangularity} = \frac{A}{A_b} \quad (4)$$

Where A is the area of shape, A_b area of the minimum bounding rectangle of shape.

Convexity

Convexity is defined as ratio of the perimeter of the convex hull to the original contour of perimeter of shape.

$$\textit{Convexity} = \frac{P_{convexhull}}{P} \quad (5)$$

Where, P is the perimeter of shape and convex hull perimeter of given shape.

Solidity

Solidity measures the density of the object. It is defined as the ratio of area of shape to the convex area of shape.

$$\textit{Solidity} = \frac{A}{CH} \quad (6)$$

Where, where, A is the area of the shape region, CH is the convex hull area of the shape. The solidity of a convex shape is always 1.

Centre of gravity, circularity ratio, rectangularity, convexity, solidity are the shape descriptors. Information for feature extraction is calculated from those shape features. It is discriminate the shape from the other shape. In machine learning based approach feature extraction is the major part.

C. Classification

This is another approach of medical imaging using shape based feature extraction for classification, where we assign the images to the class labels as per there appearance and features extracted by the shape feature technique. In this approach, MRI images extracted from MICCAI BRaTs 2015 dataset are used to classify the brain tumor as benign and malignant tumor. Proposed shape feature based classification technique is used to extract the features from brain images (MRI) and to classify them into benign and malignant.

Support vector machine is supervised learning algorithm. It is useful for classification and regression of data. There is binary and multiclass classification. It separates data in particular classes.

In two class problem, data is linearly separable. Let the dataset D as (x_1, y_1)…., (x_D, y_D). Where training tuples are X i and Labels are y i . Each label can take values +1 or -1. {Labels yi $\in$ [+1,-1]}.

Corresponding to the classes of Malignant= yes, Benign=no.

The support vector machine tries to search from all hyper planes that minimize the training error. In simple word, shortest distance from the hyperplane is equal to the shortest distance from other side margin. Margins are parallel to the hyperplane. Consider with maximum margin with the shortest distance hyperplane gives better accuracy.

$$W.X + b = 0 \quad (7)$$

where, weight vector is W, bias parameter is b.

Consider two input attributes, A1, A2, b is additional tuples. Then two cases are occurs, any point lies above the separating hyperplane and lies below the separating hyperplane. So, the weights can be adjusted to define the hyperplane sides.

For defining the maximal margin hyperplane (MMH) following formula is used for fulfill the constraint.

$$\frac{1}{2}\,||w||^2 \text{ with } (w.xi + b) \geq 1 \quad (8)$$

By using Lagrange multipliers formulation and then solved by the karush-kuhn-Tucker (KKT) theorem.

$$\text{Max} \sum_{i=l}^{l} a_i - \frac{1}{2}\sum_{i,j=1}^{l} y_i\, y_j a_i a_j \; x_i^T x_j \quad (9)$$

$$\text{Where} \sum_{i=l}^{l} a_i y_i = 0, \; a_i \geq 0 \quad (10)$$

And w is:

$$W=\sum_{i-l}^{l} a_i y_i x_i \quad (11)$$

y_i is the calss labels and $x_i x^T$ is a test tuples. l is the number of support vectors. In above case dot product between support vector $x_i x^T$. This will prove very useful for finding the MMH and support vector for the data are nonlinearly separable.

Random Forest is generally works on decision tree. Firstly it forms a number of decision trees and then each tree decides the number of attributes on basis of voting and after that selects the major numbered class. That is the final class for that attribute. RF tree is base Random Forest is nothing but another ensemble method (EoC). Decision tree classifier contains each of the classifiers in the ensemble. So, the collection of classifiers is a “forest”. By using a random selection of attributes, the individual decision trees are generated at each node to determine the split. Random vector sampled is depend on the each tree values independently. For all trees in random forest with the same distribution. In classification technique, separately voting done for each tree, and then finally which class get more vote that class is selected for final class. [14]

Consider M is training set, and n is the number of given tuples. d is the number of generated decision trees.

For decision tree, firstly choose one branch and select tuples. If needed again split that branch in next branch till we achieved exact class. This process repeats till cover n number of tuples and make the decision trees.

For each node having f number of attributes. Number of nodes is always smaller than f.

Random forest tree is typically grown using CART method. For construction of decision tree, Mi is the iteration, randomly select, at each node. For the split at the node F attributes as candidates. Maximum size trees are grown and are not pruned. This is the way of working and the selecting input for random forest.

There is another method of random forest which uses linear combination for the input method it uses random selection of attributes. By using given attributes, it creates new linear attributes. Selection of attributes and addition with the constant in all cases. The main goal is searching for the best division of parts with linear combination. When the only few attributes are present then this form of random forest is useful. So, they can reduce the correlation between the individual classifiers. The accuracy of random forest is more robust than AdaBoost classifier. Random forest algorithm works on very large databases. It also used for the regression method.

Performance Measures

The accuracy and the performance of the proposed system can be tested using confusion matrix. Outcomes of the prediction can be True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN). Low grade glioma images and High-grade glioma images are categorized correctly as True Negative (TN) and True Positive (TP). In some cases, test says no Low-grade glioma tumor, but it has High grade glioma tumor, then it is False Negative (FN) and test says it is High grade glioma tumor, but it has Low grade glioma tumor, then it is False Positive (FP). The performance of the binary classification test has two major statistical measures in medical field and they are sensitivity and specificity.

TABLE I. Evaluation Measures

Performance Measures	**Formulae**
Accuracy	TP+TN/TP+TN+FP+FN
Sensitivity	TP/TP+FN
Specificity	TN/TN+FP

IV. DATASET

The Multimodal Brain Tumor Segmentation (BraTS) is a challenge held annually since 2012 in conjunction with MICCAI conference [12]. All BraTS datasets share four MRI modalities. BraTS images are available in .mha format which we convert them to 2D images for image processing. Brain tumor dataset contains low grade gliomas and high-grade gliomas. It uses for the classification purpose.

V. RESULTS AND EVALUATION

As discussed in previous sections for segmentation of brain tumor. Image processing techniques are applied on the brain tumor images. As shown in Fig. 2.

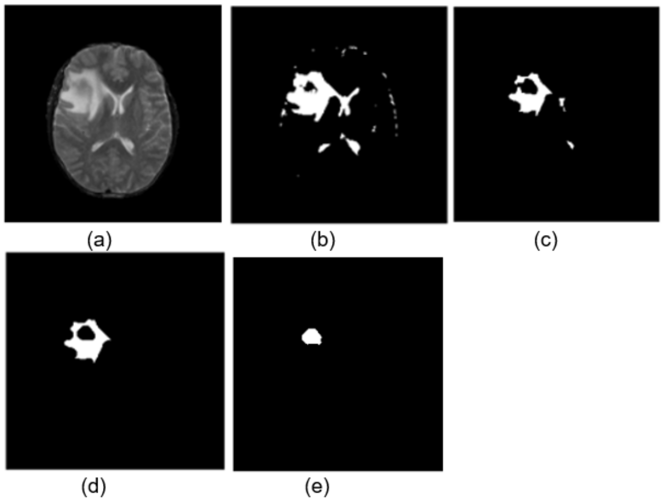

Fig. 2. Results of segmentation of brain tumor

MICCAI BraTS 2015 dataset is used for the experimentation. For the shape feature extraction region of interest of tumor part of brain tumor is used. Proposed work uses .mha file. Firstly, it converted in 2 directional format to apply image processing for the segmentation of only tumor part. Shape features like center of gravity, circularity ratio, rectangularity, convexity, and solidity are calculated by using segmented part of brain tumor. 210 images are used to classify the brain tumor as benign and malignant brain tumor. Binary classification is used for this work. Shape features are provide to support vector machine and random forest classifier for classification of brain tumor.

TABLE II. Classification Results

Algorithm	**Accuracy**	**Sensitivity**	**Specificity**
Support vector machine	81.42%	80.15%	83.33%
Random forest	86.66%	84.44%	85.86%

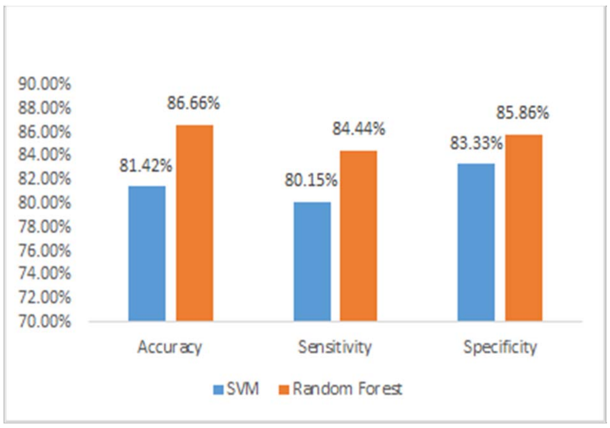

Fig. 3. Results of comparison of classifiers

VI. CONCLUSION

Computer aided diagnosis system results the diagnosis of disease fast as compared to traditional way of diagnosis. In brain tumor diagnosis is at early stage must result the quality of life. The aim of study is to analyze the effectiveness of shape feature in classification of brain tumor as benign and malignant tumor. Shape feature plays a very important role in brain tumor classification. By using different shape feature extraction and classification with support vector machine and random forest algorithms. Proposed method achieved 86.66% accuracy with random forest algorithm. Applications related to this approach are like object detection, medical, classification, security etc.

REFERENCS

[1] American Brain Tumor Association Adolescent and Young Adult Primary Brain and Central Nervous System Tumors Diagnosed in the United States published as a supplement to the Society for Neuro-Oncology official journal, Neuro-Oncology (www.abta.org).

[2] Devos, A. and Lukas, L. (2014) Does the Combination of Magnetic Resonance Imaging and Spectroscopic Imaging Improve the Classification of Brain Tumors IEMBS ' 26th Annual International

[3] Yang Mingqiang, Kpalma Kidiyo and Ronsin Joseph, A Survey of Shape Feature Extraction Techniques Pattern Recognition Techniques, ISBN 978-953- 7619-24-4, pp. 626, November 2008, I-Tech, Vienna, Austria

[4] Devi, Nilakshi, and Kaustubh Bhattacharyya. "Automatic Brain Tumor Detection and Classification of Grades of Astrocytoma." Proceedings of the International Conference on Computing and Communication Systems. Springer, Singapore, 2018.

[5] Ismael, Mustafa R., and Ikhlas Abdel-Qader. "Brain Tumor Classification via Statistical Features and Back-Propagation Neural Network." 2018 IEEE International Conference on Electro/Information Technology (EIT). IEEE, 2018Patel, Mitisha Narottambhai, and Purvi Tandel. "A Survey on Feature Extraction Techniques for Shape based Object Recognition." Image 137.6 (2016).

[6] Zhang, Lichi, et al. "Malignant Brain Tumor Classification Using the Random Forest Method." Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR). Springer, Cham, 2018.

[7] Sehgal, Aastha, et al. "Automatic brain tumor segmentation and extraction in MR images." Advances in Signal Processing (CASP), Conference on. IEEE, 2016.

[8] Sunit Sivaraj, Dr. B. Surendiran, Hariharan. M et al. " MRI Brain Tumour Image Retrieval Using Low Level Features and High Level Semantics"IEEE (2015

[9] Gupta, Sheifali, et al. "Segmentation, Feature Extraction and Classification of Astrocytoma in MR Images." Indian Journal of Science and Technology 9.36 (2016).

[10] Rezaei, Kimia, and Hamed Agahi. "SEGMENTATION AND CLASSIFICATION OF BRAIN TUMOR CT IMAGES USING SVM WITH WEIGHTED KERNEL WIDTH." Computer Science & Information Technology 39.

[11] Joshi, Anjali, V. Charan, and Shanthi Prince. "A novel methodology for brain tumor detection based on two stage segmentation of MRI images." Advanced Computing and Communication Systems, 2015 International Conference on. IEEE, 2015.

[12] https://www.smir.ch/BRATS/Start2015

[13] Conference of the IEEE Engineering in Medicine and Biology Society , 1-5 September 2004, 407 410.

[14] Nedjar, Imane, et al. "Random Forest Based Classification Of Medical X-ray Images Using a Genetic Algorithm For Feature Selection." Journal of Mechanics in Medicine and Biology 15.02 (2015): 1540025

[15] Sled JG, Zijdenbos AP and Evans AC, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," IEEE

[16] Trans. Med. Imag., vol. 17, no. 1,pp. 87-97, Feb. 1998.

[17] Tustison N., Gee J., "N4ITK: Nick's N3 ITK Implementation For MRI Bias Field Correction," Published in The Insight Journal - 2009 January-June.

[18] Perona P and Malik J, "Scale-space and edge detection using Anisotropic Diffusion," IEEE Trans. On Pattern Analysis and Machine Intelligence, vol. 12, no. 7, July 1990.

[19] Nandpuru, Hari Babu, S. S. Salankar, and V. R. Bora. "MRI brain cancer classification using support vector machine." 2014 IEEE Students' Conference on Electrical, Electronics and Computer Science. IEEE, 2014.

[20] Rathi, V. P., and S. Palani. "Brain tumor MRI image classification with feature selection and extraction using linear discriminant analysis." arXiv preprint arXiv:1208.2128 (2012).