Brain Tumor Segmentation Based On a Various Classification Algorithm

A.Udhaya Kunam
Research Scholar, PG & Research Department of Computer Science, Raja Dooraisingam Govt. Arts College, Sivagangai, TamilNadu, India.

Dr.N.Sujatha
Assistant Professor, PG & Research Department of Computer Science, Raja Dooraisingam Govt. Arts College, Sivagangai, TamilNadu, India.

Abstract – Brain tumor is a mass of tissue due to the abnormal replication of cells in the brain. It is a mass of tissues which results in hormonal changes results in mortality. The prediction and prevention of brain tumor is a complex one in order to conserve human life. Now-a-days identifying of brain tumor is more flexible by means of advanced medical image processing methodologies. In order to do this, the well-known technique for brain imaging is Magnetic Resonance Imaging (MRI) due to the absence of ionizing radiations. The major drawbacks in existing brain image processing techniques were more false positive rates with low accuracy. In this paper we intended to identify the best brain tumor detection system by comparing the performance of well known classifiers. For that, we approach a perfect brain tumor detection system by overcoming the existing shortcomings which are occur during the traditional methodologies. In the proposed method the preprocessed section is performed by median filtering and segmentation by means of morphological technique. The implementation of Genetic Algorithm (GA) based feature optimization is applied to our proposed system, for achieving higher accuracy by selecting best subset from the input images. These selected features were implied to three well known classifiers such as Decision Tree J48, k-Nearest Neighbor (KNN) and Multi-Layer Perceptron (MLP). The performance of all these classifiers were discussed in the paper and the accuracy of proposed brain tumor detection system are achieved by specificity, sensitivity, accuracy and error rates.

Index Terms – Brain tumor, Genetic Algorithm (GA), Decision Tree J48, Multi-Layer Perceptron (MLP) and k-Nearest Neighbor (KNN).

1. INTRODUCTION

Brain tumor is a crucial type of cancer that is caused by the growth of suspicious tissues due the excessive multiplication of cells. The brain tumor is common in both children and adults. The tumors cause severe pressure on the brain and spreads throughout the entire region of the brain. The brain tumor is majorly classified as primary and metastatic brain tumor. The tumor caused in the brain is primary brain tumor, whereas the tumor in the brain resulted due to the spread of cancer from other body parts is termed as metastatic cancer. The metastatic tumors are common that the primary brain tumors. The primary tumors are further classified as benign and malignant tumors, in which the former is the early stage and the latter leads to mortality. The major causes of the brain tumors are not yet surely identified. But, they are caused due to the exposure of radiations, while medical imaging. The presence of brain tumors are confirmed by the following symptoms: frequent headaches, loss of motor skills, vision changes, seizures, cognitive disorders, and numbness and speech problems. The brain tumors may also result in hormonal disorders and personality changes.

The brain tumors are diagnosed by the medical imaging techniques to detect the tumor in an efficient way. The medical images can be obtained using several imaging techniques such as ultrasound, Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). Among these techniques MRI techniques is used for brain image acquisition, as the other techniques exposes the brain to emitting radiation that leads to cancer. In Magnetic Resonance Imaging (MRI), the three dimensional images of the hidden organ are obtained using a strong field of magnetism and radio waves. The major significance of MRI technique is that there is no risk of ionized radiations. The absence of ionizing radiation is the major advantage of MRI scans. The diagnostic ability of the MR images are enhanced by the application of image processing techniques. Brain image processing is a crucial task in medical image processing. Medical image processing makes medical diagnosis easier by disclosing the internal structures of the unseen organs in the human body. The medical image processing reduces the burden of both the patients and the doctors. It is one of the major applications of digital image processing, in which mathematical operations are applied on an image to enhance its quality.

Further improvements on classification are done by enabling feature optimization step. By doing these it minimize the dimensionality of the feature set and also consumes minimal time with low computational complexity. The feature selection consists of three stages such as screening, ranking, and selection. In the screening stage the irrelevant features were

removed and then other remaining features were sorted during the ranking stage. During the selection stage feature optimization algorithm is processed by means its, obtains the feature subsets. At every feature set selection it generates subsets for the goodness, further it is used as the classification inputs.

An effective image classification technique is obtained by feeding the set of features into them. Generally the brain tumors were classified based on its features such as size, location and severity. During classification the most common challenges need to overcome are;

- Minimal accuracy
- High computational complexity
- It's hard to classify the pixels near the boundaries

To overcome the above mentioned facts in this paper we applies GA based feature optimization, according to this it process into three different classifier such as Decision Tree J48, MLP and kNN. To find the best among these, the performance of these classifiers were considered as per its features. During the proposed work the preprocessing is done by median filtering and morphological technique is used for segmentation.

In this paper the remaining sections were structured as follows: The traditional techniques for feature selection and classification of brain image processing were reviewed at section II and Section III are about proposed technique in brain tumor detection system. Section IV deals in the comparison of results between J48, kNN and MLP classification techniques. Finally Section V describes about the conclusion of the proposed work.

2. RELATED WORK

In this section we discussed various existing brain tumor detection techniques in the field. *Gordilo, et al.* [1] explains in his works about the traditional techniques in brain images which is mainly for tumor segmentation. During segmentation it classifies the white matter, grey matter, and the cerebrospinal fluid of brain. His works composed with the semiautomatic and full automatic approach on segmentation. He also showed the application of Magnetic Resonance Imaging (MRI) in medical imaging. There are four segmentation techniques included such as region-based, pixel-based, model-based and thresholding-based. According to Adegoke, et al [2] he reviewed the fundamentals of image processing and discussed about feature extraction techniques in detail. Extraction is nothing but from the selected images relevant features were extracted which is so helpful to study among the hidden features of the images. *Sharma, et al.* [3] included mathematical model in his work for morphological reconstruction features which are extracted from the brain MR images. Initially by means of global thresholding technique the raw MR images were preprocessed. Next the entire brain image is segmented from the tumors; it can be carried out by the suggested approach. Additionally to detect the tumors, non-uniform intensity regions were also altered. The morphological operators are applied for removing the pepper noise in the brain images. At *Selvakumar, et al.* [6] a hybrid clustering technique is used by combining FCM and K-means clustering for segmenting the brain tumors. According to his suggested algorithm the shape and sizes were evaluated. But he cannot able to get a exact diagnosis of tumor to achieve that he computed aided technique which reduces the computational time and indentify the location accurately.

Dahab, et al. [8] applies a novel method of neural network techniques in processing medical images. He proposes image classification by suggesting Learning Vector Quantization (LVQ) based modified Probabilistic Neural Network (PNN) model. The properties of the medical images were improved by undergoing smoothing, filtering and edge detection techniques. The accuracy is achieved by canny edge detection algorithm and its attained 100% accuracy with sensitivity compared to the other methods. *John, et al.* [9] handles texture based NN and wavelet approaches for the brain tumor classification. He progress this approach at three stages such as wavelet decomposition, texture based feature extraction and NN based classification. The image decomposition is done by Daubechies wavelet method, in addition to it the homogeneity, entropy, contrast and correlation is applied for image extraction as per the suggested techniques. The result of NN based classification achieves nearly 100% accuracy. According to *Mustaqeem, et al.* [10], he proposes a hybrid brain tumor detection algorithm by merging watershed and thresholding based segmentation techniques. In preprocessing technique the sharpening and enhancing tasks were performed in which the watershed segmentation segments the region of the images into the pixel intensity values. This proposed method was more efficient and for intensity based segmentation morphological operators is used.

Baboulaz and Dragotti [11] deals with the low resolution and quality images. His works were concentrated at the corners and step edges, which consist of low level features of the images, were identified by sub-pixel extraction method. To compute a straight continuous step edge mathematically, he uses Heaviside function technique. It is comparatively best approach gives better results compared to the other techniques. At *Sridhar, et al.* [13] the image classification is done by combining the Discrete Cosine Transform (DCT) and PNN. Here the DCT is applied for minimizing the dimensionality of the feature set by extraction. It results in rapid processing speed with reduced computational complexities. *Shao, et al.* [14] describes multi objective genetic programming technique for classifying the medical images which are automatically generated with a set of image features. The performance of this method is studied by measuring the fitness, classification error and tree complexities. The *Chu, et al.* [15] proposes feature

selection approach for improving the effectiveness of image classifications and attains accuracy compared to other traditional techniques.

3. PROPOSED METHOD

This section explains about proposed technique used in brain tumor detection system. The below fig 1 demonstrates the overall flow of the brain tumor detection system.

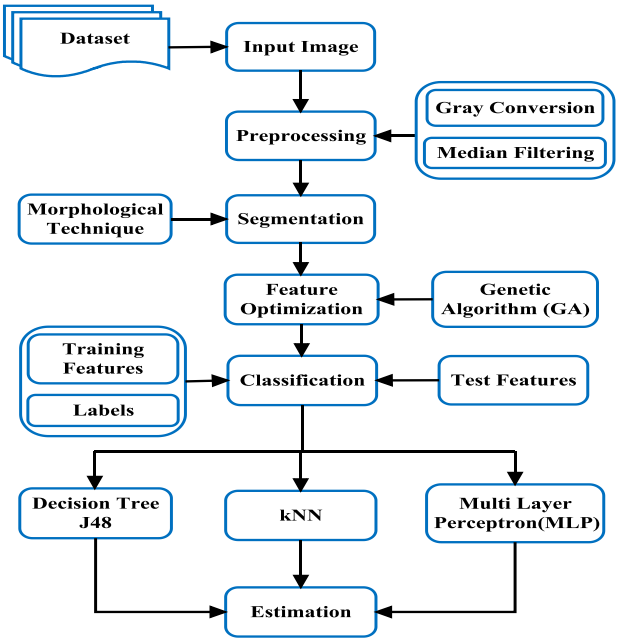

Fig 1. Overall flow of the brain tumor detection system

The steps involves in proposed brain tumor detection system is explained below:

- Preprocessing
- Image segmentation
- Feature selection
- Image classification

A. Preprocessing

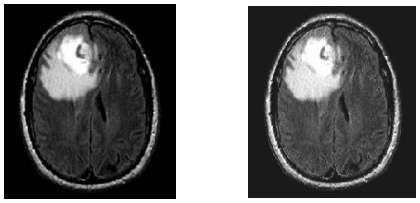

Fig 2 (a) Raw Image (b) Preprocessed image

Due to noisy, inconsistent and incomplete data, pre-processing [39] plays an important role. It is one of the preliminary steps that are required to acquire the high accuracy of steps. CT and MRI images consist of artifacts; patient specific and equipment based artifacts; others are ring, staircase and volume effect artifacts. Before analyzing all these are removed by pre-processing procedures. We have proposed different denoising approaches Which is done by applying median filtering technique [34].

B. Segmentation

It is the process of dividing an image into regions with several properties such as Color, texture, brightness, contrast and gray level [35]. The input to the process is a digital gray scale image. (e.g., CT or MRI). The output of the process is abnormalities. The use of segmentation is to give greater information than which exists in medical images. Various techniques like neural networks, decision tree, and rule based algorithm and Bayesian networks are used to get desired output data in segmentation. On that stage the tumor area is categorized as per the characteristic features such as size, location and shape. The result of segmentation is shown in Fig 3.

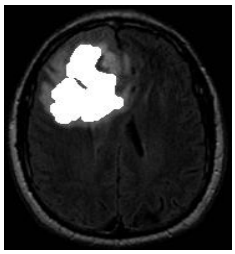

Fig 3 Segmented Images

C. Feature Optimization

Feature Optimization is the combined process of feature extraction and selection which plays a vital role in brain image processing. The initial feature selection process minimized the dimensionality of the feature set, by doing this it takes minimum time for detecting as well as computational costs. Then to extract the best set of features from the raw dataset Genetic Algorithm (GA) is used.

1) Genetic Algorithm (GA)

In image processing a best known optimization technique is GA which selects accurate feature from a large set. These features were obtained after distinguishing the good and bad features. The GA consists of the following steps:

- Initialization
- Evaluation
- Selection
- Recombination
- Mutation
- Replacement

a) Initialization

The raw dataset undergoes random selection of features and population initialization is done according to the domain specific knowledge. The huge sized population requires high computational time and also as much as accuracy is not obtained. By means of keeping population size always static cumulatively the probability must be one.

b) Evaluation

The evaluation is nothing but examining the fitness by validating the goodness of the features with the every feature in the population. The resultant features after the evaluation were grouped using fitness functions as subset, namely, offspring population.

c) Selection

The selection process is the combination of several factors such as roulette-wheel selection, ranking based selection, tournament selection and stochastic universal selection for achieving perfect fitness values. It is also known as reproduction and the probability of the fitness function involves in choosing of features.

d) Recombination

As the name states in this stage one or more features grouped by means of crossover operator. In this section any two features were selected randomly from the previous stage and those characteristics between the features were interchanged from one another. Finally the perfect characteristics from two features were gained and combined to get a best feature.

e) Mutation

The best feature gained at previous stage is operated on mutation to maintain the diversity in population. To be a better one some changes were done locally and these mutations were repeated until it reaches the satisfied conditions.

Genetic Algorithm

Step 1: Initialize the population

Step 2: Compute the fitness function

Step 3: Apply selection, crossover and mutation operations

Step 4: Select the best feature set

Step 5: Repeat steps 2, 3 and 4 until the stopping criteria is met.

D) Classification

Classification means labeling the images as per its features. Among that the best feature is indentified by applying GA. The further process is progressed by implying the GA into three classifiers such as KNN, J48 and MLP for result comparisons.

1) KNN Classification

KNN classifier is a well known classifier applicable for both classification and regression which compares the features with its neighboring feature. The input samples which are processed into KNN classification technique were the training samples with closest neighbors in order to achieve the desire result. The distance between the individual instances in the feature set with its neighbors were calculated by the pixels before computing. The main intention of applying KNN classification technique is assigning the unlabeled sample data according to its class as per its k nearest neighbors. Then by means of nearest neighbors it forms cluster and changes locally due to the high sensitivity in the feature set.

2) Decision Tree J48

As the name itself implies J48 is the best known decision tree based classification technique. Initially it classifies the images as per the attributes and forms tree structure respectively. The tree hierarchy is explained in an understandable way. The Decision Tree J48 is extended from ID3 and it is performed mainly for its simple methodology in identifying the hidden pixels in the images. Under classification the images were arranged in a leaf structure and get pruned. By labeling these pixels were grouped and on each pixel the information's were extracted then tested. From resultant pixel the perfect one is selected and these classifiers are appreciated for handling both discrete and continuous values.

3) Multi-Layer Perceptron (MLP)

The Multi-Layer Perceptron (MLP) is a feed forward ANN technique, here by mapping the classification were done on the input images. The mapping is done on the features of the training and testing dataset. Here the mapping is done by applying back propagation algorithm. By means of that the MLP constructs nodes as a directed graph and then connected to each other. Each individual node in the graph is provided with non-linear activation function. Additionally the datasets of MLP were trained by supervised learning techniques which are also helpful in classifying non-linear data's. It operates fitness function in a stochastic manner for solving the complexities.

4. PERFORMANCE ANALYSIS

In this section the works of three classification techniques such as kNN, J48 and MLP were compared by means of its features extracted using GA based feature selection. The factors involved in performance comparison were listed below;

- Accuracy
- Sensitivity
- Specificity
- Error Rate

A. Accuracy

Accuracy is calculating the ratio of number of correct assessment to the total number of assessments. In the entire dataset initially the number of relevant images were extracted and compared to entire dataset by applying the below mentioned formula in which data quality and errors were the important factors which are measure in terms of percentages (%).

$$Accuracy = \frac{(TN+TP)}{(TN+TP+FN+FP)} \quad (1)$$

Where, TN-True Negative, TP-True Positive, FP-False positive and FN-False Negative.

B. Sensitivity

Initially total number of true positive and false negative assessments is extracted from that ratio of the number of true positive assessments to the total number is calculated to get the sensitivity result. The correctly identified data's declare the degree of positive values. The calculation part of sensitivity is done by applying the below formula and it's measured in terms of percentage (%).

$$Sensitivity = \frac{TP}{(TP+FN)} \quad (2)$$

C. Specificity

The specificity helps in predicting the impact of changes in the output because of its changes in input dataset. The correctly identified negative values give the Specificity which is also measured in terms of percentage (%). The representation of the specificity formula is the ratio of the number of true negative assessments to the total number of true negative and false positive assessments.

$$Specificity = \frac{TN}{(TN+FP)} \quad (3)$$

Table 1 Comparison table for J48, kNN an MLP classification techniques based on accuracy, sensitivity and specificity

Classification Techniques	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**
J48	95.23	97.02	85
Knn	97.08	98.98	83
MLP	94.33	96.07	86.8

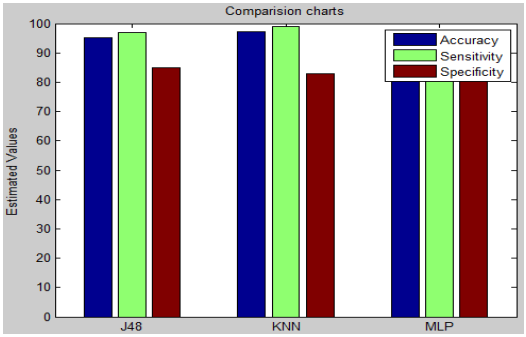

Fig 3. Comparison of J48, kNN and MLP classifiers in terms of accuracy, specificity and sensitivity

The fig 3 shows the graphical representation of comparing three classification techniques such as J48, kNN and MLP by measuring the parameters accuracy, sensitivity and specificity. According to the table 1 kNN achieves higher accuracy compared to J48 and MLP classification techniques. The accuracy value obtained by kNN is 97.08% where as by J48 and MLP are 95.23% and 94.33% respectively. Not only accuracy on sensitivity also kNN proves more effective than the other two classifiers. The gained values on sensitivity by these three classifiers are 98.98%, 97.08% and 96.07% respectively. On dealing with specificity the obtained result proves MLP classification technique is better than the kNN and J48 techniques. The MLP technique gained 86.6% specificity but kNN is of 83% and the specificity of J48 is 85%.

D. Error Rate

For a system if the error rate is in minimal level it's declared as efficient system and the error rate is calculated by defining the ratio of the number of errors to the unit time. The formula for calculating error rate is represented below;

$$Error\ Rate = \frac{Number\ of\ errors}{Unit\ time} \quad (4)$$

The table 2 clearly shows the error rates of three classifications such as kNN is 2.91, MLP is 5.66 and J48 is 4.76. These values were plotted on graphical view and shown exactly in below mentioned fig 4. Among these it is clear that compared to J48 and MLP techniques the error rate for kNN is very low that achieves overall efficiency than others.

Table 2 Error rate comparison for J48, kNN and MLP classification techniques

Classification Techniques	**Error Rate**
J48	4.76
kNN	2.91
MLP	5.66

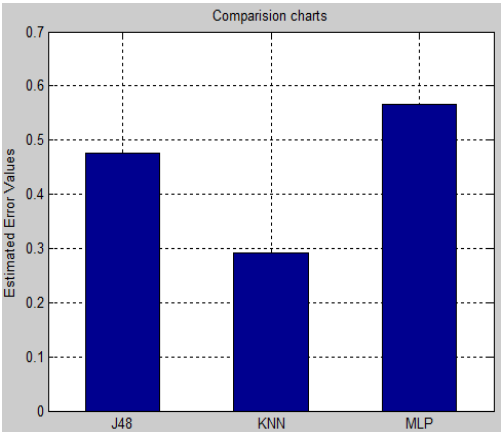

Fig 4 Comparison of error rate between J48, kNN and MLP classifiers

5. CONCLUSION

This paper proposed an enhanced novel brain tumor detection system at several stages. The initial preprocessed section applies median filtering techniques for preprocessing MR brain images. Then it undergoes segmentation by means of morphological segmentation technique were tumor affected region are segmented perfectly in which the GA based feature optimization technique is applied to reduce the overall complexity by providing best feature sets. These gained features were fed as input for three known classifiers such as kNN, J48 and MLP. Here these classifiers were intended to separate the normal and abnormal brain MR images based on some factors which are explained above. The three classifiers performance were compared accordingly in terms of accuracy, specificity and sensitivity. Among which the kNN classification technique is proved by providing best result in the factors of accuracy, specificity and sensitivity along with low error rate. Hence it is proved among all, the kNN is the most efficient classifier for brain tumor classification.

REFERENCES

[1] N. Gordillo, E. Montseny, and P. Sobrevilla, "State of the art survey on MRI brain tumor segmentation," *Magnetic resonance imaging*, vol. 31, pp. 1426-1438, 2013.

[2] B. Adegoke, M. Omotayo, A. Olarinoye, and A. Olawuni, "Medical image Feature Extraction: A Survey," 2013.

[3] Y. Sharma and Y. K. Meghrajani, "Brain tumor extraction from MRI image using mathematical morphological reconstruction," in *Emerging Technology Trends in Electronics, Communication and Networking (ET2ECN), 2014 2nd International Conference on*, 2014, pp. 1-4.

[4] A. Sivaramakrishnan and M. Karnan, "A novel based approach for extraction of brain tumor in MRI images using soft computing techniques," *Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, pp. 1845-1848, 2013.

[5] E.-S. A. El-Dahshan, H. M. Mohsen, K. Revett, and A.-B. M. Salem, "Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm," *Expert systems with Applications*, vol. 41, pp. 5526-5545, 2014.

[6] J. Selvakumar, A. Lakshmi, and T. Arivoli, "Brain tumor segmentation and its area calculation in brain MR images using K-mean clustering and Fuzzy C-mean algorithm," in *Advances in Engineering, Science and Management (ICAESM), 2012 International Conference on*, 2012, pp. 186-190.

[7] A. Jayachandran and R. Dhanasekaran, "Automatic detection of brain tumor in magnetic resonance images using multi-texton histogram and support vector machine," *International Journal of Imaging Systems and Technology*, vol. 23, pp. 97-103, 2013.

[8] D. A. Dahab, S. S. Ghoniemy, and G. M. Selim, "Automated brain tumor detection and identification using image processing and probabilistic neural network techniques," *International journal of image processing and visual communication*, vol. 1, pp. 1-8, 2012.

[9] P. John, "Brain tumor classification using wavelet and texture based neural network," *International Journal of Scientific & Engineering Research*, vol. 3, p. 1, 2012.

[10] A. Mustaqeem, A. Javed, and T. Fatima, "An efficient brain tumor detection algorithm using watershed & thresholding based segmentation," *International Journal of Image, Graphics and Signal Processing*, vol. 4, p. 34, 2012.

[11] L. Baboulaz and P. L. Dragotti, "Local feature extraction for image super-resolution," in *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, 2011, pp. V-401-V-404.

[12] Z. Camlica, H. Tizhoosh, and F. Khalvati, "Medical Image Classification via SVM using LBP Features from Saliency-Based Folded Data," *arXiv preprint arXiv:1509.04619*, 2015.

[13] D. Sridhar and I. M. Krishna, "Brain tumor classification using discrete cosine transform and probabilistic neural network," in *Signal Processing Image Processing & Pattern Recognition (ICSIPR), 2013 International Conference on*, 2013, pp. 92-96.

[14] L. Shao, L. Liu, and X. Li, "Feature learning for image classification via multiobjective genetic programming," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, pp. 1359-1371, 2014.

[15] C. Chu, A.-L. Hsu, K.-H. Chou, P. Bandettini, C. Lin, and A. s. D. N. Initiative, "Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images," *Neuroimage*, vol. 60, pp. 59-70, 2012.