Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2019) 27: 1704 – 1718
© TÜBİTAK
doi:10.3906/elk-1804-13

Optimized bilevel classifier for brain tumor type and grade discrimination using evolutionary fuzzy computing

Kavitha SRINIVASAN[1], **Mohanavalli SUBRAMANIAM**[2,*], **Bharathi BHAGAVATHSINGH**[1]

[1]Department of Computer Science and Engineering, SSN College of Engineering, Kalavakkam, India

[2]Department of Information Technology SSN College of Engineering, Kalavakkam, India

Received: 03.04.2018 • **Accepted/Published Online:** 16.11.2018 • **Final Version:** 15.05.2019

Abstract: In this paper, an optimized bilevel brain tumor diagnostic system for identifying the tumor type at the first level and grade of the identified tumor at the second level is proposed using genetic algorithm, decision tree, and fuzzy rule-based approach. The dataset is composed of axial MRI of brain tumor types and grades. From the images, various features such as first and second order statistical and textural features are extracted (26 features). In the first level, tumor type classification was done using decision tree constructed with all features. Further evolutionary computing using genetic algorithms (GA) was applied to select the optimal discriminating feature set (5 features) and classification using the decision tree constructed with the reduced feature set resulted in better performance. In the second level, grade classification, a fuzzy rule-based approach was used to resolve the uncertainty in discriminating the tumor grades II and III. Membership functions of all grades were defined for all features extracted from brain tumor grade images, to derive the fuzzy inference rules for grade discrimination. Similar to type classification with GA, better grade discrimination performance was exhibited with fuzzy inference rules derived using optimal feature set (13 features) using GA. Overall performance comparison of the proposed bilevel classifier with all features vs GA-based feature selection, shows that evolutionary computing combined with fuzzy rule -based approach is successful in reducing false positives, thereby enhancing classifier performance.

Key words: Fuzzy rule-based approach, brain tumor, decision tree, optimal feature set, genetic algorithm, magnetic resonance images

1. Introduction

Medical imaging technology has revolutionized health care over the decades, enabling physicians to diagnose disease earlier and improve patient health. Medical images can be acquired through various modalities such as magnetic resonance imaging (MRI), computed tomography (CT), single photon emission computed tomography, ultrasonography, and positron emission tomography. MRI-based medical image analysis is popularly used in brain image analysis in recent days as it is more suitable for efficient and objective evaluation of large data. Brain tumors are abnormal and uncontrolled proliferations of cells and it is believed to be the most lethal disease [1]. Hence, timely diagnosis of this disease is important for proper treatment for the patients; it crucially determines their lifetime.

*Correspondence: mohanas@ssn.edu.in

1.1. Tumor classification in brain MRI images

Extensive research is progressing in the area of brain image analysis and diagnosis. Computer aided disease diagnosis (CADD) system for brain related images are being developed in two aspects (i) bilevel classification of brain images as normal or abnormal using linear classifier without considering the disease type, (ii) multilevel classification of images into specific disease types or normal. The algorithms used for classifying brain images are using support vector machine (SVM), decision tree (DT), neural network, naive Bayes, and linear discriminant analysis (LDA) [2–6].

In the survey article [1] on MRI-based image analysis of brain tumor images, the authors have given a comprehensive overview of brain tumors. The state of the art techniques used in segmentation, registration, and modeling techniques to perform tumor image analysis have been detailed with a focus on gliomas. Also a critical assessment of the current state and limitations in clinical application is given along with future developments in radiological tumor assessment.

Ahmed Kharrat et al. [7] have proposed a hybrid approach to classify brain MRI into normal, benign or malignant tumor. They have used wavelet based feature extraction technique, and extracted 44 features and reduced it to 5 features using genetic algorithm. A SVM classifier has been trained employing RBF kernel and 5-fold cross-validation to classify the brain images. Though the performance reported is high, the drawback of the hybrid approach is that the SVM has to be trained afresh for every new image in the database.

Naik and Patel [8], in their research, proposed a system to classify tumor type of the CT scan brain images. They used gray level cooccurrence matrix (GLCM) of the preprocessed CT images to extract the features. Decision tree and naive Bayesian classifiers were used to classify the images and their performances also were compared. Geethu and Monica [9], in their research, made an elaborate review on segmentation, and classification methods for brain tumor MR images are discussed. The survey revealed state-of-the-art high-accuracy classification techniques suitable to classify the tumor type and discussed the scope for design of better image processing and machine learning approaches to grade the type of tumor specifically for gliomas: astrocytomas. Firat et al. [10] assessed the contribution of multiparametric MRI features from multiregion of interest, only for grading of gliomas. Their research shows a promising scope for applying machine learning algorithms for classifying type and grade of brain tumors.

A study of the related works in the literature shows that classifying the specific tumor type with grade or normal using multilevel classifier has scope for extensive research.

1.2. Evolutionary computing for feature selection

Evolutionary computing is based on the central concept of ”natural selection” leading to better and better solutions. Genetic algorithm (GA) is a programming technique that applies biological evolution as a problem-solving strategy. Researchers have applied GA to optimize the feature selection process, to enhance the performance of both linear and multiclass classification [5]. Sharma et al. [11], in their research, used an adaptive fuzzy and neural network to segment astrocytoma, a type of brain tumor grade I to IV. Feature extraction was done using gray level coccurrence matrix of the brain MRI. Genetic algorithm was applied to select relevant features to derive the fuzzy rules for tumor segmentation. Kavitha et al. [12], in their research, proposed a classification system for tumor types using SVM and DT with GA. Experiments carried out using two classifiers (with and without GA) for brain tumor classification show that DT with GA gives maximum accuracy compared to SVM with GA for optimal feature set.

1.3. Fuzzy approaches for tumor classification and discrimination

Classification algorithms are generally not suitable for grade discrimination due to uncertainty and vagueness in the feature set. To address this issue, various research works have been proposed using understanding-based techniques namely ontology-based approach, case-based reasoning, graph grammar, and fuzzy-based reasoning [13].

Samuel et al. [14] developed a web-based fuzzy inference system for diagnosis of typhoid fever. Chen Nian-yi et al. [15] proposed a fuzzy rule-extraction algorithm based on fuzzy min-max neural network to perform grade classification of glioma tumor. Nedeljkovi [16] made a detailed description about image classification using fuzzy logic. A prior knowledge of spectral information about a land area was used to build a maximum likelihood (ML) classifier and a fuzzy inference system.

Zarandi et al. [17] proposed a type-II fuzzy image processing expert system to diagnose brain tumors, especially astrocytoma tumors in T1-weighted MR images with contrast. In their proposed system, they applied an image filter to preprocess the MR images, segmentation using PCM fuzzy clustering method. El-Melegy et al. [18] proposed a prior information-guided fuzzy Cmeans algorithm to perform automated segmentation of tumor region in brain MRI images. A fuzzy clustering technique has been used in a combination of region and contour-based methods to detect and segment tumors in 3D brain images in the research work proposed by [19].

From the existing research works, it is understood that fuzzy approach has wider scope for improvising tumor grade discrimination, with extraction of appropriate features. In the present research study, an optimized bilevel classification system is proposed with the following objectives: reducing the false positive rate during tumor type and grade classification, identification of optimal feature set for attaining maximum classification accuracy using decision tree and fuzzy rule-based approach with GA and reducing the uncertainty between the grades II and III of astrocytoma brain tumor.

2. Proposed methodology

In this work, a bilevel classification of brain tumor for identifying its type at the first level and grade at the second level is proposed. The dataset contains brain tumor type and grade images. The dataset is partitioned into training and test sets for both type and grade images. Features are extracted from the images of training dataset. The large feature set is optimized by identifying the best discriminating features using genetic algorithm. Decision tree is constructed using both the entire feature set (DT) and optimal feature set (DT with GA). The rule set obtained from both are tested with the features extracted from test dataset. In type classification, which constitutes the level 1 of the proposed work, performance of the rule set resulting from DT and (DT with GA) DTGA were compared. Level 2 involves construction of fuzzy system for grade classification. The fuzzy inference rules of the system are constructed using entire feature set and optimal feature set obtained from the grade training images and their performances were compared. Finally, the two levels (tumor type and grade) are integrated by choosing the best techniques with respect to each, i.e the best among DT or DTGA for type classification and fuzzy or fuzzy with GA for grade classification. The overall design of the system is shown in Figure 1.

2.1. Dataset

Axial MRI brain tumor type and grade images were collected from sources such as Harvard Medical School (http://www.med.harvard.edu/aanlib/home.html) and Radiopedia (http://radiopaedia.org/articles/normal-brain-

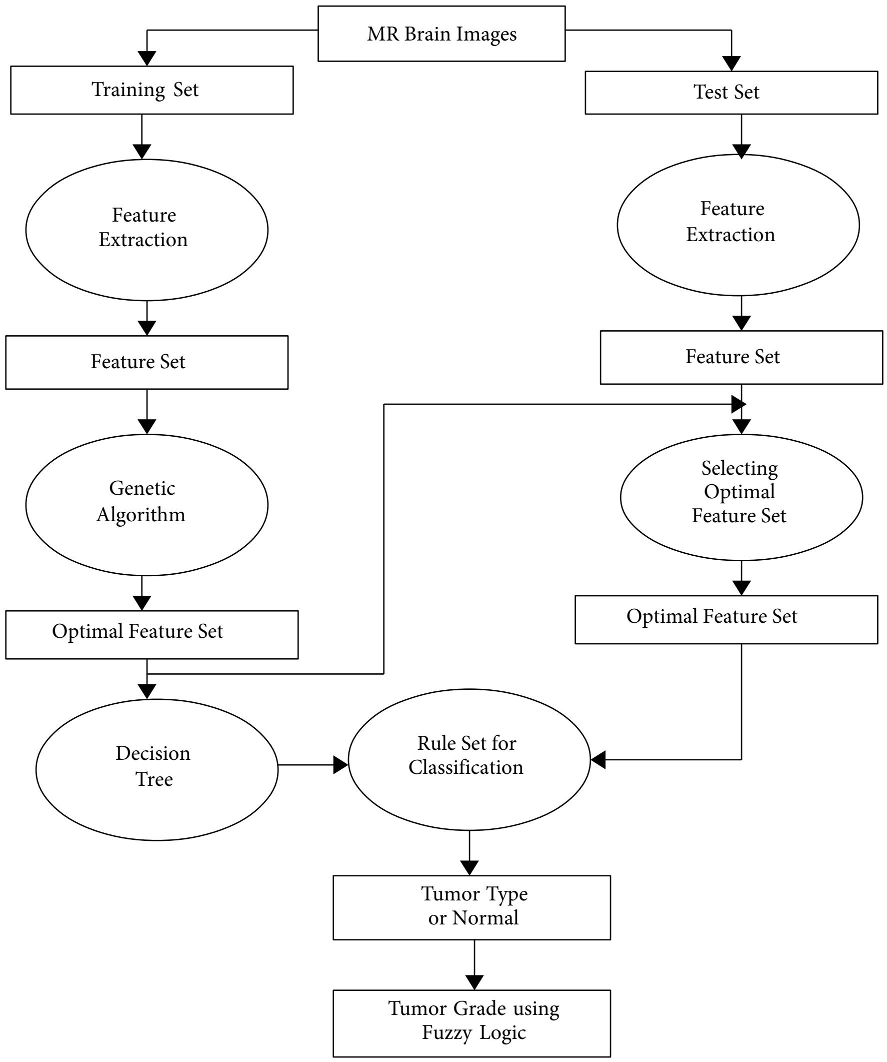


Figure 1. Proposed bilevel diagnostic system design.

imaging-examples-1). The dataset for brain tumor type contains 239 images in which 70% of total images (171) are taken for training and the remaining 30% (68) are taken for testing and the split up of the images among different types is given in Table 1. The dataset for brain tumor grade contains 80 images, 70% of which (56) are taken for training and the remaining 30% (24) are taken for testing and the split up of the images among different grades is given in Table 2.

2.2. Feature extraction

Statistical features, namely first-order features (6), second-order features (4), and textural features (16) are extracted from the images [7]. The first-order features are calculated using the histogram and the second-order features are extracted from the GLCM (gray level cooccurrence matrix) of input images. A total of 26 different features are extracted from the images that are listed in Table 3. From the extracted features, the optimal feature set is identified using genetic algorithm.

Table 1. Dataset statistics for brain tumor type.

Brain tumor type	No. of training images	No. of testing images	Total images
Type1	56	24	80
Type2	48	16	64
Type3	26	11	37
Type4	41	17	58
Total	171	68	239

Table 2. Dataset statistics for brain tumor grade.

Brain tumor type	No. of training images	No. of testing images	Total images
Grade1	14	6	20
Grade2	14	6	20
Grade3	14	6	20
Grade4	14	6	20
Total	56	24	80

Table 3. Extracted features.

Feature Type	Feature Name
First-order statistical	Mean, variance, standard deviation, skewness, kurtosis, entropy
Second-order statistical	Correlation, contrast, homogeneity, energy
Textural	Sum of energy, auto correlation, cluster prominence, cluster shade, dissimilarity, max probability, sum average, sum variance, sum entropy, difference variance, difference entropy, information measure of correlation1, information measure of correlation2, inverse difference, inverse difference normalized, inverse difference moment normalized

2.3. Genetic algorithm–optimal feature set

Genetic algorithm (GA), a larger branch of evolutionary computing, is a search-based optimization technique based on the principles of genetics and natural selection. In GA, we have initial populations which are subject to recombination and mutation operations, producing new solutions. The fitness value of each candidate solution decides its participation in the generation of better solutions. In our proposed system, to classify the type and grade of the tumor, there is a need to select best discriminating features and we used GA to evolve the optimal feature set.

The sequence of steps of genetic algorithm is given below:

- The initial population is generated with M = 20 individuals.
- Each individual is a vector of F = 26 bits representing the extracted features of the brain images taking the values as either 0 or 1 (0 - rejected feature, 1 - selected feature).
- The fitness value of an individual is the accuracy obtained from the classifier constructed for the dataset using the features represented by that individual.

- Similarly the fitness value is calculated for each individuals and the best 10 individuals are selected for the next iteration (generation).
- From these individuals two parent chromosomes (individuals) are selected using tournament selection.
- One point crossover is done to produce two new individuals based on crossover probability as 0.3.
- The generated new individuals are applied to bit string mutation based on mutation probability as 0.5.
- The evolutionary process is repeated until a termination condition is reached (global optimal solution or maximum iteration of 100). At the end of GA, optimal feature set with maximum accuracy is identified. The process flow is shown in Figure 2 [20].

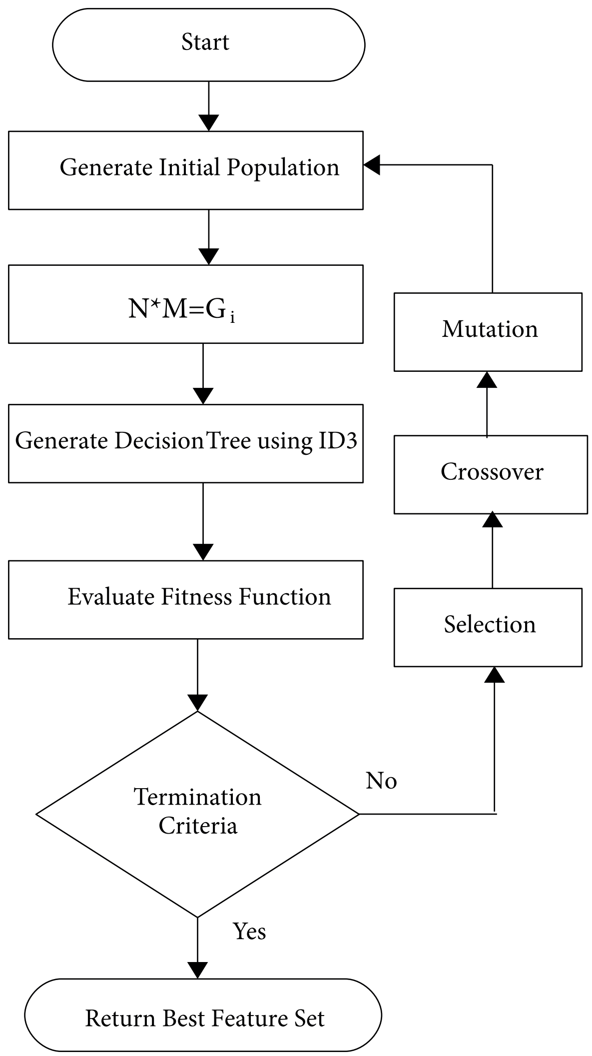


Figure 2. Process flow of GA.

2.4. Decision tree classifier

The decision tree is constructed for the extracted features of brain tumor images and optimal feature set identified using the genetic algorithm as input data for all possible splits. The best split attribute which has

Algorithm 1 GeneticAlgorithm identifies the optimal feature set from the entire feature set.

Input: Features_set, with 26 features extracted from the training images
Output: OFS Optimal Feature Set

```
function GeneticAlgorithm(Features)
    Initialise the initial population containing M individuals each of F bits taking values 0 or 1   ▷ (0-not considered, 1- considered)
    for all i in M individual do
        fitness_value(i) = fitness(individual(i))
    end for
    while Termination criteria(global accuracy or maximum iteration) not achieved do
        Select the best N individuals
        Select the parents based on selection method
        Perform cross over based on crossover-probability
        Perform mutation based on mutation-probability
        for all i in N + new children individual do
            fitness_value(i) = fitness(individual(i))
        end for
    end while
end function
procedure Fitness(Features)   ▷ Computes the fitness of each individual
    Extract the features represented by the individual
    Call the classifier to construct the model file or rule set for the considered features
    Calculate the accuracy and return
    return OFS
end procedure
```

the highest information gain is chosen to partition the data in two subsets and the partition occurs recursively. The algorithm stops when any one of the following condition is achieved: all the data of one partition belongs to the same type, there are no remaining features on which the data may be further partitioned and there are no data to be partitioned [21].

In decision tree construction, K-fold cross-validation is applied for estimating the performance of a classifier with K = 10 and m=239 (number of training images). The process for single run of 10-fold cross-validation is given below:

1. Arrange the training examples in a random order.
2. Divide the training examples into 10 folds each with the size of 239/10.
3. for i = 1, . . . , K
    (a) Train the classifier using all the examples that do not belong to Fold i.
    (b) Test the classifier on all the examples in Fold i.
    (c) Compute ni, the number of examples in Fold i that were correctly classified.
4. Compute the accuracy A_i as follows: $A_i = n_i/(m/K)$

Algorithm 2 Decision tree algorithm

Input 1: Data partition D, which consists of training tuples and their associated class labels
Input 2: Attribute_List, the set of candidate attributes.
Output: Decision tree

function DecisionTree(D, Attribute_List)
  Create a node N
  If tuples in D are all of the same class C then return N as a leaf node labelled with the class C
  If *Attribute_List* is empty then return N as a leaf node labelled with the majority class in D ▷ majority voting
  Apply $attribute\_selection\_method(D, Attribute\_list)$ [21] to find the best split point
  Label node N with split attribute
  If splitting attribute is discrete-valued and multiway splits allowed then $Attribute\_List < -Attribute\_List - splitting\_attribute$
  for all outcome j of split point **do** ▷ partition the tuples and grow subtrees for each partition
    Let D_j be the set of data tuples in D satisfying outcome j
    If D_j is empty then attach a leaf labelled with the majority class in D to node N
    Else attach the node returned by $Generate\_decision\_tree(D_j, Attribute\_List)$ to node N
  end for
  return N
end function

To obtain an accurate estimate of the classifier, the maximum accuracy obtained is selected from 10 runs. Let $(A_1, ..., A_t)$ be the accuracy estimates obtained in 10 runs from which the accuracy is calculated as $A = max(A_1, A_t)$.

2.5. Fuzzy rule-based approach

The fuzzy rule-based approach is used in identifying the grade of a specific tumor type. Decision tree algorithm exhibits good accuracy for classifying the type of the brain tumor, but does not perform well in grade discrimination. Particularly, in discrimination of grade 2 and grade 3, for which the tumor texture is similar, fuzzy approaches may be suitable to resolve the uncertainity.

In type1 (astrocytoma), 26 features are extracted from each image, which are are divided into 4 input-value set Ii :j based on their grade classes where $1 <= i <= 4$ and $1 <= j <= 26$. In addition, from the extracted features optimal feature set is identified using GA, for which the fuzzy system is built with $1 <= i <= 4$ and $1 <= j <= 13$.

The features in each input-value sets are sorted. The triangular membership function $A_{i,j}$ which has the triplets $b_{i,j}$ (center), $c_{i,j}$ (left) and $a_{i,j}$ (right). Triplets of the membership function can be calculated as given in Eq. (1) to (3).

$$b_{i,j} = \frac{x_{i,\min} + x_{i,\max}}{2}, \tag{1}$$

$$a_{i,j} = b_{i,j} - \frac{b_{i,j} - x_{i,\min}}{1 - \alpha}, \tag{2}$$

$$c_{i,j} = b_{i,j} + \frac{x_{i,\max} - b_{i,j}}{1 - \alpha}, \tag{3}$$

where $x_{i,\min}$ and $x_{i,\max}$ are the minimum and maximum elements of the input fuzzy set. The fuzzy rules are generated based on the hierarchical relationship between the class and the corresponding input fuzzy set $A_{i,j}$ and further simplified by computing the equality between two MF's as given in Eq. (4).

$$E(A_{i1,j1}, A_{i2,j1}) = \frac{\|A_{i1,j1} \cap A_{i2,j1}\|}{\|A_{i1,j1} \cup A_{i2,j1}\|} \tag{4}$$

If the equality E value is above the cut-off (a = 3), then new membership function is defined by merging the two membership function as given in Eqs. (5) to (7). These membership functions are used to generate fuzzy inference rules to build the Mamdani fuzzy model for grade discrimination.

$$a_{new,j} = \frac{a_{i1,j} + a_{i2,j}}{2}, \tag{5}$$

$$b_{new,j} = \frac{b_{i1,j} + b_{i2,j}}{2}, \tag{6}$$

$$c_{new,j} = \frac{c_{i1,j} + c_{i2,j}}{2}. \tag{7}$$

Algorithm 3 Fuzzy rul-based approach to generate and simplify the fuzzy rule set

Inputs: GradeFIS (Name), mamdani (Type), min(andMethod), max(orMethod), centroid(defuzzMethod), min(impMethod), max(aggMethod)

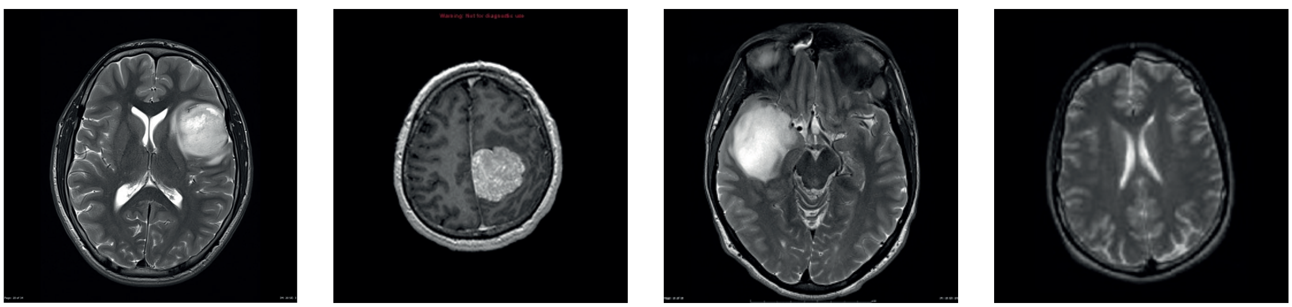
Output: FS Fuzzy rule set $[1X4$ struct$]$

function Fuzzy Rule Generation( )
    Define the input membership function
    Divide the dataset into input value set based on the grade class
    for all input value set **do**
        for all feature **do**
            Determine the triangular input membership function
        end for
    end for
    for all grade class **do**
        Define the rule list
    end for ▷ Simplifying the rule list based on the equality
    ▷ Similarity between membership function
    for all feature j **do**
        for all input membership function of the grade classes **do**
            Calculate the equality between the input membership function
            If $(equality >= cutoff)$ then
            Define new membership function which is the average of the input membership function for the two grade classes
            Modify the rule list
        end for
    end for
    return FS
end function

3. Experiments and results

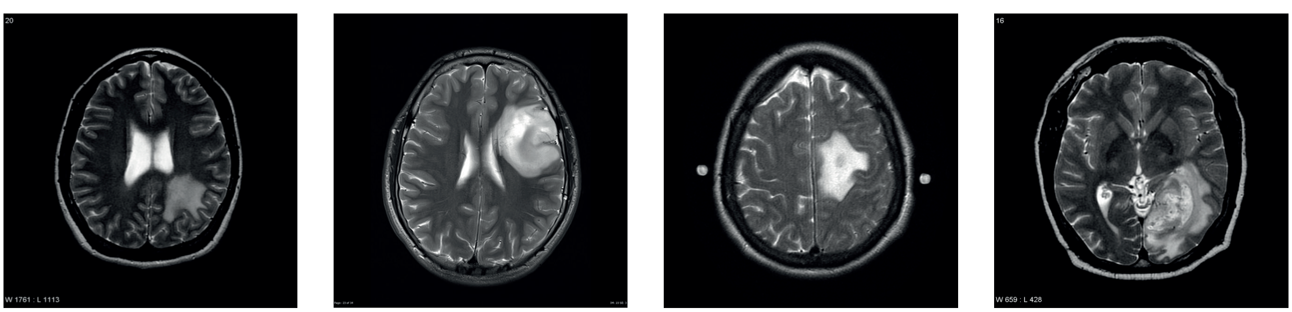
This research work attempts to improve the classification performance by selecting the suitable discriminating features for type and grade classification of tumors. The implementation is done using R and MATLAB 2015a version in windows environment. For decision tree construction, the extracted 26 features are passed to RPART and the accuracy is calculated using GA-DT fitness evaluation function for each random subset. From the accuracy estimation, a feature subset with 5 features are selected as optimal set for first level type classification. For fuzzy rule construction, each input and output variable (feature) a triangular membership function (MF) is created with its specific range. From the created MFs, fuzzy rules are generated for each grade and evaluated. From the accuracy estimation, a feature subset with 13 features are selected as optimal set for second-level grade classification. The proposed fuzzy model is developed using MATLAB fuzzy GUI and coding.

The different level of experiments carried out are: (i) Brain tumor type classification using decision tree, (ii) Brain tumor type classification using decision tree with genetic algorithm, (iii) Brain tumor grade classification using fuzzy, (iv) Brain tumor grade classification using Fuzzy with genetic algorithm, and (v) Bilevel system for brain tumor type and grade classification. The different types of tumor images and the grades of astrocytoma brain tumor are shown in Figures 3 and 4.



"Glioma-Type1" "Meningioma-Type2" "Metastasis-Type3" "Normal-Type4"

Figure 3. Brain tumor types



"Grade1" "Grade2" "Grade3" "Grade4"

Figure 4. Grades of Astrocytoma Brain tumor

3.1. Performance analysis

The classifier performances for different approaches are compared for the measures using accuracy, sensitivity, and specificity. Accuracy is estimated as the ratio of number of correctly classified images to the total images. Sensitivity and specificity are calculated using the confusion matrix.

3.2. Result of brain tumor type classification using decision tree

From the extracted 26 features, decision tree is constructed and validated using test image set features until the maximum accuracy is obtained. Then the rule set is derived for the decision tree. The confusion matrix of decision tree is given in Table 4. Optimal features (5 features—Mean, homogeneity, max probability, sum variance, and sum entropy) are selected using genetic algorithm, from which decision tree is constructed and its corresponding confusion matrix is given in Table 4.

Table 4. Confusion matrix of type classification.

Confusion	Type1		Type2		Type3		Type4	
matrix	DT	DT with GA	DT	DT with GA	DT	DT with GA	DT	DT with GA
Type1	24	23	0	1	0	0	0	0
Type2	0	0	14	15	0	1	2	0
Type3	0	0	1	0	10	11	0	0
Type4	3	1	2	4	1	0	11	12

From the confusion matrix, the specificity and sensitivity are calculated and tabulated in Table 5.

Table 5. First level classification performance using decision tree

Tumor types &	Decision tree		Decision tree with GA	
Classifier measures	Specificity in %	Senstivity in %	Specificity in %	Senstivity in %
Type1	92.10	100.00	97.43	95.83
Type2	93.75	87.50	90.19	93.75
Type3	98.00	90.90	98.03	100.00
Type4	96.00	64.70	100.00	70.58
Average	94.96	85.77	96.41	90.04

The average specificity, average sensitivity, and accuracy of classification with and without GA for brain tumor type classification is given in Table 6.

Table 6. First-level classification–performance analysis.

Classifier/measures	Specificity in %	Senstivity in %	Accuracy in %
Decision tree	94.96	85.77	86.76
Decision tree with GA	96.41	90.04	89.70

3.3. Result of grade classification using fuzzy rule-based approach

The confusion matrix of fuzzy rule-based approach for brain tumor grade classification using the entire set of 26 features is given in Table 7. The confusion matrix of fuzzy rule-based approach with GA, for brain tumor grade classification for the optimal feature set of 13 features (mean, standard deviation, skewness, entropy, contrast, energy, sum of square, cluster prominence, dissimilarity, max probability, sum variance, difference variance, inverse difference moment normalized) are given in Table 7.

Table 7. Confusion matrix of grade classification.

Confusion	Type1		Type2		Type3		Type4	
Matrix	Fuzzy Rules (FR)	FR with GA	Fuzzy Rules (FR)	FR with GA	Fuzzy Rules (FR)	FR with GA	Fuzzy Rules (FR)	FR with GA
Grade1	4	4	2	2	0	0	0	0
Grade2	0	0	4	4	2	2	0	0
Grade3	0	0	0	0	6	6	0	0
Grade4	0	0	0	0	6	5	0	1

Table 8. Second-level classification performance using fuzzy approach.

Tumor types &	Fuzzy rule-based approach		Fuzzy rule-based approach with GA	
Classifier measures	Specificity in %	Senstivity in %	Specificity in %	Senstivity in %
Grade1	100	66.66	100	66.66
Grade2	83.33	66.66	84.61	66.66
Grade3	50.00	100	56.25	100
Grade4	100	0	100	16.67
Average	83.33	58.33	85.21	62.50

From the confusion matrix, the specificity and sensitivity are calculated and tabulated in Table 8.

The fuzzy inference rule set generated can be represented in three different ways. They are verbal, symbolic, and indexed. The average specificity, average sensitivity and accuracy of classification with and without GA for brain tumor grade classification is given in Table 9.

Table 9. Second-level classification–performance analysis.

Classifier/measures	Specificity in %	Senstivity in %	Accuracy in %
Fuzzy	83.33	58.33	58.33
Fuzzy with GA	85.21	62.50	62.50

3.4. Result of bilevel classification

The rule set generated from Decision Tree with GA and the rule set generated from fuzzy with GA are integrated for classifying the type and the grades of the tumor respectively as a bilevel classification system. Presently, the developed system is tested for glioma type only in the second level.

The performance measures evaluated for the overall system are shown in Table 10, for the classifiers designed with GA for feature selection and without feature selection. The sensitivity and specificity measures (true-positive and true-negative rates) shows the significant performance improvement obtained by using GA. Figure 5 illustrates the improvement in the classifier performance for both type and grade classification using GA.

Thus, experiments have been carried out with MRI of brain tumors for its type and grade classification. The type classification was tested with decision tree classifier with all features and also using only optimal features selected by applying GA. Similarly, grade classification was experimented initially with decision tree but resulted in very poor performance. Hence, fuzzy approach with all features and optimal features selected using GA was used for grade discrimination. The overall performance of the classification of tumor type and

Table 10. Overall performance comparison.

Tumor types & Classifier measures	Classification without GA			Classification with GA		
	Accuracy	Specificity	Senstivity	Accuracy	Specificity	Senstivity
First-level type classifier	86.76	85.77	94.96	89.70	90.04	96.41
Second-level grade classifier	58.33	58.33	83.33	62.50	62.50	85.21
Bilevel classifier	72.55	72.55	89.15	76.1	76.27	90.81

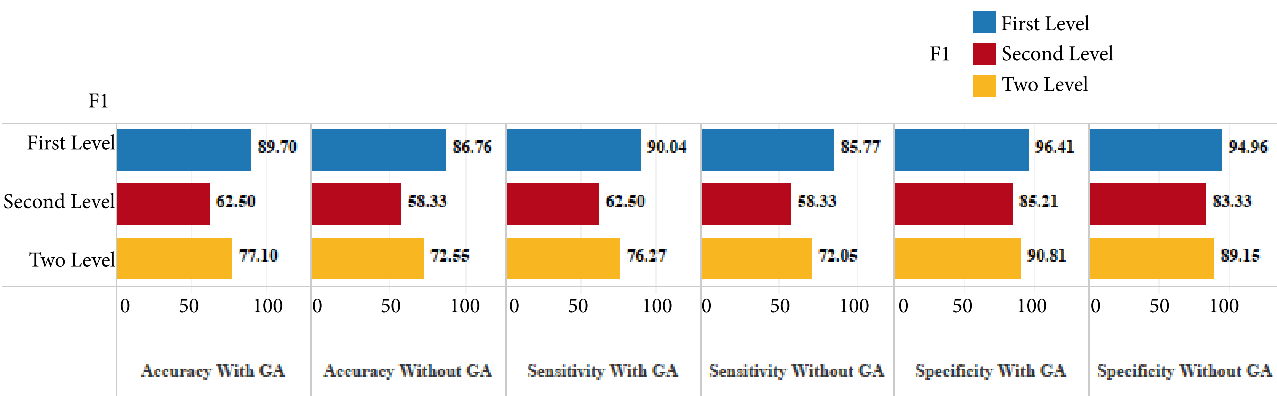


Figure 5. Performance measures for classification with all features and with GA based feature selection.

grade is observed to be better with optimal features selected using GA. Also, from the sensitivity and specificity scores attained, it is evident that the proposed bilevel classification system with GA is successful in reducing false positives and also the uncertainty in discriminating the grade of the tumor type.

4. Conclusion and discussion

In this work, we have explored the importance and advantage of combining classifiers with selection of appropriate discriminating features using an evolutionary approach (GA), for improving the performance of the classifier. We have designed a bilevel classification system to identify the type of brain tumor in the first level and grade of one particular type in the second level. The classifiers built using decision tree algorithm and fuzzy rule-based approach gave decent accuracies for type and grade classification respectively, but when combined with genetic algorithm for selection of optimal feature set, showed significant improvement in the performance measures which is very much necessary for an automation system developed for medical diagnosis.

The objectives of the proposed system were: (i) to reduce the false positive rate in tumor type and grade classification, (ii) to identify the discriminating feature set for attaining maximum classification accuracy, and (iii) to reduce the uncertainty (ambiguity) in discriminating the grades II and III of astrocytoma during classification. In the proposed bilevel classifier, decision tree was used to generate the rule set for classifying the brain tumor type using the optimal features identified by genetic algorithm. The performance of the classification using this technique resulted in a better accuracy (89.70%) than that of decision tree constructed using all extracted features (86.76%). Similarly, classification with the fuzzy inference rules generated for brain tumor grade using the optimal feature set selected using genetic algorithm resulted in an accuracy of 62.50%, which was better than that obtained using fuzzy inference rules generated with all the extracted features (58.33%). The overall accuracy of the proposed bilevel system is 72.55% with all features and 76.1% with GA-based optimal feature set selection.

The sensitivity and specificity measures of the proposed system clearly shows that fuzzy rule-based approach combined with GA performs better in grade discrimination compared to the fuzzy approach with all features. The future challenge is to further reduce the false-positives during grade classification by identifying relevant rules. The system can be developed as an expert tool by collecting real time features such as age and sex and including it along with expert's rules. It can also be extended to include other brain tumor types and grades of all the types and thus can be used as a diagnostic assistant by physicians and radiologists.

References

[1] Bauer S, Wiest R, Nolte LP, Reyes M. A survey of MRI based medical image analysis for brain tumor studies. Physics in Medicine and Biology 2013; 58 (13): R97-R129. doi: 10.1088/0031-9155/58/13/R97

[2] El-Sayed AE, Heba MM, Kenneth R, Abdel-Badeeh MS. Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm. Expert Systems with Applications 2014; 41 (11): 5526–5545. doi: 10.1016/j.eswa.2014.01.021

[3] Zacharaki EI, Wang S, Chawla S, Soo YD, Wolf R, Melhem ER, Davatzikos C. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. Magnetic Resonance in Medicine 2009; 62 (6): 1609–1618. doi: 10.1002/mrm.22147

[4] Jiawei H, Micheline K. Data Mining Concepts and Techniques. San Francisco, CA, USA: Morgan Kaufmann, 2007.

[5] Kavitha S, Thyagharajan KK. Efficient DWT-based fusion techniques using genetic algorithm for optimal parameter estimation. Soft Computing 2017; 21 (12): 3307-3316. doi: 10.1007/s00500-015-2009-6

[6] Komal S, Akwinder K, Shruti G. Brain tumor detection based on machine learning algorithms. International Journal of Computer Applications 2014; 103 (1): 7-11. doi: 10.5120/18036-6883

[7] Ahmed K, Karim G, Mohamed BM, Nacera B, Mohamed ABID. Hybrid approach for automatic classification of brain MRI using genetic algorithm and support vector machine. Leonardo Journal of Sciences 2010; 17: 71-82.

[8] Naik J, Patel S. Tumor detection and classification using decision tree in brain MRI. International Journal of Computer Science and Network Security 2014; 14: 49-53.

[9] Geethu M, Monica SM. MRI based medical image analysis: Survey on brain tumor grade classification. Biomedical Signal Processing and Control 2018; 39: 139-161. doi: 10.1016/j.bspc.2017.07.007

[10] Citak-Er F, Firat Z, Kovanlikaya I, Ture U, Ozturk-Isik E. Machine-learning in grading of gliomas based on multi-parametric magnetic resonance imaging at 3T. Computers in Biology and Medicine 2018; 99: 154-160. doi: 10.1016/j.compbiomed.2018.06.009

[11] Sharma M, Sourabh M. Brain tumor segmentation using genetic algorithm and artificial neural network fuzzy inference system (ANFIS). In: Meghanathan N, Nagamalai D, Chaki N, editors. Springer Advances in Computing and Information Technology. Advances in Intelligent Systems and Computing, Berlin, Heidelberg, Germany: Springer, 2013, pp. 329-339.

[12] Kavitha S, Bharathi B, Mohanapriya K, Mohanapriya S, Nithus SVC. Optimal feature set selection for brain tumor classification using genetic algorithm with support vector machine and decision tree. Asian Journal of Research in Social Sciences and Humanities 2016; 6 (9): 660-670. doi: 10.5958/2249-7315.2016.00825.X

[13] Tzu-Ping W, Shyi-Ming C. A new method for constructing membership functions and fuzzy rules from training examples. IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics) 1999; 29 (1): 25-40. doi: 10.1109/3477.740163

[14] Samuel OW, Omisore MO, Ojoko BA. A web based decision support system driven by fuzzy logic for the diagnosis of typhoid fever. Expert Systems with Application 2013; 40 (10): 4164-4171. doi: 10.1016/j.eswa.2013.01.030

[15] Ye CZ, Yang J, Geng DY, Zhou Y, Chen NY. Fuzzy rules to predict degree of malignancy in brain glioma. Medical and Biological Engineering and Computing 2002; 40 (2): 145-152. doi: 10.1007/BF02348118

[16] Nedeljkovic I. Image classification based on fuzzy logic. The International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences 2004; 34: pp. 3-7.

[17] Fazel ZMH, Zarinbal M, Izadi M. Systematic image processing for diagnosing brain tumors: A type-II fuzzy expert system approach. Applied Soft Computing 2011; 11 (1): 285-294. doi: 10.1016/j.asoc.2009.11.019

[18] El-Melegy MT, Hashim MM. Tumor segmentation in brain MRI using a fuzzy approach with class center priors. EURASIP Journal on Image and Video Processing 2014; 21: 1-14. doi: 10.1186/1687-5281-2014-21

[19] Khotanlou H, Atif J, Colliot O, Bloch I. 3D brain tumor segmentation using fuzzy classification and deformable models. In: Bloch I, Petrosino A, Tettamanzi AGB, editors. Springer Lecture Notes in Computer Science, Berlin, Heidelberg, Germany: Springer, 2006, pp. 312-318.

[20] Kavitha S, Thyagharajan KK. Features based mammogram image classification using weighted feature support vector machine. In: Krishna PV, Babu MR, Ariwa E, editors. Springer Communications in Computer and Information Science, Berlin, Heidelberg, Germany: Springer, 2012, pp. 320-329.

[21] Kavitha S, Mohanapriya K, Bharathi B. Brain tumor types and grades classification based on statistical feature set using support vector machine. In: IEEE 2016 ISCO 10th International Conference on Intelligent Systems and Control; Coimbatore, Tamilnadu, India; 2016. pp. 412-419.