BIOMEDICAL ENGINEERING | RESEARCH ARTICLE

An improved tumor segmentation algorithm from T2 and FLAIR multimodality MRI brain images by support vector machine and genetic algorithm

Aswathy Sukumaran, Devadhas G. Glan and S. S. Kumar

Cogent Engineering (2018), 5: 1470915
An improved tumor segmentation algorithm from T2 and FLAIR multimodality MRI brain images by support vector machine and genetic algorithm

Aswathy Sukumaran1*, Devadas G. Glan2 and S. S. Kumar3

Abstract: This paper puts forth a framework of a medical image analysis system for brain tumor segmentation. Image segmentation helps to segregate objects right from the background, thus proving to be a powerful tool in medical image processing. This paper presents an improved segmentation algorithm rooted in support vector machine (SVM) and genetic algorithm. SVMs are the basis of a machine-learning technique that has been used for segmentation and classification of medical images. The database consists of two weighted images: T2 and FLAIR. The proposed system mainly consists of two stages. The first stage performs pre-processing the MRI image, followed by block division. The second stage includes feature extraction, feature selection, and finally, the SVM-based training and testing. The feature extraction is done by first-order histogram and co-occurrence matrix; GA using wrapper method is used to select optimum subset features. Experimental outcomes for 10 patients with brain tumor are evaluated using the proposed method. The results of the algorithm have been compared with other algorithms. The performance of the algorithm has been evaluated using various measures i.e. Dice coefficient, Sensitivity, Specificity, and Jaccard coefficient.}

ABOUT THE AUTHORS
S. U. Aswathy, research scholar, Computer Science and Engineering Department, Noorul Islam University, Tamil Nadu, India. She completed B.Tech. (Electronics and Instrumentation) from Noorul Islam University and M.Tech. (Computer and Information Technology) From MS University 2005, 2009 respectively. Presently Working as HOD in Muslim Association college of Engineering, Kerala.

G. Glan DevaDhas is presently working as a Professor in the Department of Electronics and Instrumentation at Vimal Jyothi Engineering College, India. He received his B.E. Degree in Instrumentation and Control Engineering, M.E. Degree in Process Control and Instrumentation and Ph.D. in Intelligent Controller Design during 1998, 2001, 2013 respectively.

S. S. Kumar is presently working as HOD in the Department of Electronics & Instrumentation at Noorul Islam University, Tamil Nadu, India. He has received his B.E. Degree in Electronics and Instrumentation Engineering, M.E. Degree in Applied Electronics and Ph.D. in Information and Communication Engineering, Intelligent Controller Design during 2002, 2004, 2013 respectively.

PUBLIC INTEREST STATEMENT
Brain tumor detection from MR images is the main objective of this research work. In medical image segmentation, hand labeling of brain pathologies by a radiologist is often regarded as the gold-standard technique to segment abnormalities in the brain. This study provides an effective algorithm for detecting and segmenting edges of brain tumors that will help neurosurgeons to identify the boundary of the tumor region and to differentiate the clear-cut margin of the tumor from the rest of the brain tissue. This work exploits the importance of magnetic resonance imaging as a tool for the identification of brain tumor. This work proposes a new method using a combination of wrapper-based genetic algorithm and support vector machine to classify brain tumor. The superiority of the algorithm is established by comparing it with the state of the art algorithms such as level set method and fuzzy-based methods.

© 2018 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.
over 250 MRIs are proof of the proficiency of our work. The performance of the proposed system is evaluated in terms of specificity, sensitivity, accuracy; time elapsed and figure of merit. This approach also permits the segmentation of image volumes based on training sets selected on a single slice.

Subjects: Computer Engineering; Computer Science (General); Engineering Education; Biomedical Engineering; Electrical & Electronic Engineering

Keywords: segmentation; support vector machine; genetic algorithm; K nearest neighbors; wrapper method

1. Introduction

Brain tumor segmentation of MRI is very significant in medical image diagnosis, as it yields information related to anatomical structures as well as abnormal tissues which are absolutely necessary for treatment planning and patient follow-up. The brain image segmentation can also be helpful for general modeling and construction of pathological brains and brain atlases Olabarriago and Smeulders (2001). Although there have been numerous efforts and promising developments in the medical imaging community, characterization of abnormalities are still considered to be a challenging and difficult task, due to its widespread variety of possible shapes, locations, structure, and image intensities of various types of tumors Duncan and Ayache (2000). Some sections of the tumor may also deform the surrounding structures or the edema associated with it. The tumor might also affect necrosis that changes the image intensity around the tumor. Present methods set aside scope for increased automation, applicability, and accuracy. This paper aspires to subsidize to this area, by proposing a distinct method, which is universal enough to address a large class of tumor types.

There exist different methods for the image segmentation and processing Priyanka (2013). These methods can be supervised, semi-supervised, and unsupervised. Zhang, Brady, and Smith (2001) proposed a new method known as Hidden Markov Random Field (HMRF) model. This model has the ability to encode both the spatial and statistical properties of a given image. When compared with existing methods, it is more flexible for image modeling. The limitation of this system is that its preliminary estimations is based on threshold value which is purely heuristic and its time consuming. It gives inaccurate results most of the time. Tolba, Mostafa, Gharib, and Salem (2003) proposed a new method for MRI brain image segmentation known as Gaussian Multi-Resolution Expectation Maximization algorithm. This algorithm is based on EM algorithm and the multi resolution analysis of the given image. The limitation of this technique is that miss-classified pixel is generated when this algorithm is applied to pixel laying in the edges of boundaries. Agarwal and Kumar (2014) introduced a model based on a threshold value that uses level set methods for 3D brain tumor segmentation.

In this proposed model, the speed function of the level set was designed using a general threshold. Zhang et al. (2001) in their paper explain the number of fuzzy clustering methods based on fuzzy set theory. Fuzzy c mean (FCM) algorithm (Shen, Sandham, Granat, & Sterr, 2005) assigns every pixel to the clusters without any label; however, the algorithm fails to succeed completely to segment image having noise, images having the intensity difference, artifacts, etc. To overcome these draw back, modified FCM algorithm was proposed. Selvaraj and Dhanasekharan (2013) proposed a method to overcome the intensity inhomogeneity. Though this algorithm work well in medical image segmentation, it has some limitation such as time required for segmentation is more due to its complexity in spatial and additional terms. In our proposed system, we try
to overcome the limitation on the existing system such as time complexity, figure of merit absolute volume segmentation.

This paper recommends a fully automatic and effective tumor segmentation based on SVM classifier and wrapper-based genetic algorithm (GA). SVM classifier is used to classify the segmented region of the MRI image obtained by the block integrated process and truth image of the input is also taken, followed by feature extraction by using texture features and histogram-based method. Feature selection done by using GA and the optimum features are selected by wrapper-based method, and are used to train the SVM classifier Tolba, Mostafa, Gharib, and Salem (2003). The sum of the total block output gives the classified region. Basic block diagram is shown in Figure 1.

The remaining part of this paper is outlined as follows. Section 2 gives an idea of method overview. Section 3 explains about experimental discussion with manual delineation. A comparative analysis with the existing system is discussed in Section 4. Finally, Section 5 concludes the paper.

2. Method overview

2.1. Pre-processing
In pre-processing, the image is enhanced so as to get an improved form of the finer image and to remove unwanted noise present in the image. In the proposed method, importance is given to background separation so as to get the region of interest. The main processes in pre-processing are thresholding, region filling and morphological operation Vanitha, Prabhu Deepak, PonNagesharan, and Saucepan (2015).

Algorithm for pre-processing
Step 1: calculate the mean.
Step 2: Perform the thresholding operation.
Step 3: Region filling is done to remove the holes.
Step 4: Morphological operation is done to remove noise and small objects from the background.
Step 5: Region filling operation is done to fill the unfilled region
Step 6: Morphological Opening-Remove noise & small objects from the background

2.2. Feature extraction
Extracting a good feature set for classification is a challenging task. There are many features extraction techniques such as texture features, Gabor features, feature based on wavelet transform, principal component analysis, minimum noise fraction transforms, discriminant analysis, decision boundary feature extraction, non-parametric weighted feature extraction and spectral mixture analysis. Texture feature extraction is used for our recommended system.

2.2.1. Texture features extraction
The proposed system uses two approaches for texture feature extraction. The first-order histogram is the basis of the first method, and co-occurrence matrix-based method is the core of the second method.

2.2.1.1. First-order histogram-based feature extraction. A histogram of an image gives a summary of the statistical information about the image Ghosh and Mitchell (2006). On dividing the value of intensity level histogram with the total number of pixels in the image, the probability density of occurrence of the intensity levels can be given as follows:
\[ P(i) = \frac{\text{number of pixels with grey level } h(i)}{\text{Total no of pixels in the region } (N \times M)}, i = 0, 1, \ldots, G - 1 \tag{1} \]

where \( N \) represents the number of the resolution cells in the horizontal spatial domain and \( M \) represents the resolution cell number in the vertical spatial domain. The total gray level of the image is represented by \( G \). Useful features of the image can be obtained from the histogram for describing first-order statistical features of the image are mean, variance, skewness, kurtosis, energy, and entropy Seara (2002).

### 2.2.1.2. Co-occurrence matrix-based feature extraction

Here the feature is extracted based on gray-level spatial co-occurrence matrix \( h(i,j) \) Ong and Khoo (2009). This is also termed as second order histogram-based features. These extracted features are based on the joint probability distribution of pairs of pixels. To compute the joint probability distribution between pixels, distance \( d \) and angle \( \theta \) within a given neighbourhood are used. Normally, \( d = 1,2 \) and \( \theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ \) are used for calculation. Texture features can be labeled using this co-occurrence matrix. The following features are extracted angular second moment, correlation, inertia, absolute value, inverse difference, entropy, and maximum probability.

All such features are extracted from each block of the T2 and FLAIR modal slice of the same patient and are arranged in a matrix of dimension \( m \times n \), where \( m \) is the number of blocks and \( n \) is the number of features. It is clear that there are \( n/2 \) features each in T2 and FLAIR. From Tables 1 and 2, the first column contains the index of the block randomly selected for the sake of demonstration. The columns 2 through 6 contain, respectively, the mean, variance, kurtosis and entropy. The 7th column contains the Target, which stipulates whether the block belongs to tumor regions or not. “1” in the Target means it belongs to the tumor region Padma and Sukanesh (2011).

### Table 1. Feature extraction for randomly selected blocks (FLAIR IMAGES)

| Random block index | Mean | Variance | Skewness | Kurtosis | Entropy | Target |
|--------------------|------|----------|----------|----------|---------|--------|
| 14,222             | 385  | 2.1E-05  | -255280.5| 3.2E+13  | 2.9E-12 | 0      |
| 10,458             | 128  | 2.1E-05  | -85129.86| 3.2E+13  | 2.9E-12 | 0      |
| 5,401              | 818.7| 14190.2  | -0.80904 | 1.65461  | 0.89604 | 1      |
| 5,979              | 128  | 2.1E-05  | -85129.86| 3.2E+13  | 2.9E-12 | 0      |
| 8,050              | 385  | 2.1E-05  | -255280.5| 3.2E+13  | 2.9E-12 | 0      |
| 22,611             | 385  | 2.1E-05  | -255280.5| 3.2E+13  | 2.9E-12 | 0      |
| 9,667              | 899  | 2.1E-05  | -594385.5| 3.2E+13  | 2.9E-12 | 1      |
| 15,763             | 128  | 2.1E-05  | -85129.86| 3.2E+13  | 2.9E-12 | 0      |
| 1,196              | 128  | 2.1E-05  | -85129.86| 3.2E+13  | 2.9E-12 | 0      |
| 5,882              | 642  | 2.1E-05  | -425104.5| 3.2E+13  | 2.9E-12 | 1      |
| 24,137             | 385  | 2.1E-05  | -255280.5| 3.2E+13  | 2.9E-12 | 0      |
| 15,834             | 385  | 2.1E-05  | -255280.5| 3.2E+13  | 2.9E-12 | 0      |
| 5,108              | 128  | 2.1E-05  | -85129.86| 3.2E+13  | 2.9E-12 | 0      |
| 11,408             | 368.9| 3870.06  | -3.614785| 14.0676  | 0.33729 | 1      |
| 20,312             | 385  | 2.1E-05  | -255280.5| 3.2E+13  | 2.9E-12 | 0      |
2.3. Feature selection

To find the most significant feature set, optimized feature selection process is to be used, for which GA is a good approach (Frohlich, Chapelle, & Scholkopf, 2003). The probability of getting an optimal feature subset for classification is high when GA is employed for feature selection with suitable fitness functions and possible considerations. Taking these advantages, a feature selection scheme based on GA is proposed.

Input pattern of GA population of chromosome transformed pattern wavelet transform.

2.3.1. Genetic algorithm-based feature selection

The important phases in GA are chromosome encoding, fitness evaluation, selection mechanism, genetic operators, and the criteria to stop the iteration Ghosh and Mitchell (2006) from Figure 2, the GA operates on the binary search space as the chromosomes are taken as bit strings. To begin with, an initial population is created randomly and evaluated using the

---

**Table 2. Feature extraction for randomly selected blocks (T2 IMAGES)**

| Random block index | Mean | Variance | Skewness | Kurtosis | Entropy | Target |
|--------------------|------|----------|----------|----------|---------|--------|
| 14,222             | 385  | 2.1E−05  | −255,281 | 3.2E+13  | 2.9E−12 | 0      |
| 10,458             | 128  | 2.1E−05  | −85129.9 | 3.2E+13  | 2.9E−12 | 0      |
| 5,401              | 931  | 7224.11  | 2.267787 | 6.14311  | 0.54356 | 1      |
| 5,979              | 128  | 2.1E−05  | −85129.9 | 3.2E+13  | 2.9E−12 | 0      |
| 22,611             | 513  | 16512.3  | −1.5E−08 | 1.00005  | 1       | 0      |
| 9,667              | 1,413| 2.1E−05  | −930,466 | 3.2E+13  | 2.9E−12 | 1      |
| 11,520             | 128  | 2.1E−05  | −85129.9 | 3.2E+13  | 2.9E−12 | 0      |
| 1,196              | 128  | 2.1E−05  | −85129.9 | 3.2E+13  | 2.9E−12 | 0      |
| 5,882              | 546  | 15480.2  | −0.5164  | 1.26672  | 0.95443 | 1      |
| 24,137             | 385  | 2.1E−05  | −255,281 | 3.2E+13  | 2.9E−12 | 0      |

**Table 3. Volumetric measurement of manual and automatic segmentation for FLAIR and T2 images using GA and SVM-based method**

| Image type | Image no’s | Manual segmentation | Automatic segmentation | Absolute volume segmentation | Figure of merit |
|------------|------------|---------------------|------------------------|-----------------------------|-----------------|
|            | 1          | 22,301              | 27,007                 | 21.102                      | 0.82651         |
|            | 2          | 12,223              | 9,232                  | −24.47                      | 0.799067        |
| FLAIR      | 3          | 30,383              | 22,821                 | −24.889                     | 0.978903        |
|            | 4          | 9,682               | 11,453                 | 18.292                      | 0.787337        |
|            | 5          | 10,630              | 3,271                  | −69.229                     | 0.232455        |
|            | 6          | 22,301              | 22,732                 | 1.9326                      | 0.61226         |
|            | 7          | 12,223              | 6,437                  | −47.337                     | 0.939949        |
| T2         | 8          | 30,383              | 13,480                 | −55.633                     | 0.967449        |
|            | 9          | 9,682               | 8,282                  | −14.46                      | 0.653481        |
|            | 10         | 10,630              | 1,425                  | −86.595                     | 0.477234        |
fitness function. In the bit string which represents the chromosome, a bit value “1” depicts that the particular feature indexed by the position has been selected. Only those features selected by the 1s in the chromosome are taken in the testing. The ranking is based on the accuracy of classification based on an already classified test data. Based on the ranking, the top \( n \) fittest chromosomes are selected in order to survive to the next generation. The remaining chromosomes are passed through the functionals such as mutation and crossover which produces new and probably better chromosomes Faulkner (1998). The fitness of the new chromosomes is found. These steps continue until the fitness is reached. Basic Block diagram for fitness evaluation by GA is shown in Figure 3.

2.3.2. Initial population

Initial population is generally a matrix of dimension \( m \times n \), where \( m \) is the population size and \( n \) is the chromosome length. In this work, the chromosome length (genome length) is equal to the number of features extracted from a block of the image. The population size is the number of chromosomes taken in the population. Each bit of a chromosome as discussed earlier represents the position of a feature in the feature matrix. A “1” in the chromosome represents the selection of the corresponding feature from the feature matrix.

Algorithm: Initial population

Step 1: Procedure popFunc ()
Step 2: Pop \( \leq \) randomly created binary matrix of size, population Size \times GenomeLength
Step 3: Return pop
Step 4: End procedure

2.3.3. Fitness evaluation

For a GA to select subset features, a fitness function is defined to evaluate the discriminating capability of each subset. The KNN algorithm is used to evaluate the fitness function; it solves the classification problem by looking for the shortest distance between the test data and the training sets in the feature space. The KNN algorithm computes the Euclidean distance between test data and the training sets and then find the nearest point of the training set to the test set:

\[
D(x_{\text{test}}, x_i) = \sqrt{\sum_{m=1}^{M} (x_{\text{test}} - x_i)^2}
\]  

(2)

As the GA iteration proceeds, the individual chromosomes in the current population are evaluated and the fitness is ranked based on the kNN classification error. The main aim behind the iteration is to reduce the error rate and pick the individual with the least (best) fitness value:

\[
\text{fit} = \alpha + \exp\left(\frac{-1}{N_f}\right)
\]

(3)

where \( \alpha \) is the kNN-based classification error and \( N_f \) is the cardinality of the selected features. The algebraic structure of this equation ensures the learning of the GA, error minimization, and reduced number of features selected.

Algorithm: fitness function evaluation

Step 1: Procedure fit().
Step 2: featureIndex = indices of ones from binary chromosome.
Step 3: new DataSet = dataset indexed by featureIndex.
Step 4: featCount = number of elements in featureIndex
Step 5: kNNneighborCount = 3
Step 6: kNN error \( \leq \) classifierkNN(DataSet, classInfo, kNNneighborCount).
Step 7: return kNNerror
Step 8: end procedure
2.3.4. Generation of children for new population
After fitness evaluation, a new population is created using genetic operator such as crossover and mutation. Three types of children are created, which are Elite children, crossover children and mutation children.

2.3.4.1. Elite children. These children are pushed automatically into the next generation as these are the best children with greatest fitness so far. The count of elite chromosome is bounded by population size. This implies \( \text{EliteCount} \leq \text{populationSize} \), with size 2, GA picks the two best chromosomes and pushes them automatically to the next generation. Consider the case where the population size is 100 and the EliteCount = 2. This means there are 98 (100-EliteCount) individuals in the population apart from the elite kids. Crossover and mutation kids are then produced.

2.3.4.2. Crossover children. The proportion of the next generation apart from the left over kids that are produced by crossover is called crossover fraction. If the fraction is set to one, then there are no mutation kids. The crossover fraction used in this work is 0.8, then the number of crossover children will be \( \text{COcount} = \text{round}(98*0.8) = 78 \).

2.3.4.3. Mutation children. Finally, the number of mutation children is \( \text{Mcount} = 100 - \text{EliteCount} - \text{COcount} = 100 - 2 - 78 = 20 \).
This implies \( \text{EliteCount} + \text{COcount} + \text{Mcount} = 100 \).

2.3.5. Selection mechanism: tournament
The aim of selection mechanism in GA is to make sure that the population is being constantly improved in accordance with the overall fitness values. The selection mechanism helps the GA in discarding bad designs and keeping only the best individuals. Tournament selection of size 2 is used in this proposed work due to its speed, simplicity, and efficiency. Also, tournament selection enforces higher selection pressures on the GA and makes sure the worst individual does not get into the next generation. Two functions are needed to perform in tournament selection. The first function generates the players (parents) needed in the actual tournament function, while the second function outputs the winner of the tournaments. In the tournament selection of size 2, two chromosomes are selected from the population, after the elite kids are taken out and the best of the two chromosomes is selected. This process is performed iteratively until the newly generated population is filled up.

2.3.6. Crossover function
The crossover operator in the GA genetically combines two individuals named parents to form children for the next generation. Two chromosomes are taken from tournament selection. The crossover fraction specifies the number of kids produced by the crossover functional, after Elite kids are removed from the current population being evaluated. The crossover fraction is bounded by 0 and 1. The value of crossover fraction used in this work is 0.8 and the crossover function is arithmetic type. An XOR operation is performed on the two parent chromosomes since they are binary. For two parents p1 and p2, such as p1 = 1 0 0 0, 1, p2 = 0 1 1 0, 1, crossover kid = p1 XOR p2 = 1 1 1 0 0.

2.4. Wrappers for feature subset selection
Wrapper approach is a random search technique for subset generation, which takes the advantage of the supervised learning algorithm to evaluate the significance of the feature subset and employs the GA to optimize the searching of features in the feature selection process. Karegowda, Jayaram, and Manjunath (2010). In our proposed system wrapper approach along with GA is used for random search subset generation. Here, from Figure 4 subset feature generation consists of four steps: (1) generate candidate subset by GA, (2) subset evaluation function to evaluate the subset generated in the previous step by using wrapper approach, (3)
stopping Condition (until the optimal subset according to some evaluation function is obtained), and (4) validation procedure to check whether the feature subset selected is valid.

2.5. SVM segmentation

SVM is generalized statistical learning theory Vapnik (1995), which is used for classifying the set of inputs based on supervised learning theory Kim, Jung, Park, and Kim (2002). It separates the set of inputs into two classes with class labels −1 and 1. The classifier separates into classes by defining a hyper plane defined by\( f(x) = β^T x + α \), where \( β \in \mathbb{R}^n \) is orthogonal to the hyper plane and \( α \in \mathbb{R} \) is a constant. For the set of training feature set \( f(a_i, b_j) \) defined in the form

\[
f(a, b) = \{ (a_i, b_i) / a_i \in \mathbb{R}^m, b_i \in [0, 1] \}_{i=1}^n.
\]

(4)

Here, \( a_i \) is the input vector and \( b_i \) is the target vector with two classes with class labels 0 and 1. SVM training will maximize the distance between the classes by defining the hyperplanes with a small empirical risk Li, Zhou, Wang, and Wu (2011) and Nandpuru, Salankar, and Bora (2014). The hyper planes for the classes can be represented as \( b_i (β^T a_i + α) \geq 1 \). The objective function for optimizing the distance between the classes defined by the hyper planes can be represented as

\[
\text{minimize: } P(β, α, ε) = 0.5 \| β \| + C \sum_{i=1}^{n} e_i
\]

(5)

where \( e_i \) is a set of slack variables and \( C \) is a cost variable for each slack variable.

2.5.1. Training and testing phase

Proposed SVM-based segmentation consists of two parts: (1) testing and (2) training. Firstly, the image is divided into blocks and features extracted from each block are arranged in the feature matrix. A truth vector is formed from the corresponding truth image. This matrix is used to train the classifier. SVM classifies data by finding the best hyperplane that separates all specific data points of one class from those of the other shown in Figure 5. The finding of the hyper plane and the corresponding support vector is called Training Nandpuru et al. (2014). For segmentation, the test image undergoes the feature extraction. Feature Selection is done with the help of wrapper-based GA. These selected features are used for SVM classifier training Niyogi, Burges, and Ramesh (1999). Here SVM classifier is trained to classify a tumor region and abnormal region. This classification takes place block wise, and the sum total of the classified blocks can give out the classified region Niyogi et al. (1999).

**Algorithm for training phase**

- **Step 1:** Divide the input database image into blocks of size \( m \times m \)
- **Step 2:** Extract features of each block and arrange them in a feature matrix. Features are calculated.
- **Step 3:** The extracted feature vector of each block is arranged on feature matrix of size \( N \times 5 \), \( (N = \text{no of block}) \).
- **Step 4:** Construct the truth vector from the Truth image.
- **Step 5:** SVM is trained using the feature matrix and the corresponding Truth vector.

**Algorithm for testing phase**

- **Step 1:** Divide the images into blocks of size \( m \times m \)
- **Step 2:** Extract features of each block and place them in a feature matrix and features are calculated.
- **Step 3:** Construct an \( M \times N \) Zeros matrix where, \( M \) represent rows and \( N \) represent column size.
- **Step 4:** feed feature vector of each block into already trained SVM. If the SVM output is 1, The resultant block in the Zero matrix is made 1.
- **Step 5:** repeat 2–4. This makes the Zeros matrix change each time, the SVM gives 1 output.
Finally gives a map of the tumor location.

3. Experimental results and discussion with manual delineation

For the performance evaluation, a series of 30 images (T2 and FLAIR) were taken in the same class of tumor and segmented with a common threshold value, in association with the radiologist segmented image. For the numerical analysis, the number of false negatives (FN) and false positives (FP) are calculated, based on the number of pixels in the interested region (ROI). The four parameters—TP, FP, TN, and FN—are evaluated by the logically AND-ing the ground truth and segmented image Zeljkovic et al. (2014). The proposed methods have been implemented by using MATLAB software with a computer specification of core 2duo having a processor speed of 1.6 GHz. The proposed algorithm is implemented on real human brain MRI data set available on the web. In our proposed method a data set consisting 30 medical images (23 abnormal and 7 normal) were taken into consideration, of these 95 objects are used in the training and 30 objects were used in testing phase. The images in the database have been exported into a set of 2D axial slice in the MHA format of pixel size 256 × 256. In our proposed system the images are divided into a block of size (m X m). The Supervised evaluation is carried out to figure out the difference between the ground truth and the segmentation result using the evaluation metric. The ground truth is obtained from manual segmentation. The results of the proposed SVM-based algorithm are evaluated by comparing it with the result of manually segmented data done by medical experts Wang (2006). The degree of similarity of the segmented results gives the accuracy of the segmented image. The result is obtained in terms of specificity, accuracy, absolute volume measurement error (AVME) and figure of merit (ε), helps to demonstrate the validity of the proficiency:

Specificity \( \delta = \frac{TN}{TN + FP} \)  

Accuracy \( A = \frac{(TN + TP)}{(TN + TP + FN + FP)} \)  

Absolute Volume Measurement Error (AVME) \( = \frac{V_{\text{automatic}} - V_{\text{manual}}}{} \times 100\% \)  

Figure of Merit \( (\varepsilon) = 1 - \frac{|V_{\text{manual}} - V_{\text{automatic}}|}{V_{\text{manual}}} \)  

Table 1 illustrates the features extracted from the randomly selected blocks of an image, FLAIR, and T2. First column of Table 1 contains the index of the block randomly selected for the sake of illustration. The columns 2 through 6 contain, respectively, the mean, variance,
kurtosis and entropy. The 7th column contains the Target, which specifies whether the block belongs to tumor regions or not. A “1” in the Target means it belongs to the tumor region. Table 2 shows the comparison result of volume measurements of manual segmentation with the proposed system for FLAIR and T2 images. The degree of accuracy is compared with the manual segmentation done by the radiologist. The time required for processing each image for 10 patients for T2 and FLAIR is plotted as graphs in Figure 6. Figure 7 shows the statistical evaluation, showing specificity and accuracy of multimodality images of FLAIR and T2. The performance of the classifier is evaluated from the confusion matrix, specificity and accuracy by testing it with the features extracted from the feature set such as mean; histogram; variance; standard deviation; skewness; kurtosis; entropy. The proposed scheme gives out a better performance on the basis of sensitivity, which is 96.799 in the case of FLAIR modality whereas 91.2326 in the case of T2 modality. The specificity was reported to be 98.3240 for
FLAIR, and 99.2615 for T2. The accuracy was seen to be 98.2257 for FLAIR and 98.7428 for T2. Thus, it is clear that both FLAIR and T2 are suitable for this scheme for tumor detection. Moreover, the Figure of Merit has hiked up to 0.8257 which is really a high value. The experimental results obtained from the simulation of the given database, show the effectiveness of the advised system.
4. Comparative analysis with existing system

Proposed method is compared with two existing methods such as expectation maximization using level set (Aswathy, Kumar, & Gan Devadhas, 2014, 2015) and FCM method (Agarwal & Kumar, 2014; Gordillo, Montseny, & Sobrevilla, 2013). The comparison diagram of proposed SVM-based system with the existing method such as combination of EM & level set and FCM using T2 and FLAIR images is shown in Figure (9c & f). Figure (9b) shows the segmented image output for the method such as EM and level set, FCM, and SVM for same image input is shown. The results of the proposed system are evaluated by comparing it with the result of manually segmented data done by medical experts; the degree of similarity of the segmented results gives the accuracy of the segmented image. The proposed method is evaluated in terms of specificity, accuracy, AVME and figure of merit (ε). The evaluation metrics of sensitivity, specificity, and accuracy can be stated in the terms of TP, FP, FN, and TN based on their pixel values. A comparison of the proposed SVM-based system with the existing method such as combination of EM and level set and FCM using T2 and FLAIR is shown in Tables 4 and 5.
Table 4. Statistical comparison for the FLAIR and T2 input images for existing system

| Input image | Specificity | Accuracy | Time elapsed |
|-------------|-------------|----------|--------------|
|             | EM + level set | FCM | GA + SVM | EM + level set | FCM | GA + SVM | EM + level set | FCM | GA + SVM |
| 1           | 97.202       | 57.7914 | 100   | 98.6777 | 90.4845 | 99.2594 | 10.679036 | 0.10644 | 0.687467 |
| 2           | 97.1383      | 57.0374 | 100   | 98.8108 | 92.7889 | 99.2405 | 11.280992 | 0.10334 | 0.701177 |
| 3           | 96.8692      | 54.8225 | 100   | 98.8223 | 91.524  | 99.3658 | 10.734473 | 0.095835 | 0.683585 |
| 4           | 97.2801      | 58.4199 | 99.9413 | 98.5446 | 93.6795 | 98.8777 | 10.589031 | 0.108671 | 0.652442 |
| 5           | 97.3814      | 58.8072 | 99.9568 | 98.7558 | 94.221  | 99.0586 | 11.895542 | 0.107736 | 0.960127 |
| 6           | 98.6256      | 79.3182 | 99.9382 | 98.2523 | 96.3182 | 98.3838 | 11.326163 | 0.075599 | 0.236125 |
| 7           | 98.6285      | 80.3202 | 99.8274 | 98.1481 | 96.2335 | 98.2735 | 11.905631 | 0.079258 | 0.311487 |
| 8           | 98.3767      | 78.481  | 99.6854 | 98.3333 | 96.8589 | 98.4303 | 11.230366 | 0.072311 | 0.226894 |
| 9           | 97.9919      | 74.1494 | 99.6623 | 98.588  | 94.8741 | 98.8482 | 11.577909 | 0.106921 | 0.248339 |
| 10          | 98.2436      | 77.1106 | 99.6908 | 98.7847 | 95.6559 | 98.9983 | 11.697833 | 0.081979 | 0.242185 |
| Image type | Image Input | Manual | Automatic segmentation | Absolute volume measurement | Figure of merit |
|------------|-------------|--------|------------------------|-----------------------------|-----------------|
|            |             |        |                        |                             |                 |
| FLAIR      | 1           | 22,301 | 78,407                 | 18,432                      | 27,007          |
|            |             |        |                        | 251,5851                    | 21,102          | 0.788978       |
|            |             |        |                        | -17.35                      | -24.47          | 0.795297       |
|            |             |        |                        | 20,093                      | -24.889         | 0.751111       |
|            |             |        |                        | 16,45677                    | 0.817083        |
|            |             |        |                        | 0.755297                    |
|            |             |        |                        | 0.980674                    |
|            | 2           | 12,223 | 74,201                 | 14,679                      | 9,232           |
|            |             |        |                        | 507,0605                    | 10,943          |
|            |             |        |                        | 20,093                      | -24.47          |
|            |             |        |                        | 0.751111                    |
|            |             |        |                        | 0.817083                    |
|            |             |        |                        | 0.795297                    |
|            | 3           | 30,383 | 35,383                 | 29,742                      | 22,821          |
|            |             |        |                        | 16,45677                    | 18,292          |
|            |             |        |                        | 0.788978                    |
|            |             |        |                        | 0.799067                    |
|            |             |        |                        | 0.835434                    |
|            |             |        |                        | 0.978903                    |
|            | 4           | 9,682  | 957,832                | 7,623                       | 11,453          |
|            |             |        |                        | 979,2915                    | 18,292          |
|            |             |        |                        | 0.817083                    |
|            |             |        |                        | 0.978903                    |
|            |             |        |                        | 0.787337                    |
|            | 5           | 10,630 | 108,800                | 18,789                      | 3,271           |
|            |             |        |                        | 923,5183                    | -69,229         |
|            |             |        |                        | 0.307714                    |
|            |             |        |                        | -8.23518                    |
|            |             |        |                        | 0.978903                    |
|            | 6           | 22,301 | 120,369                | 13,654                      | 22,732          |
|            |             |        |                        | 439,7471                    | 1,9326          |
|            |             |        |                        | 0.980674                    |
|            |             |        |                        | -3.39747                    |
|            |             |        |                        | 0.61226                     |
|            | 7           | 12,223 | 101,747                | 12,957                      | 22,732          |
|            |             |        |                        | 64,377                      |
|            |             |        |                        | 732,4225                    |
|            |             |        |                        | 60,051                      |
|            |             |        |                        | 732,4225                    |
|            |             |        |                        | 47,337                      |
|            |             |        |                        | 0.52663                     |
|            |             |        |                        | -6.32422                    |
|            |             |        |                        | 0.939949                    |
|            | 8           | 30,383 | 108,938                | 29,394                      | 13,480          |
|            |             |        |                        | 258,5492                    | -55,633         |
|            |             |        |                        | 0.443669                    |
|            |             |        |                        | -1.58549                    |
|            |             |        |                        | 0.967449                    |
|            | 9           | 9,682  | 85,949                 | 8,327                       | 8,282           |
|            |             |        |                        | 787,7195                    |
|            |             |        |                        | -34,65                      |
|            |             |        |                        | 0.855402                    |
|            |             |        |                        | -6.87719                    |
|            |             |        |                        | 0.653481                    |
|            | 10          | 10,630 | 109,225                | 16,187                      | 1,425           |
|            |             |        |                        | 927,5165                    |
|            |             |        |                        | 52,277                      |
|            |             |        |                        | -86,595                     |
|            |             |        |                        | 0.134055                    |
|            |             |        |                        | -8.27516                    |
|            |             |        |                        | 0.477234                    |
Tables 4 and 5 show the statistical evaluation of the methods, such as combination of EM and level set, FCM, and SVM method for T2 and FLAIR inputs. The time required for processing each image for T2 and FLAIR for the methods is shown in Table 4. Table 3 gives the volumetric analysis for manual and automatic segmentation for the combination of EM and level set, FCM, and SVM-based method for T2 and FLAIR input image is shown. The degree of similarity between the manual segmentation and the machine segmented results gives the similarity of the segmented results. Theoretical studies prove that the combination of EM and level set algorithm helps to detect the tumor boundaries more accurately. Figure 9(a) shows that it fails to produce so. Table 4 shows that the computational time required for FCM is very less compared with the other two algorithms. Statistical comparison and tumor volume calculation from Tables 4 and 5 for the proposed algorithm show the efficiency. The performance of the classifier is evaluated from the confusion matrix, specificity, and accuracy by testing it with the features extracted from the feature subset by a wrapper method such as mean, histogram, variance, standard deviation, skewness, kurtosis, and entropy. The experimental results show the effectiveness of the proposed system and this algorithm turns out to an efficient fully automatic brain tumor segmentation.

5. Conclusion and future works
Fully automatic brain tumor segmentation is one of the snags faced in medical image processing. Even if much exertion has been made in discovering out a good solution for this problem, it still remains unsolved. In our proposed system, authors are using wrapper-based GA for random search technique in identifying the best features and SVM as a classifier. The comparative analysis of the proposed system with other existing system such as FCM and EM and level set is carried out and the performance of the proposed method is analyzed in terms of number of subset features selected, figure of merit, sensitivity, specificity, AVME, and the classification accuracy. From the conducted experiments, it is evident that the proposed method outperforms other methods compared and produces better accuracy for the classifier which is adopted as the feature evaluation mechanism in the feature selection process. However, experiment results show that employing feature subset selection surely enhances the classification accuracy. Hence, this proposed method is well suited for the medical application where the classification algorithm is predefined. In the future, matrix transformations of the image may be used which could reduce the complexity and time. In order to avoid manual error, author’s tried to propose a fully automatic intelligent system by making use of the advantage of finding the global optimum by using GA, thereby reducing the error rate of brain tumor segmentation.

List of abbreviations
- MLE: maximum-likelihood estimator
- FCM: fuzzy C mean
- EM: expectation maximization
- SVM: support vector machine
- GA: genetic algorithm
- KNN: K-nearest neighbor
- TN: true negative
- TP: true positive
- FN: false negative
- FP: false positive

Acknowledgements
I would like to express my sincere gratefulness to Dr. G. Glan Deva Dhas, professor (HOD), Department of Electronics and Instrumentation, Vimal Jyothi College of engineering, Kannur, Kerala and Dr. Kumar S.S., associate professor (HOD), Department of Electronics and Instrumentation, Noorul Islam University, Thuckalay, Tamilnadu for their valuable guidance and for inspiring me to do the research work continuously.

Funding
No funding from any organization.
Competing Interest
All authors declare that they have no conflicts of interest, including no financial, personal, or other relationships with other people or organizations.

Author details
Aswathy Sukumaran1
E-mail: aswathy.su@gmail.com
ORCID ID: http://orcid.org/0000-0003-0459-9260
Devadas G. Glan2
E-mail: glandeva@gmail.com
S. S. Kumar3
E-mail: kumar_s_s@hotmail.com
1 Department of Computer Science, Noorul Islam University, India.
2 Electronics and Instrumentation, Vimal Jyathi Engineering College, Kerala, India.
3 Electronics and Instrumentation, Noorul Islam University, India.

Data sharing
Please contact author for data requests.

Authors’ contributions
SU carried out the segmentation work and SS and G carried out the analysis part, i.e., comparison with existing method such as FCM and combination of EM and level set method.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Citation information
Cite this article as: An improved tumor segmentation algorithm from T2 and FLAIR multimodality MRI brain images by support vector machine and genetic algorithm, Aswathy Sukumaran, Devadas G. Glan & S. S. Kumar, Cogent Engineering (2018), 5: 1470915.

Cover image
An improved tumor segmentation algorithm from T2 and FLAIR multimodality MRI brain images by support vector machine and genetic algorithm

References
Agarwal, V. P., & Kumar, M. (2014, May). Tumor extraction by level set method using threshold algorithm. International Journal of Computer Applications (0975–8887), 94(15), 38–41.
Aswathy, S. U., Kumar, S. S., & Glan Devadas, G. (2014). A survey on detection of brain tumour from MRI brain images. IEEE Explore Digital Library, 871–877. doi:10.1109/ICC ICTCT.2014.6993081
Aswathy, S. U., Kumar, S. S., & Glan Devadas, G. (2014). A survey of segmentation of brain tumor from MRI brain images. International Journal of Applied Engineering Research, 9(26), 9267–9275. ISSN 1993-4562
Aswathy, S. U., Kumar, S. S., & Glan Devadas, G. (2015, December). Quick detection of brain tumor using a combination of E-M and level set method. Indian Journal of Science and Technology, 8(34). ISSN: 0974-6846. doi:10.17485/ijst/2015v8i34/ 85361
Duncan, J. S., & Ayache, N. (2000). Medical image analysis: Progress over two decades and the challenges ahead. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22, 85–106. doi:10.1109/ 34.824822
Faulkner, E. (1998). Genetic algorithms and grouping problems. Boston, MA: Wiley.
Frohlich, H., Chapelle, O., & Scholkopf, B. (2003). Feature selection for support vector machines by means of genetic algorithm. In 15th IEEE international conference on tools with artificial intelligence (pp. 142–148). ISBN: 1-59593-186-4, @ USA, IEEE.
Ghosh, P., & Mitchell, M. (2006, July 8–12). Segmentation of medical images using a genetic algorithm. In GECCO (pp. 1171–1175).
Gopal, P., & Sukanesh, R. (2013). Wavelet statistical texture features-based segmentation and classification of brain computed tomography images. IET Image Processing, 7, 25–32. doi:10.1049/iet- ipr.2012.0073
Gordillo, N., Montseny, E., & Sobrevilla, P. (2013). State of the art survey on MRI brain tumor segmentation. Magnetic Resonance Imaging, 31(8), 1426–1438. doi:10.1016/j.mri.2013.05.002
Karegowda, A. G., Jayaram, M. A., & Manjunath, A. S. (2010). Feature subset selection problem uses wrapper approach in supervised learning. International Journal of Computer Applications (0975–8887), 1(7), 13–19.
Kavitha, A. R., DivyaMeena, M., Gayathri, K., & Raghav, K. P. (2015). Brain tumor segmentation using genetic algorithm with modified region growing method. International Journal of Emerging Technology in Computer Science and Electronics, 13, 1–7.
Kim, K. I., Jung, K., Park, S. H., & Kim, H. J. (2002, November). Support vector machines for texture classification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(11), 1542–1550. doi:10.1109/TPAMI.2002.1046177
Li, M., Zhou, X., Wang, X., & Wu, B. (2011). Genetic algorithm optimized SVM in object-based classification of quickbird imagery. In Spatial Information Research Center of Fujian Province, Fuzhou University, Fuzhou 350002, China, 2011 IEEE (pp. 348–352).
Logeshwaran, C., Bharathi, P., & Gowthami, M. (2015, May). Brain tumor detection uses hybrid techniques and support vector machines. International Journal of Advanced Research in Computer Science and Software Engineering, 5, 1–8.
Nandpuru, H. B., Salankar, S. S., & Bora, V. R. (2014). MRI brain cancer classification using support vector machine. In IEEE students’ conference on electrical, electronics and computer science, March, Bhopal, India, IEEE.
Niyogi, P., Burges, C., & Ramesh, P. (1999). Distinctive feature detection using support vector machines. In ICASSP 99 (Vol. 1, pp. 425–428), USA, IEEE. ISSN: 1520-6149.
Olabarriaga, S. D., & Smuclers, A. W. M. (2001, June). Interaction in the segmentation of medical images: A survey. Medical Image Analysis, 5(2), 127–142. doi:10.1016/S1361-8415(00)00041-4
Ong, H.-C., & Khoo, H.-K. (2009). Improved image texture classification using gray level co-occurrence probabilities with support vector machines postprocessing. European Journal of Scientific Research, 36(1), 56–64. ISSN 1650-216X.
Padma, A., & Sukanesh, R. (2011). Automatic classification and segmentation of brain tumor in CT images using optimal dominant gray level run length texture
features. *International Journal of Advanced Computer Science and Applications*, 2(10). doi:10.14569/issn.2156-5570

Priyanka, B. S. (2013, July). A review on brain tumor detection using segmentation. *International Journal of Computer Science and Mobile Computing (IJCSCM)*, 2(7), 48–54. ISSN 2320–088X.

Seara, R. (2002). Image segmentation by histogram thresholding using fuzzy sets. *IEEE Transactions on Image Processing*, 11(12), 1457–1465. doi:10.1109/TIP.2002.806231

Selvaraj, D., & Dhanasekharan, R. (2013, January–February). MRI brain tumor detection by histogram and segmentation by modifying GVF model.

Sharma, M., & Mukharjee, S. Brain tumor segmentation using genetic algorithm and Artificial Neural Network Fuzzy Inference System (ANFIS). In *Advances in computing & information technology, AISC 177* (pp. 329–339). Retrieved from springer link.com, 2013.

Shen, S., Sandham, W., Granat, M., & Sterr, A. (2005). MRI fuzzy segmentation of brain tissue using neighborhood attraction with neural-network optimization. *IEEE Transactions on Information Technology in Biomedicine*, 9(3), 459–467. doi:10.1109/TITB.2005.847500

Suken, J. A. K., & Vandewalle, J. (1999). Least square support vector machine classifier. *Neural Processing Letter*, 9,1–10.

Tolba, M. F., Mostafa, M. G., Gharib, T. F., & Salem, M. A. (2003). MR-brain image segmentation using gaussian multi resolution analysis and the EM algorithm. In *ICEIS, Research gate* (pp. 1–11).

Vanitha, U., Prabhu Deepok, P., PonNagesharon, N., & Saucepan, R. (2015). Tumor detection in brain using morphological image processing. *Journal of Applied Science and Engineering Methodologies*, 1, 1–6.

Vapnik, V. (1995). The nature of statistical learning theory. New York, NY: Springer-Verlag.

Wang, G. S. (2006). Properties and construction methods of kernel in support vector machine. *Computer Science of China*, 33(6), 172–178.

Zeljkovic, V., Druzgalski, C., Zhang, Y., Zhu, Z., Xu, Z., Zhang, D., & Mayorga, P. (2014, April 7–12) @ Brazil. Automatic brain tumor detection and segmentation in MR images. ISBN: 978-1-4799-3555-0. IEEE Catalog Number: CFP1418G-ART.

Zhang, Y., Brady, M., & Smith, S. (2001, January). Segmentation of brain MR images through a Hidden Markov Random Field Model and the expectation-maximization algorithm. In *Proceedings of the IEEE transaction on medical images*. doi:10.1109/42.906424