Feature selection based on hybrid optimization for magnetic resonance imaging brain tumor classification and segmentation

Ahmed KHARRAT[1,*] and Mahmoud NEJI[2]

[1]University of Sfax, MIRACL Laboratory ISIMS, SakietEzzeit, BP 242-3021 Sfax, Tunisia.
[2]University of Sfax, MIRACL Laboratory FSEG, Elmatar, BP 1088-3018 Sfax, Tunisia.
E-mails: ahmed.kharrat@isims.usf.tn; mahmoud.neji@fsegs.rnu.tn

*Author to whom correspondence should be addressed; Tel.: 00216-21-400-664; Fax: 00216-74- 862-432

Abstract
With the health information technology being infused into clinical health, e-health is becoming a key factor in delivering improvements in the health sector. Brain tumor data feature selection is crucial for the development of a viable cancer detection system based on brain tumor data. Our study aimed to obtain an optimal feature subset through a hybrid algorithm of Simulated Annealing-Genetic Algorithms (SA-GA). Two real datasets of brain tumor Magnetic Resonance Images are used to assess the performances of the proposed approach. The first dataset was freely downloaded from the Harvard Medical School brain atlas. The second brain tumor dataset was created from Nanfang Hospital, Guangzhou, China, and General Hospital, Tianjing Medical University, China from 2005 to 2012. The proposed approach is compared to the methods of simulated annealing, genetic algorithm and with the state-of-the-art methods used separately. The obtained results show that SA-GA exceeds simulated annealing and genetic algorithms when they are applied in isolation, in terms of accuracy and computing time. The evaluation shows that our method overtakes the state-of-the-art methods with a segmentation accuracy rate of 97.82%±0.74 for glioma tumor and 95.12% ±3.21for pituitary tumor.

Keywords: Simulated Annealing; Genetic Algorithms; Feature Selection; Computing time; Segmentation

Introduction

Cancer is a leading cause of morbidity and mortality globally. A rise in the number of worldwide cancer cases will lead to an increase in the number of cancer deaths, and soon cancer is expected to overtake heart disease as the leading cause of death. In fact, it is estimated that 18 million new cancer cases will be reported in 2018 and about 9.6 million people are expected to die of cancer this year.

Brain tumors are responsible for the increased mortality rate among different age categories. Brain tumors are manifested through the abnormal growth of the cells inside or around the brain [1]. The National Brain Tumor Foundation (NBFT) reported that in the last three decades, the total number of people that developed brain tumors and died from them has almost tripled [2]. Detecting the brain tumor in its early stages is of great importance and a major challenge for further studies. Therefore, the analysis of the tumor and its area is done by computers and image processing devices.

In the last two decades, Computer-Aided Detection (CAD) has developed rapidly. CAD systems drastically improved radiological diagnostic accuracy and minimized the time and effort necessary for

diagnosis, decreased missed cancerous case detection and improved inter- and intra-reader variability [3]. For this reason, pattern recognition techniques, including machine learning, are vital to the development of CAD systems [4]. To create a CAD system, various image processing techniques such as image segmentation, feature extraction and selection, and classification were integrated. Feature selection represents an active research domain in pattern recognition [5], machine learning [6] and data mining [7]. Irrelevant and redundant features invite further search as they make patterns less detectable and rules necessary for forecasting or classification less evident, in addition to the high overfitting risk. The selection of feature subsets requires determining the appropriate feature to maximize the accuracy of prediction or classification. The principal aim of this study was to determine an optimal feature subset. Selecting features is usually based on the parameters of computational time and the quality of the generated feature subset solutions. In fact, fast and accurate classification, using the minimum number of features is often opted for. This can apparently be obtained through feature selection. We proposed a novel hybrid algorithm for an optimal selection of feature subsets able to classify brain tumors as benign or malignant.

Material and Method

Review of Existing Techniques

Existing feature selection methods can be classified into filter models and wrapper models. Filter approaches score and rank features according to certain statistical criteria and select the highest ranking features. Filter models include t-test [8], chi-square test [9], Wilcoxon Mann-Whitney test [10], mutual information [11], Pearson correlation coefficient [12], and principal component analysis [13]. Despite its speed, the filter model is not robust against interactions among features and feature redundancy and may not produce the optimal subset of features.

The wrapper technique takes into account the feature subset and the regression model interactions. Wrapper methods use a learning algorithm to look for the most appropriate subset of features and to assess the accuracy of possible subsets in predicting the target.

Two broad categories of wrapper methods have been identified, greedy and stochastic.

Sequential backward selection (SBS) (also called backward stepwise elimination) and Sequential forward selection (SFS) (also called forward stepwise selection) are the two most common wrapper methods using the strategy of greedy hill-climbing search. Stochastic algorithms are developed to solve large-scale combinatorial problems. Particle swarm optimization (PSO), ant colony optimization (ACO), genetic algorithm (GA), and simulated annealing (SA) are leading feature subset selection research issues [14]. This type of algorithms can effectively capture feature redundancy and interaction. It does not depend on the restrictive monotonicity assumption. They yield the best feature subset, but they are computationally expensive.

Several hybrid approaches were proposed. For instance, a feature selection algorithm based on correlation and a genetic algorithm [15], t-statistics and a genetic algorithm [16], mutual information and a genetic algorithm [17], principal component analysis and an ACO algorithm [18], chi-square approach and a multi-objective optimization algorithm [19] rely on filter and wrapper methods. Interestingly, these approaches first apply filter methods to select a feature pool, then the wrapper method to obtain the optimal feature subset from the selected feature pool. Thus, feature selection becomes faster as in the filter method, the adequate number of considered features is rapidly reduced. Despite the low probability of proper predictor elimination by filter methods, hybrids of filter and wrapper methods can be inaccurate because an isolated relevant feature can be as discriminating as an irrelevant one in the presence of feature interactions.

Wrapper schemes use the K-Nearest Neighbor (K-nn) as a learning algorithm. In this approach, feature selection is "wrapped" in a learning algorithm. K-nn can be used for numerous training sets and provide accurate information about distance, weighted average and pixels. Meanwhile, an accurate K-nn algorithm depends on the presence of noisy or irrelevant features, or feature scales inconsistent with their importance. Moreover, the choice of k affects the K-nn algorithm. Empirical evidence suggests that its memory is intensive while its classification is slow [20, 21].

Proposed Algorithm

We proposed, first of all, to classify human brain MR images and then to segment them according to a methodology that consists in three steps expandable in four during the segmentation: feature extraction, feature selection, classification, and segmentation.

For each image, we extract 26 features using WT-SGLDM. To check the performance of the proposed method, nine additional features were extracted. A total of 44 features were thus obtained. Then, we employ SVM classifier, especially RBF kernel as an effective option for kernel function [26-28]. The choice of RBF is explained by the fact that it can classify multi-dimensional data unlike a linear kernel function and that it has fewer parameters to set than a polynomial kernel. The SA-GA parameters served to reduce the number of extracted features as described in Table 1.

**Table 1.** Parameters related to the SA-GA algorithm

Parameter	Value/Method
Initial Temperature	$T0=75$
Temperature change	$Ti+1=0.09 \times Ti$
Number of iterations for each temperature	50
Selectivity function	Fitness
SA Stop condition	$Tstop=0.01$
Generation number	100
Size of initial population	Size of the solution obtained by SA
Selection Method	Tournoi
Probability of crossover	$Pc=0.9$
Mutation probability	$Pm=0.1$
Crossing method	Crossing to a point

AnSA-GA hybrid algorithm is proposed to efficiently select the optimal feature subsets. It relies on simulated annealing, a genetic algorithm, a greedy search algorithm and a support vector machine. Our hybrid approach is characterized by avoiding being trapped in a local minimum of SA, a high GA crossover operator convergence. At the same time, it guarantees a support vector machine (SVM) with high computational efficiency and a robust local greedy search algorithm.

Interestingly, our SA-GA approach selects feature sets without recourse to the filter steps. It couples the mutation-based search SA algorithm with good global searchability and the GA capacity to implement both the crossover and the mutation operations. Therefore, GA succeeds in solving the convergence issue, but the low fixed mutation rate and the crossover often trap the search in a local minimum. Furthermore, both SA and GA have weak local search capabilities. In the meantime, the greedy algorithms have good local searchability, but they lack strength in their global search.

SA-GA performs three search stages.

- **Stage 1:** SA-GA employs an SA to avoid the risk of a random bad choice of the initial population as shown in Fig.1. This figure illustrates the difficulties of bad and random initialization of generating the initial population by GA. The first case demonstrates the risk of deviating from the desired optimum because of a random bad choice of initial population. By contrast, in the second case, the initialization is better thanks to SA. Hence, SA generates an initial population for the GA better than the population generated randomly by GA. This process results in a better exploration and exploitation of search space. As SA is a global search algorithm, it guarantees the convergence to a global optimum.

  Due to a very high temperature, SA tolerates new solutions; which brings about a near random search through the search space. However, at a low temperature close to zero, it only accepts improvements.
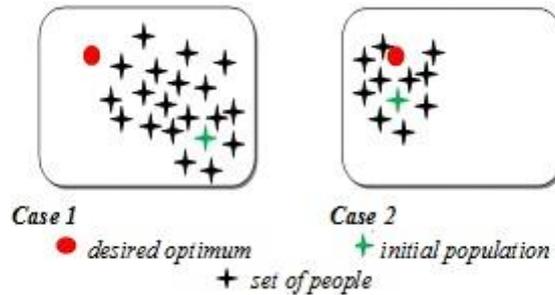
Figure 1. One kernel Exploration of research space by genetic algorithm

- **Stage 2:** Our proposed algorithm (SA-GA) performs optimizations using a GA. We set the GA population at 100. SA detects the best solutions and makes up the initial population. Crossover in GA is aimed at forming new and hopefully better solutions by exchanging information between pairs of good solutions. The crossover operator facilitates rapid convergence to a good solution. Thanks to the mutation operator, new genes are introduced into the population and genetic diversity is retained.
- **Stage 3:** The greedy algorithm locally searches the k-best solutions provided by SA and GA and chooses the best neighbors who are defined regarding the fitness function. As computational efficiency is essential, we employ a fast and robust supervised learning algorithm (SVM) to analyze data, recognize patterns and assess candidate solutions.

In order for SVM to perform effectively, the feature selection method should be reliable concerning discarding noisy, irrelevant and redundant data and preserve the discriminating power of data. Without feature selection, SVM input space is ample and disturbed; which lowers the SVM performance. SVM accuracy rate depends on the quality of the feature's dataset, and other factors including the kernel function and the two parameters C and γ as well.

SVM provides the optimal solution depending on several kernel functions and the most frequently used functions are the polynomial kernel, sigmoid kernel, and radial basis kernel function (RBF) [22, 23]. Our study employs the RBF kernel function to discover the optimal solution. The RBF classifies multi-dimensional data based on fewer parameters than a polynomial kernel. Another significant merit of SVM is that its performance is significantly different from other kernel functions.

In RBF, C and γ should be appropriately set. The C parameter refers to the penalty cost. The value of C influences the classification outcome.

The γ parameter affects the outcome more than C, as the partitioning outcome in the feature space depends on its value. The classification outcomesare inappropriate if the parameter values are not properly set [24].

Hence, good global search capability, rapid convergence to a near optimal solution, along with excellent local search ability and high computational efficiency are achieved through our proposed algorithm (SA-GA).

Computing Fitness of Feature Subsets

The estimation of all the features is based on the fitness function in equation (2). A fitness value is used to measure the 'fitness' of a feature to a population. The initial genetic process population encompasses the best solutions detected by SA. GA and high fitness ones discard low fitness. In our algorithm (SA-GA), SVM evaluates candidate feature subset solutions. Before this step, each feature is scaled between 0 and 1 for normalization purposes. A 5-fold cross validation was necessary to estimate the SVM classifier testing accuracy. Solution fitting evolves in parallel with the accuracy. In case of equal accuracy rate of two solutions, the solution that relies on fewer features wins.

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \times 100\,\% \tag{1}$$

where TP, TN, FP and FN are as defined in Table 2:

Table 2. Contingency table

| Case	Classifier	Reality
False Positive (FP)	Tumor	Normal
False Negative (FN)	Normal	Tumor
True Positive (TP)	Tumor	Tumor
True Negative (TN)	Normal	Normal

$$fitness = W_A \times Accuracy + W_{nb} \times \frac{1}{N} \tag{2}$$

where W_A is the weight of accuracy and W_{nb} is the weight of N feature participated in classification where $N \neq 0$.

Our paper contributes to ensuring to enhance the already obtained classification and segmentation results by using a hybrid SA-GA algorithm in the feature selection phase. Our contribution also consists in finding an optimal texture feature set that classifies normal brain, benign tumor, and malignant tumor by using block classification to segment the different types of a tumor using SVM.

Data

The proposed methodology is validated on two brain tumor datasets. The first were freely downloaded from the Harvard Medical School brain atlas [25]. The data set for this task comprises a total of 83 transaxial images: 29 images are normal, 22 malignant tumors suffering from a low-grade Glioma, Meningioma and 32 benign tumors suffering from a Bronchogenic Carcinoma, Glioblastoma multiform, Sarcoma, and Grade IV tumors. All images were considered to belong to seven persons (four men and three women aged between 22 and 81 years). The second, following the experimental setup in [30], we randomly split the 3064 T1-weighted contrast-enhanced images from 233 patients, with 2134 malignant tumors suffering from a meningioma (708 slices), glioma (1426 slices), and 930 benign tumors suffering from a pituitary tumor (930 slices), into 5 subsets of roughly equal size. Partitioning according to the patient ensures that slices from the same patient will not simultaneously appear in the training set and test set. For all the experiments, fivefold cross-validation is used. In fivefold cross-validation, one subset is sequentially used as the test set (query images), whereas the remaining four subsets are used as training sets. We tested our classification and segmentation algorithms for several normal brains and pathological brain MR images from the Harvard Medical School brain atlas (Fig. 2) and Nanfang Hospital and Tianjin Medical University training data (Fig. 3).
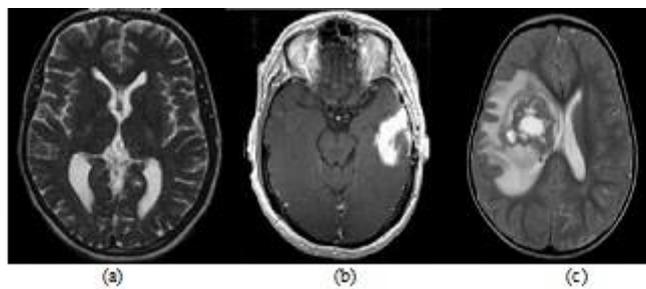
Figure 2. Sample of brain MRIs from the Harvard Medical School brain atlas: **(a)** Normal brain; **(b)** Pituitary tumor; **(c)** Glioma tumor
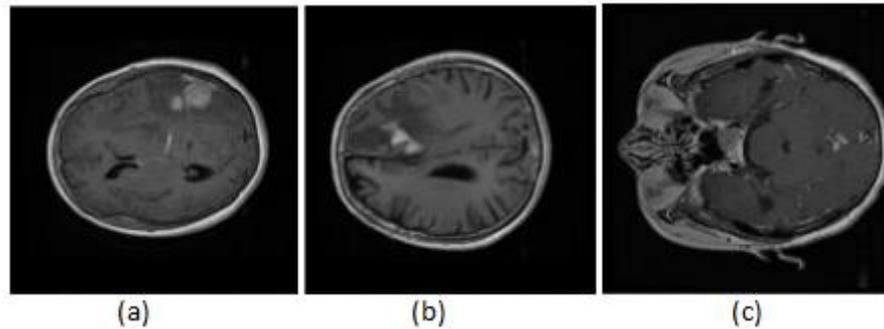
Figure 3. Brain images from the Nanfang Hospital and Tianjin Medical University training dataset: **(a)** Meningioma, **(b)** glioma, **(c)** pituitary tumors

The suggested classification algorithm was designed to segregate normal from benign or malignant brain tumor MR images. As tumor regions may be scattered all over the image, we applied pixel classification rather than classical segmentation methods. Among all pixel classification methods, the Support Vector Machines (SVM) was adopted as a segmentation method.

Results and Discussion

We suggested, first of all, to classify human brain MR images and then to segment them. For each image, we extract 26 features using WT-SGLDM. To check the performance of the proposed method, nine additional features were extracted. A total of 44 features was thus obtained. Then, we employed the SVM classifier by using RBF kernel.

The best SA-GA selected features during the execution are illustrated in Table 3. The classification performance of 95.65% was obtained with 4 of the whole available features, thus classifying normal and pathological brains such as benign or malignant tumor employing the least features and reducing the classifier cost.

Table 3. Results of feature selection performed by SA-GA

Feature selection	Feature set	Classifier accuracy (%)
SA-GA	**7 features:** Mean of contrast (M.CON), Mean for Information measure of correlation I (M.IMC I), Mean dissimilarity (M.DISS), Range of correlation (R.CORR), Range of variance difference (R.DVAR), Range Information measure of correlation I (R.IMC I), Range Information measure of correlation II (R.IMC II).	95.65
	5 features: Mean of contrast (M.CON), Mean dissimilarity (M.DISS), Mean of homogeneity (M.HOMO), Range of correlation (R.CORR), Range Information measure of correlation II (R.IMC II).	95.65
	4 features: Mean of contrast (M.CON), Mean of energy mat (M.ENER MAT), Mean of homogeneity (M.HOMO), Range of correlation (R.CORR).	95.65

Using only four features: mean of contrast, mean of homogeneity, mean of energy and range of correlation, the classification was obtained at an accuracy rate of 95.65. The SA-GA algorithm selected features according to the appearance of images of the tumors database. The area of the tumor in abnormal brain images was brighter, with a regular color distribution. The selection of the contrast and the correlation features as descriptive characteristics of the tumor was thus significant.

Tables 4 and 5 show the classification rates for performing the proposed approach on hybrid optimization (SA/GA) from the Harvard Medical School brain atlas and the second dataset brain tumor from Nanfang Hospital and Tianjin Medical University respectively, by using the most common kernel functions including linear, polynomial of degree and RBF.

Classification accuracy varies from 92.69±1.2 to 94.41±1.2 and 95.69±1.2 to 96.78±1.2 from the one brain atlas and the second dataset brain tumor respectively, with polynomial and radial basis function. Both tools could achieve satisfactory classification results for brain tumor but we prefer the application of RBF. In this case, the classification accuracy varies from 93.25 to 95.65 % and 95.69 to 98.08 % in the mean standard deviation format (Mean SD) of 94.45±1.2 % and 96.89±1.2 from the one brain atlas and the second dataset brain tumor respectively.

Note that increasing numbers of learning images can succeed in our approach to convert and lead to good results. This makes our approach an efficient clinical image analysis tool for doctors or radiologists to classify MRI tumor and to further obtain MRI tumor location.

Profile regularity and repetition in a signal can be detected thanks to the correlation. Furthermore, color distribution at a tumor area was regular as the values are relatively close. Homogeneity is then chosen as a descriptive feature of the tumor according to these aspects. In particular, the homogeneity has an opposite behavior of the contrast. The homogeneity characteristic is related to the homogeneous texture regions. As for the energy feature, it is responsible for extracting a regular image contour.

Table 4. Classification accuracies based on hybrid optimization (SA/GA) using the Harvard Medical School brain training data

Kernel used	Total number of images	Number of images in the training			Number of images in the testing			Images misclassified	Classification accuracy±SD(%)
		N	B	M	N	B	M		
Linear	83	12	9	16	29	18	36	7±1	92.69±1.2
Polynomial	83	12	9	16	29	18	36	5±1	94.41±1.2
Radial basis function	83	12	9	16	29	18	36	5±1	94.45±1.2

N is a normal, B is a Benign and M is a Malign

Table 5. Classification accuracies based on hybrid optimization (SA/GA) using Nanfang Hospital and Tianjin Medical University training data

| Kernel used	Total number of images	Number of images in the training		Number of images in the testing		Images misclassified	Classification accuracy ± SD (%)
		B	M	B	M		
Linear	3064	800	1900	130	234	4±1	95.69±1.2
Polynomial	3064	800	1900	130	234	3±1	96.78±1.2
Radial basis function	3064	800	1900	130	234	3±1	96.89±1.2

B is a Benign, M is a Malign

This comparison gives importance to three features more than others because they are selected several times by simulated annealing, genetic algorithm and the SA-GA process (Table 6). These three features are Mean of contrast (M.CON), Mean of homogeneity (M.HOMO) Range of correlation (R.CORR). They lead us to extract the abnormal areas of a brain MRI image and specifically distinguish tumors from noise in the image, which facilitates and optimizes the classification and segmentation system. Thus, contrast, homogeneity, and correlation features present the most distinctive features of a tumor, since they combine light distribution (correlation and contrast) with extracting texture-homogeneous zones (homogeneity). Therefore, the SA-GA approach achieves better results than SA and GA in isolation regarding reducing effective and reliable data.

Table 6. Comparison of results obtained by the genetic algorithm, simulated annealing, and SA-GA algorithm

Feature set	GA[26-28]			SA [29]			SA-GA		
	7	6	5	9	8	7	7	5	4
Mean of energy (M.ASM)									
Mean of contrast (M.CON)		×	×			×	×	×	×
Mean of correlation (M.CORR)	×								
Mean of variance (M.VAR)									
Mean of inverse difference moment (M.IDM)				×					
Mean of entropie (M.ENT)									
Mean of sum average (M.SAVG)			×						
Mean of sum variance (M.SVAR)			×			×			
Mean of sum entropie (M.SENT)					×				
Mean of difference variance (M.DVAR)	×								
Mean of difference entropie (M.DENT)									
Mean of Information measure of correlation I (M.IMC I)	×	×					×		
Mean of Information measure of correlation II (M.IMC II)									
Mean of maximal correlation coefficient (M.MAX CORR)				×					
Mean of correlation mat (M.CORR MAT)					×				
Mean of cluster Prominence (M.CP)				×					
Mean of cluster Shade (M.CS)									
Mean of dissimilarity (M.DISS)							×	×	
Mean of energy mat (M.ENER MAT)					×	×			×
Mean of homogeneity (M.HOMO)		×	×	×	×	×		×	×
Mean of Maximum probability (M.MAX PROB)	×								
Mean of inverse difference moment (M.IDM)	×	×							
Range of energy (R.ASM)									
Range of contrast (R.CON)	×			×					
Range of corrélation (R.CORR)		×	×			×	×	×	×
Range of variance (R.VAR)									
Rang of inverse difference moment (R.IDM)						×			
Range of entropy (R.ENT)									
Range of average sum (R.SAVG)									
Range of sum variance (R.SVAR)				×					
Range of sum entropy (R.SENT)									
Range of difference variance (R.DVAR)				×	×		×		
Range of difference entropy (R.DENT)									
Range of Information measure of correlation I (R.IMC I)					×		×		
Range of Information measure of correlation II (R.IMC II)							×	×	
Range of maximal correlation coefficient (R.MAX CORR)					×				
Range of correlation mat (R.CORR MAT)					×	×			
Range of cluster Proeminence (R.CP)									
Range of cluster Shade (R.CS)									
Range of dissimilarity (R.DISS)									
Range of energy mat (R.ENER MAT)				×					
Range of homogeneity (R.HOMO)	×	×		×					
Range of Maximum probability (R.MAX PROB)									
Range of inverse normalized difference (R.IDN)									

GA = Genetic Algorithms; SA = Simulated Annealing

Fig. 4 shows that the simulated annealing reached the maximum accuracy (99.99%) by 10, 80 and 100 generations. Meanwhile, this maximum precision is achieved by the genetic algorithm for more generations: 10, 30, 40, 70 and 100. On the other hand, our proposed SA-GA approach achieves a high degree of accuracy almost similar to that achieved by SA. We conclude that this is due to the precision of the initial search zone of the proposed SA-GA approach.
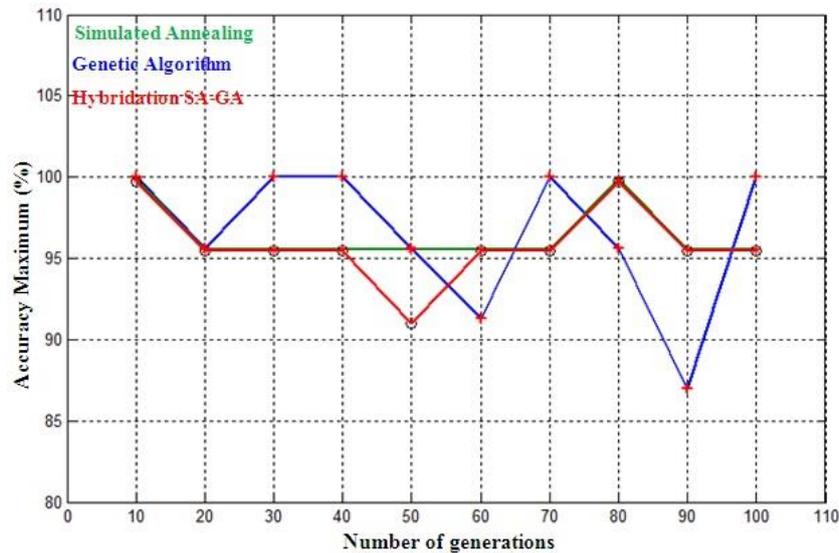

Figure 4. Accuracy evolution according to the number of generations through the implementation of SA, GA, and SA-GA

The results of convergence to an optimal solution for the three implemented methods indicate that the genetic algorithm achieves 99.99% accuracy by producing different numbers of generations. However, more stability is obtained by simulated annealing, as the maximum accuracy obtained for its solutions varies between only two values regardless of the number of generations. By adding simulated annealing before the execution of the genetic algorithm, we note a significant influence on the result of the proposed SA-GA approach revealed by the stabilization of the results obtained.

According to Fig.4, we note that the highest accuracy achieved by simulated annealing, genetic algorithm, and our proposed approach SA-GA does not depend on the number of generations produced since we can have an accuracy of 99.99 % by the three methods even with only ten generations. We can see then that the calculation accuracy depends only on the fitness function. For this, the choice of the selectivity function is essential for optimal solutions.

According to Fig.5, the three algorithms: simulated annealing, genetic algorithm, and our hybrid SA-GA algorithm need more computing time which increases the number of generations. The difference regarding computing time between simulated annealing and genetic algorithm is exceedingly remarkable. The latter spends too much computing time. Thus, the necessary time to produce a solution by genetic algorithm for ten generations is three times greater than the time required by simulated annealing for 100 generations. We also note that SA-GA pace computing time is almost near the pace of the genetic algorithm according to the number of generations. Thus, the influence of the genetic algorithm remains significant since the difference between the two curves which show these two algorithms is considered low (genetic algorithm needs further 162 min than our proposed approach SA-GA for performing 100 generations).

Simulated annealing however slightly decreased the computing time curve achieved by SA-GA compared to the genetic algorithm curve thanks to the precision and the reduction of the search area in the initialization phase. Therefore, optimization of the genetic algorithm for the selection of the most relevant features by simulated annealing is evident in Fig.5.
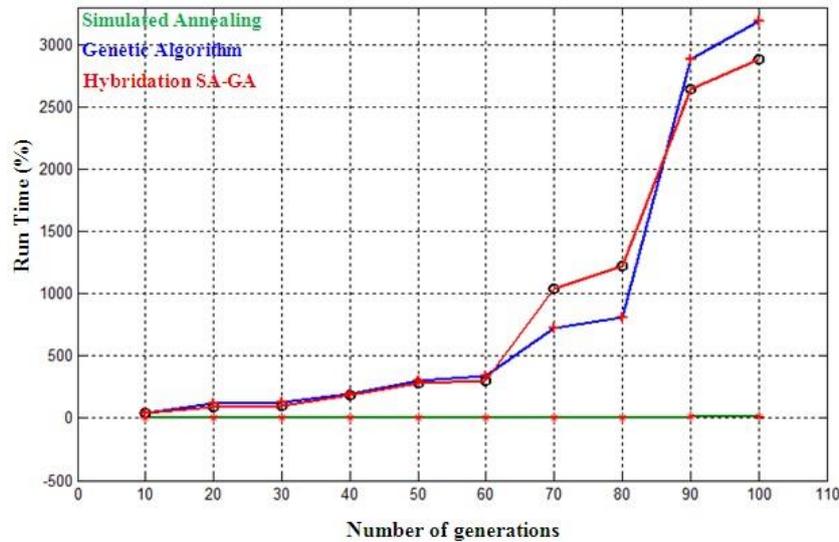
Figure 5. Computing time depending on the number of generations through the implementation of SA, GA, and SA-GA

In our method, the tumor region, which is a heterogeneous one, is segmented with high accuracy using only four optimal texture features. The results of our segmentation method using hybrid SA-GA algorithm in the feature selection phase and depicted alongside the ground truth (a manually selected image) are shown in Fig.6 and 7. According to the results shown in Table 7 our method can be used to detect benign and malignant tumors. Moreover, the segmentation accuracy applied to glioma tumors is high (97.82%±0.74) compared with a pituitary tumor because it considers all pixels concurrently and it highlights heterogeneous regions as well as homogeneous regions within the malignant tumor.
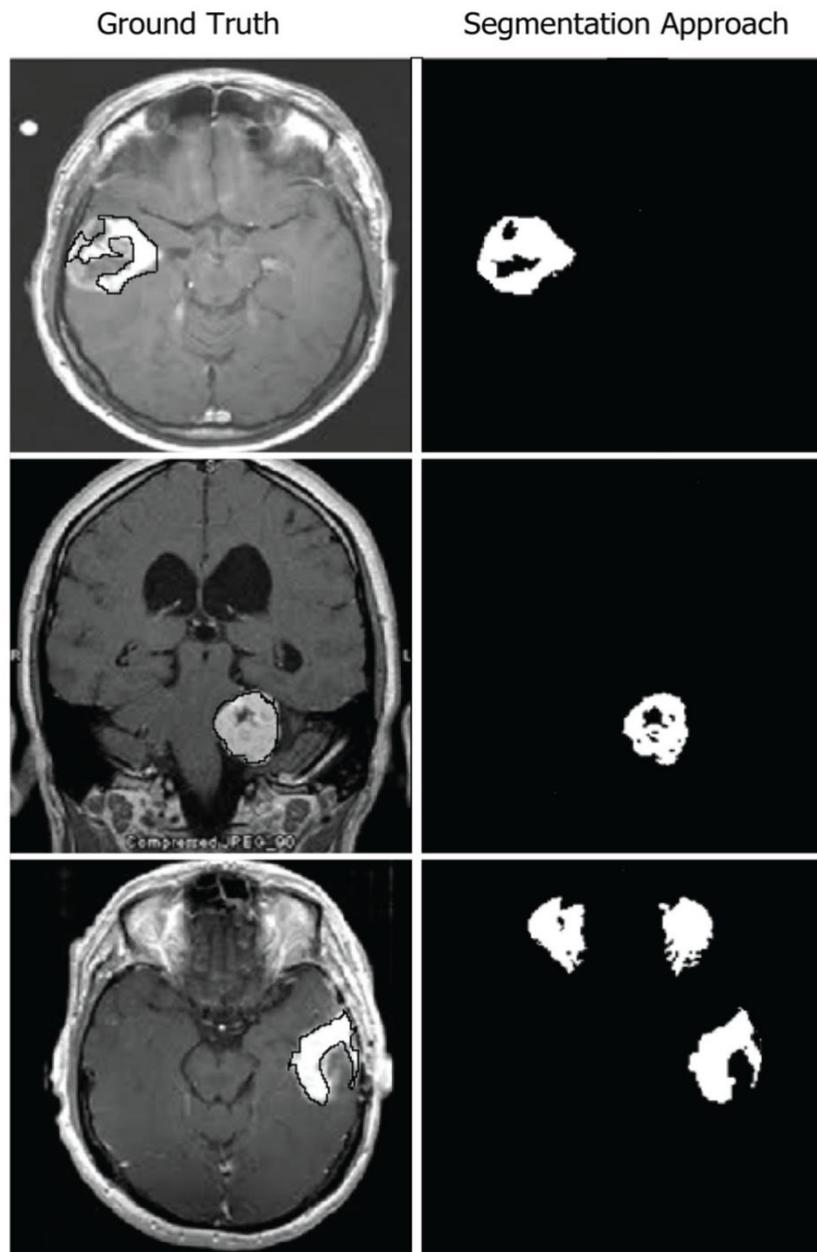
Table 7. Segmentation performance of the proposed method for different types of brain tumors based up training data Nanfang Hospital and Tianjin Medical University

Tumor type	Segmentation accuracy (mean ±std %)
Glioma	97.82±0.74
Pituitary tumor	95.12±3.21
Meningioma	87.77±3.04

In our study, we are also interested in the diagnosis of abnormalities of the brain tumor. The goal was to explore the use of our advanced technology and quantify its effectiveness in dealing with of segmentation issue. To achieve this goal, we compare our method with state-of-the-art methods.

While using the feature selection result of our hybrid SA-GA algorithm, we evaluate the segmentation performance of the proposed method for different types of tumors in this section. Table 7 summarizes the results. The segmentation performance of meningiomas is much lower than that of gliomas and pituitary tumors. One possible reason is the imbalance of data between different categories of tumors.

The accuracy rate achieved by our method (95.9%) illustrates that the latter outperforms the state-of-the-art methods. The accuracy rate already achieved by previous methods varies between 87.3% by Yang et al. [31] and 94.68% by Cheng et al. [34]. In all five methods, the same set of data is used with five-way cross-validation. The results of the segmentation of the four compared methods, taken directly from the corresponding original articles, as compared to our proposed method are presented in Table 8. The accuracy of our method significantly surpasses those of the other four methods.

Figure 6. Segmentation results on a brain MRI with benign tumor: A manually selected image (Ground Truth) and using optimal feature-based (SA-GA) segmentation

Table 8. The proposed method's segmentation results compared with the state-of-the-art method based up training data Nanfang Hospital and Tianjin Medical University (mean %)

Methods	Yang et al. [31]	Huang et al. [32]	Huang et al. [33]	Cheng et al. [34]	Proposed
Segmentation accuracy	87.3	91.0	91.8	94.68	**95.9**

Ground Truth   Segmentation Approach
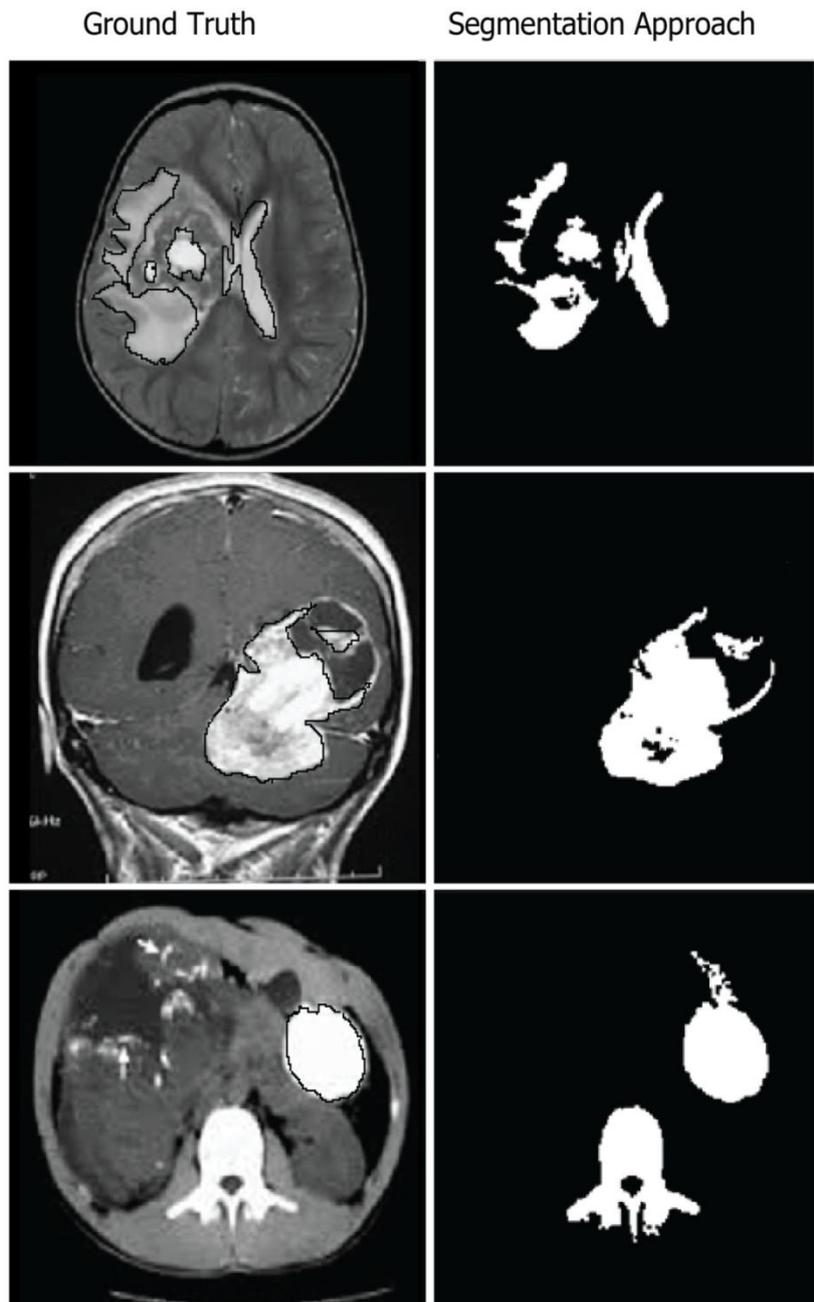
Figure 7. Segmentation results on a brain MRI with malignant tumor : A manually selected image (Ground Truth) and using optimal feature-based (SA-GA) segmentation

In this repository, we address the problem of segmenting brain tumor images in archives that have the same pathological type as the image of the query. Segmented images with diagnostic information can be used by radiologists to provide decision support. The success of image segmentation systems is based on good performance and adapted distance metrics. For example, using our hybrid SA-GA algorithm in the feature selection phase, we can achieve a segmentation accuracy of 95.90%.

In the selection phase of the most relevant features of the brain images, simulated annealing, genetic algorithm and SA-GA showed that the execution speed is highly dependent on the number of generations produced to have an optimal solution. The execution speed is a crucial criterion for classification system assessment as well as for segmentation. This makes our proposed approach important in clinical utility. It may help radiologists to detect brain tumors; to decrease missed cancerous case detection and to improve inter- and intra-reader variability. Thus it is useful for brain

cancer diagnosis by minimizing the time and effort necessary for diagnosis. However, this approach may suffer from two drawbacks: (1) some types of features may significantly dominate others in the selection phase; thus, the potential of all features cannot be fully exploited; (2) the resulting choice of irrelevant features will make the subsequent classification task computationally expensive. Our future perspective is to use the results of 2D images for 3D images segmentation as 2D images segmentation's error rate is lower than that of 3D.

Conclusions

The proposed hybrid SA-GA algorithm for selection of the optimal feature subsets from a large number of features is able serve as a tool for computer-aided diagnosis in discrimination of brain tumors on magnetic resonance images. The proposed SA-GA approach outperforms existing approaches regarding accuracy. Generally, the performance of isolated algorithms is seldom entirely satisfactory; however, a judicious combination can overcome the weaknesses of each. Our approach proved its efficiency in research applications, especially in computer-aided diagnosis and radiotherapy planning. The performance of the proposed approach provides a substantial improvement against four closely related methods, achieving a mean of 95.9%.

Conflict of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

1. Karnam G, Thirumala R. A Gui Based Hybrid Clustering Technique for Brain Tumor Segmentation and Validation in Mri Images. Aust J Basic & Appl Sci. 2016;10(8):84-96.
2. Saurabh S, Chauhan NC. Techniques for Detection and Analysis ofTumours from Brain MRI Images: A Review. Journal of Biomedical Engineering and Medical Imaging 2016;3(1):9-20.
3. Gopi K, Ramashri Th.Design and development of a computer aided diagnosis system for segmentation of brain tumor. ARPN Journal of Engineering and Applied Sciences 2015;10(22):213-224.
4. Grana M, Termenon M, Savio A, Gonzalez-Pinto A, Echeveste J, Perez JM, et al. Computer aided diagnosis system for Alzheimer disease using brain diffusion tensor imaging features selected by Pearson's correlation. Neuroscience Letters 2011;502(3):225-229.
5. Xuan Z, Xin G, Jiajun W, Hui Y, Zhiyong W, Zheru C.Eye tracking data guided feature selection for image classification. Pattern Recognition 2017;63(2):56-70.
6. Bischl B, Lang M, Kotthoff L, Schiffner J, Richter J, Studerus E, Casalicchio G, Jones Z M.mlr: Machine learning in R. Journal of Machine Learning Research 2016;17(170):1-5.
7. Dash M, Choi K, Scheuermann P, Liu H. Feature selection for clustering-a filter solution.In Proceedings of the Second IEEE International Conference on Data Mining, 2002, pp.115-122.

8. Deqing W, Hui Z, Rui L, Weifeng L, Datao W. t-Test feature selection approach based on term frequency for text categorization. Pattern Recognition Letters 2014;45(2):1-10.

9. Jin X, Xu A, Bie R, Guo P. Machine learning techniques and chi-square feature selection for cancer classification using SAGE gene expression profiles. Lecture Notes in Computer Science 2006;3916:106-115.

10. Liao C, Li S, Luo Z. Gene selection using Wilcoxon rank sum test and support vector machine for cancer. Lecture Notes in Computer Science 2007;4456:57-66.

11. Liwei W, Omar A M S. An improved Fuzzy Mutual Information Feature Selection for Classification Systems. 16th IEEE/ACIS International Conference on Computer and Information Science, 2017, pp. 119-124.

12. Chinnaswamy A, Srinivasan R. A comparative study of hybrid feature selection methods using correlation coefficient for microarray data. Journal of Network and Innovative Computing 2016;4:2160-2174.

13. Yijuan Lu, Cohen Ira, Xiang S Z, Tian Qi. Feature selection using principal feature analysis. In Proceedings of the 15th ACM international conference on Multimedia, 2007, pp. 301-304.

14. Chuan L, Wenyong W, Qiang Z, Xiaoming S, Martin K. A new feature selection method based on a validity index of feature subset. Pattern Recognition Letters 2017;92:1-8.

15. Ronen M, Jacob Z.Using simulated annealing to optimize feature selection problem in marketing applications. European Journal of Operational Research 2006;171:842-858.

16. Tan F, Fu X, Wang H, Zhang Y, Bourgeois A. A hybrid feature selection approach for microarray gene expression data. Lecture Notes in Computer Science 2006;3992:678-685.

17. Shamsul H, John Y, Herbert J, Mohammad MH, Giancarlo F, Michael B. A Hybrid Feature Selection With Ensemble Classification for Imbalanced Healthcare Data: A Case Study for Brain Tumor Diagnosis. IEEE Access 2016;4:9145-9154.

18. Fatourechi M, Birch G, Ward RK. Application of a hybrid wavelet feature selection method in the design of a self-paced brain interface system. Journal of Neuro Engineering and Rehabilitation 2007;4(11):1-13.

19. Huang J, Cai Y, Xu X. A wrapper for feature selection based on mutual information.In 18th International Conference on Pattern Recognition 2006, pp. 618-621.

20. Buttrey SE, Karo C. Using k-nearest-neighbor classification in the leaves of a tree. Computational Statistics & Data Analysis 2002;40:27-37.

21. Kuncheva LL. Editing for the k-nearest neighbors rule by a genetic algorithm. Pattern Recognition Letters 1995;16:809-814.

22. Hsuan T L, Chih JL.A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods, Technical Report, University of National Taiwan, Department of Computer Science and Information Engineering 2003, 1-32.

23. Muller KR, Mike S, Ratsch G, Tsuda K, Scholkopf B.An introduction to kernel-based learning algorithms. IEEE Trans. Neural Netw 2001;12:181-201.

24. Pardo M, Sberveglieri G. Classification of electronic nose data with support vector machines. Sens. Actuators B: Chem 2005;107:730-737.

25. Harvard medical school. 1999. Available from: http://www.med.harvard.edu/aanlib/home.html.

26. Kharrat A, Gasmi K, Ben Messaoud M, Benamrane N, Abid M. Medical Image Classification Using an Optimal Feature Extraction Algorithm and a Supervised Classifier Technique. International Journal of Software Science and Computational Intelligence 2011;3(2):19-33.

27. Kharrat A, Gasmi K, Ben Messaoud M, Benamrane N, Abid M. Automated Classification of Magnetic Resonance Brain Images Using Wavelet Genetic Algorithm and Support Vector Machine. Proc. 9th IEEE Int. Conf. on Cognitive Informatics 2010, pp. 369-374.

28. Kharrat A, Abid M. Toward Efficient Segmentation of Brain Tumors Based on Support Vector Machine Classifier Through Optimized RBF Kernel Parameters and Optimal Texture Features. International Journal of Cognitive Informatics and Natural Intelligence 2014;8(2):15-33.

29. Kharrat A, Ben Halima M, Ben Ayed M. MRI Brain Tumor Classification using Support Vector Machines and Meta-Heuristic Method. In: 15th International Conference on Intelligent Systems Design and Applications (ISDA). IEEE, 2015, pp.446-451.

30. Brain tumor dataset from Ethics Committees of Nanfang Hospital and General Hospital, Tianjin Medical University. Dataset posted on 2 April 2017. Available from: https://figshare.com/articles/brain_tumor_dataset/1512427.

31. Yang W, Feng Q, Yu M, Lu Z, Gao Y, Xu Y, et al.Content-based retrieval of brain tumor in contrast-enhanced MRI images using tumor margin information and learned distance metric. Med Phys. 2012;39:6929-42.

32. Huang M, Yang W, Yu M, Lu Z, Feng Q, Chen W. Retrieval of brain tumors with region-specific bag-ofvisual-words representations in contrast-enhanced MRI images. Computational and Mathematical Methods in Medicine 2012;280538:1-17.

33. Huang M, Yang W, Wu Y, Jiang J, Gao Y, Chen Y, et al. Content-Based Image Retrieval Using Spatial Layout Information in Brain Tumor T1-Weighted Contrast-Enhanced MR Images. PLoS One 2014;9:e102754.

34. Cheng J, Yang W, Huang M, Huang W, Jiang J, Zhou Y, Yang R, Zhao J, Feng Y, Feng Q, Chen W. Retrieval of Brain Tumors by Adaptive Spatial Pooling and Fisher Vector Representation. PLoS One 2016;11(6):e0157112.