

Classification of Brain MRI Tumor Images Using Fuzzy C Means Clustering With Firefly Algorithms Optimized Support Vector Machine

M.R.Thiyagupriyadharsan^{1*} and Dr.S.Suja²

¹Assistant Professor, ECE Department, Kathir College of Engineering, Coimbatore, 641062, India

²Associate Professor, EEE Department, Coimbatore Institute of Technology, 641014, India

¹Email: thiyagu.priyadharsan@gmail.com

²Email: sujascit@gmail.com

Abstract: In the contemporary world, many dangerous disease which are affecting human beings and new pandemic disease is also raising alarm to have an effective health care system. In this aspect the technology plays a major role in improving and optimizing the health care system. The diagnostic is done by taking blood test, urine test, and medical imaging like X-ray, CT scan, Ultrasound scan and MRI scan system. Among these, the paper focus will be emphasized on MRI imaging in identifying the brain tumor using image processing. In the proposed work the fuzzy C means(FCM) algorithm along with firefly algorithm optimized support vector machine (SVM) are used to classify the MRI brain tumor images. The results of these works are compared using the performance metrics such as accuracy, sensitivity, specificity and precision. The proposed method gives best results for the classification of MRI brain tumor images.

Keywords: Fuzzy C means (FCM), Firefly algorithm, Support Vector Machine (SVM)

1 Introduction

Brain tumor is one of the most mortal type of cancer infections. The importance comes from the fact that early detection increases significantly the possibility of curing the disease and saving the life of patients. Recently, the treatments of cancer have greatly developed especially in the early stages of infection. Survival possibilities are very high for those patients receiving early treatments compared to those people who don't have this chance in the early stages of sickness.

Brain tumor is a mass or accumulation of biological cells in the brain. Generally, tumors of the brain or any tumor can be classified into two types of tumor. The first is called benign tumor or noncancerous tumor; whereas the second is very dangerous and cancerous that is said to be malignant tumor. The growth of these two types of tumor inside the skull forces the brain and can be very harmful for life of patient. [1]

In this paper, the work has been carried out by fuzzy C means algorithm to cluster the images, then the features are extracted and given to support vector machine optimized by swarm intelligence based algorithms like moth flame, firefly and dragonfly.[2]

The paper is organized as the following sections: section 2 deals with the fuzzy C means algorithm. The section 3 discuss about the swarm intelligence Firefly algorithms. Section 4 briefs about the support vector machine (SVM). Section 5 elaborates about the hybrid swarm intelligence algorithms. Section 6 gives about the obtained results and its discussion.

2 Fuzzy C means algorithm

Fuzzy C means (FCM) clustering is a soft clustering algorithm. It overcomes the drawback of the K-means clustering which is a hard clustering algorithm. FCM can accommodate an object in more than one cluster. It is an unsupervised algorithm. [3]

The FCM is given by

$$J = \sum_{i=1}^N \sum_{j=1}^C U_{ij}^m \|y_i - C_j\|^2 \quad (1)$$

Where U_{ij} represents fuzzy membership function of i^{th} object in j^{th} cluster.

C_j represents the cluster center

m represents degree of fuzziness.

N represents the number of data elements in the image.

The updating of membership function U_{ij} and the cluster center C_j is as follows

$$U_{ij} = \frac{1}{\sum_{k=1}^C \left(\frac{\|y_i - C_j\|}{\|y_i - C_k\|} \right)^{\frac{2}{f-1}}} \quad (2)$$

$$C_j = \frac{\sum_{i=1}^N U_{ij}^f \cdot y_i}{\sum_{i=1}^N U_{ij}^f} \quad (3)$$

The iteration is terminated, when $\max_{ij} \{ |u_{ij}^{(k+1)} - u_{ij}^{(k)}| \} < \delta$,

where δ is a termination criterion between 0 and 1,

f represents the weighting exponent that represents the degree of fuzziness

k represents the iteration count.

3 Firefly algorithm

Firefly Algorithm (FA) was a natural inspired algorithm developed by Xin-She Yang in late 2007 and 2008 at Cambridge University, it was based on the flashing patterns and behaviour of fireflies. FA uses the following three rules:

- Fireflies are unisex so that one firefly will be attracted to other fireflies regardless of their sex.
- The attractiveness is proportional to the brightness, and they both decrease as their distance increases. Thus for any two flashing fireflies, the less brighter one will move towards the brighter one. If there is no brighter one than a particular firefly, it will move randomly.
- The brightness of a firefly is determined by the landscape of the objective function. As a firefly's attractiveness is proportional to the light intensity seen by adjacent fireflies,

The light intensity varies with the equation,

$$I = I_0 e^{-\gamma r^2} \quad (4)$$

where I_0 is the light intensity at $r = 0$.

The variation of attractiveness β varies with the distance r

$$\beta = \beta_0 e^{-\gamma r^2} \quad (5)$$

where β_0 is the attractiveness at $r = 0$.

γ is the media light absorption coefficient.

and r is the Cartesian distance, between any two fireflies i and j at x_i and x_j in d dimension future space respectively.

$$r = \sqrt{\sum_{k=1}^d (x_{i,k} - x_{j,k})^2}$$

The movement of a firefly i is attracted to another more attractive (brighter) firefly j is determined by

$$x_i^{t+1} = x_i^t + \beta_0 e^{-\gamma r_{ij}^2} (x_j^t - x_i^t) + \alpha_t \epsilon_i^t \quad (6)$$

where the second term is due to the attraction. The third term is randomization with $\alpha_t \epsilon_i^t$ being the randomization parameter. If $\beta_0 = 0$, it becomes a simple random walk. On the other hand, if $\gamma = 0$, it reduces to a variant of particle swarm optimization.

4 Support Vector Machine (SVM)

Support Vector Machine algorithm is a supervised machine learning algorithm mainly used for classification and regression analysis. It was developed by Vapnik. In SVM, the training data are used to train the classification model of the brain tumor, and then this model is used to classify an unknown brain tumor image. SVM has penalty parameter, and kernel parameters that have a great influence on the performance of SVM.[5]

If there is N linearly separable samples given by, $X=\{x_1, x_2, \dots, x_N\}$, where x_i is the i^{th} training sample and each sample has d attributes and belongs to any of the two classes $y_i \in \{\pm 1\}$ shown in figure 6.6. The line $w^T x + b = 0$, represents the decision boundary of the two classes, where w denotes a weight vector, b denotes bias and x is the training sample. This line divides the space into two hyperplanes H_1 and H_2 . It is shown in figure 6.5, the encircled one are the support vectors. The hyperplane H_1 is given by $w^T x_i + b = 1$ for positive class and H_2 is given by $w^T x_i + b = -1$ for negative class, and these two equations will give $y_i (w^T x_i + b) - 1 \geq 0 \forall i = 1, 2, \dots, N$.

5 Proposed work

The proposed work is done in two stages. In the first stage the brain tumor images are clustered using fuzzy C means clustering. Then the features are extracted and given to the second stage. The second stage is proceeded by the proposed work FCM based Hybrid Firefly optimized SVM. The work flow is depicted in the flow chart shown below figure 1.

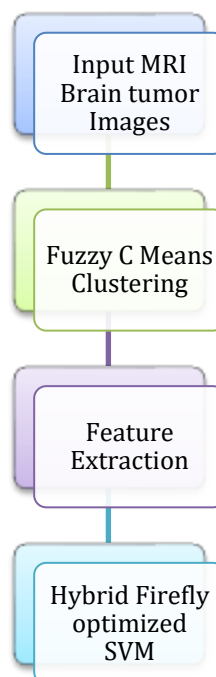


Figure 1. The proposed work flow**5.1 Hybrid Firefly optimized SVM algorithm**

The proposed algorithm can be briefly described by the following steps

Step 1 Set the values of FA parameters such as:

- The number of Fireflies ;
- The maximum number of iterations;
- The upper and the lower bounds of C and σ .

Step 2 The values of l, r , and β and the values of SVM parameters (C and σ) are initialized.

Step 3 The SVM model is trained using the training set and compute the fitness value of every firefly. The ratio between the number of misclassified samples (N_e) to the total number of testing samples (N).

The fitness function is in the form of:

$$F = \frac{N_e}{N}$$

Step 4 The values of r and β are updated.

Step 5 The fireflies are ranked according to the attractiveness.

Step 6 The firefly are sorted.

Step 7 The new positions are corrected taking into account the upper and lower values of variables C and σ .

Step 8 The stopping criterion is checked: if its criterion is achieved, go to the Step 9. Otherwise, loop to Step 3.

Step 9 The best position for all individuals comprising the optimized SVM parameters is selected, and then the SVM model was tested and evaluated.

6 Results and Discussions**6.1 Dataset Used**

The dataset used here is got from the open source figshare brain dataset. This brain tumor dataset containing 3064 T1-weighted contrast-enhanced images from 233 patients with three kinds of brain tumor: meningioma (708 slices), glioma (1426 slices), and pituitary tumor (930 slices).[6]

6.2 Performance Metrics

The metrics that are included for calculating the performance of classification of brain tumor are accuracy, precision, sensitivity, and specificity. These metrics are calculated using:

Specificity is the metric that is used to predict the true negatives of available category in a model. It is also known as True Negative Rate and calculated as,

$$\text{Specificity} = \frac{TN}{FP+TN}$$

Precision is the metric which determines how often it correctly predicts true. It is also known as Positive Predictive Value and is calculated as,

$$\text{Precision} = \frac{TP}{TP+FP}$$

Accuracy is the ability to differentiate between true positive and true negative correctly.

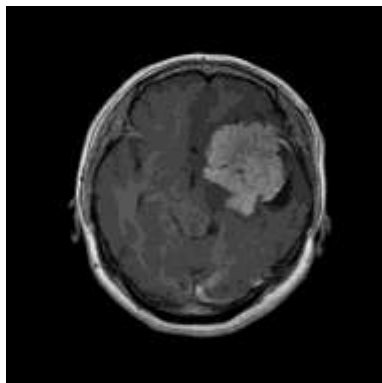
$$\text{Accuracy} = \frac{TP+TN}{N}$$

Sensitivity is a metric that is used to predict the true positive of the available category in a model. It is also known as the True Positive rate or Recall. It is calculated as,

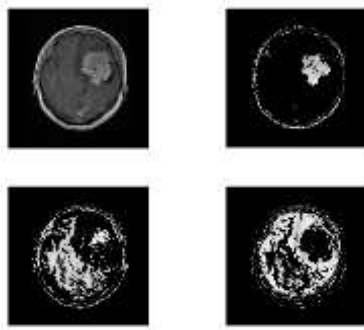
$$\text{Sensitivity} = \frac{TP}{TP+FN}$$

6.3 Results and Discussions

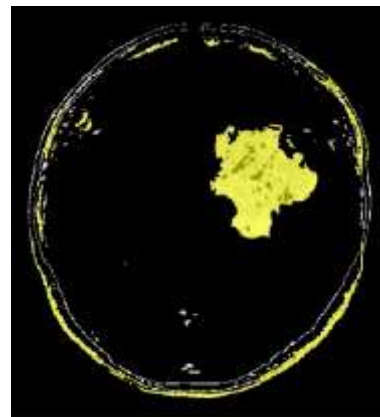
In the SVM 60% training and 40% testing is done in the dataset. The parameters C range from 0.01 to 35000 and σ ranges from 0.01 to 100. The optimal value of C lies between 970 to 1000 and σ is 0.1. The platform used to run the algorithm is Matlab 2018a. The parameters included for calculating the performance are accuracy, precision, sensitivity, and specificity.



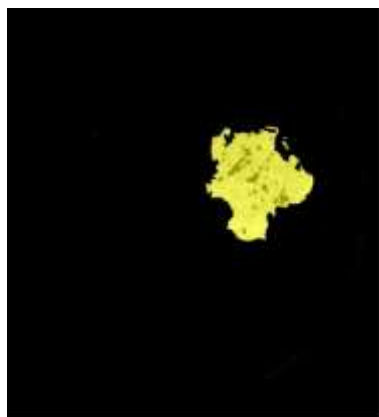
a) Input image



b) Clustered output



c) Selected cluster output



d) Tumor recognition

Figure 2: Intermediate results of the proposed algorithm for Glioma tumor

The Figure 2 shows the intermediate results of the proposed algorithm for the glioma tumor. The brain tumor image is clustered using FCM algorithm. From the clustered data, a cluster is selected which is having the tumor region located. Then morphological and dilation operation is performed to get the perfect tumor image.

Table 1: Performance Evaluation of Hybrid Firefly optimized SVM model

Metrics	Fuzzy C means with hybrid Firefly optimized SVM		
	Meningioma	Glioma	Pituitary
Accuracy	96.59%	96.59%	96.59%
Sensitivity	97.96%	98.46%	98.30%
Specitficity	99.15%	99.11%	99.12%
Precision	94.14%	95.52%	95.09%

The result shows that the accuracy of the FCM with hybrid firefly optimized SVM is 96.59 % for all the three types of tumor. It yields good result in the metrics sensitivity. The specificity that is the negative rate is significantly higher for the three types of tumor. The metrics precision also yields good result.

7 Conclusion

In this work, the fuzzy C means algorithm and firefly algorithm optimized with support vector machine (SVM) is simulated and the results are obtained. First the pre-processing of image is done. The fuzzy C means algorithm is used to cluster the images and the features are extracted. Finally the values are given to the firefly optimized support vector machine (SVM). The results shows that the fuzzy C means clustering with firefly optimized SVM gives better results. It shows significant improvement in all the four metrics such as accuracy, precision, sensitivity and specificity.

Funding Statement: The author(s) received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] Jin Liu, Min Li, Jianxin Wang, Fangxiang Wu, Tianming Liu, and Yi Pan. "A Survey of MRI-Based Brain Tumor Segmentation Methods," *Tsinghua Science and Technology*, vol. 19, no. 6, pp. 578–595, 2014.
- [2] S. Debnath, S. Baishya, D. Sen, and W. Arif, "A hybrid memory-based dragonfly algorithm with differential evolution for engineering application," *Engineering with Computers*, no. 0123456789, 2020.
- [3] J. Nayak and B. Naik, "Fuzzy C-Means (FCM) Clustering Algorithm : A Decade Review from 2000 to 2014," *Computational Intelligence in Data Mining*, vol. 2, no. January, pp. 133–149, 2015.
- [4] S. Mirjalili, "Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems," *Neural Computing and Applications*, vol. 27, no. 4, pp. 1053–1073, May 2016.
- [5] A. Tharwat, T. Gabel, and A. E. Hassanien, "Parameter Optimization of Support Vector Machine Using Dragonfly Parameter Optimization of Support Vector Machine using Dragon y Algorithm," *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2017*, vol. 639, no. Chapter 29, pp. 309–319, 2017.
- [6] "brain tumor dataset." https://figshare.com/articles/brain_tumor_dataset/1512427 (accessed May 16, 2019).