Neural Network with Bee Colony Optimization for MRI Brain Cancer Image Classification

Sathya Subramaniam and Manavalan Radhakrishnan
Department of Computer Science and applications, Periyar University, India

Abstract: *Brain tumor is one of the foremost causes for the increase in mortality among children and adults. Computer visions are being used by doctors to analysis and diagnose the medical problems. Magnetic Resonance Imaging (MRI) is a medical imaging technique, which is used to visualize internal structures of MRI brain images for analyzing normal and abnormal prototypes of brain while diagnosing. It is a non-invasive method to take picture of brain and the surrounding images. Image processing techniques are used to extract meaningful information from medical images for the purpose of diagnosis and prognosis. Raw MRI brain images are not suitable for processing and analysis since noise and low contrast affect the quality of the MRI images. The classification of MRI brain images is emphasized in this paper for cancer diagnosis. It can consist of four steps: Pre-processing, identification of Region of Interest (ROI), feature extraction and classification. For improving quality of the image, partial differential equations method is proposed and its result is compared with other methods such as block analysis method, opening by reconstruction method and histogram equalization method using statistical parameters such as carrier signal to ratio, peak signal-to-ratio, structural similarity index measure, figure of merit, mean square error. The enhanced image is converted into bi-level image, which is utilized for sharpening the regions and filling the gaps in the binarized image using morphological operators ROI is identified by applying region growing method for extorting the five features. The classification is performed based on the extracted image feature to determine whether the brain image is normal or abnormal and it is also, introduced hybridization of Neural Network (NN) with bee colony optimization for the classification and estimation of cancer affect on given MRI image. The performance of the proposed classifier is compared with traditional NN classifier using statistical measures such as sensitivity, specificity and accuracy. The experiment is conducted over 100 MRI brain images.*

Keywords: *MRI images, NN, bee colony, PDE, biological analysis, feature extraction.*

1. Introduction

Magnetic Resonance Imaging (MRI) is an imaging technique used primarily to produce the medical images of the human body for prognosis. It is very necessary to enhance the contrast of such images before further processing or analysis can be conducted. Image enhancement is a challenging research area in image processing. The goal and objectives of this paper is to increase the classification rate of brain tumour detection.

In this paper, Classification of MRI brain images is proposed for the purpose of cancer diagnosis. It can be concluded with four steps: Pre-processing, identification of Region of Interest (ROI), feature extraction and classification. In this paper, PDE enhancement method is proposed for improving the quality of the image [13]. The enhanced image is converted into bi-level image, which is utilized for sharpening the regions and filling the gaps in the binarized image using Morphological operators. ROI is identified by applying region growing method for extort the five features. The classification is performed based on the extracted image feature to determine whether the brain image is normal or abnormal.

This paper is organized as follows: Section 2 describes the preprocessing. In section 3, region extraction is introduced, section 4 describes feature extraction and section 5 describes classification of MRI images. Section 6 shows an experimental result analysis and discussion. Finally, conclusions are presented in section 7.

2. Pre-Processing

The image enhancement is the process of improving the visibility and perceptibility of the various regions in the image that can be partitioned and extract image features from the regions [4]. Image enhancement is usually followed by identification of ROI and feature extraction, which is of paramount importance in low-level vision [10, 14].

A traditional method of image enhancement is hard to read the goal out exactly from the low contrast image and it takes for more iteration, time-consuming and less obvious. An algorithm is used to enhance the weak edges. It can effectively improve the readability of the image. Besides, for its reducing the number of iterations, it had a great improvement in efficiency. This method can be applied to real-time processing of video images in the dark. The core code of this algorithm as follows:

Algorithm 1: Image enhancement.

$t1=exp((-row^2- col^2)/4/e)$;
$t2=exp ((-I^2*c^2/4/(e^{-3})*t1^2*(row + col)^2)/k^2)$;
$t3=I*8- N- S- W- E- EN- ES- WS- WN$;
$t4=(1 + w)*t2-w$;
if $(I<=e)$
  $h=0$;
else if$(I >= A * e)$
  $h=1$;
end
end

$$I=I+(I*(1-h)+ h* I* t3)*t4 \qquad (1)$$

Where *I* is one image to be processed, *w=w*, *k=k*, *e=e*0. *t*1 and *t*4 are the middle of the volumes. According to *t*1 and *t*2, we get *g(I)*, while from *t*3, we get $div_{(\nabla I \div |\nabla I|)}$ and $[(1+W)\, G\, (|\nabla G * \nabla u|) - w]$ is from *t*4. In accordance with the number of iteration, the above codes carry out in cycle.

3. Binarization

Image binarization converts an image of up to 256 gray levels to a black and white image. Frequently, binarization is used as a pre-processor. In this paper, PDE based image enhance method is used to enhance the MRI image [13]. The simplest way to use image binarization is to choose a threshold value and classify all pixels with values above this threshold as white, and all other pixels as black.

4. Region Extraction

Morphological operation is used as an image processing tools for sharpening the regions and filling the gaps for binarized image. The dilation operator is used for filling the broken gaps at the edges and to have continuities at the boundaries. In our method, after a base image is generated from one of slice image in MRI data, region growing method is applied to the selected slice image based on the base image [1]. The area which is obtained by region growing method is considered as a new base image in the next step and this extraction process is repeated for all slice images. Finally, we extracted the area of tumor and brain and both are visualized in three-dimensional domain simultaneously to understand the position relations of the tumor. Onto the dilated image a filling operator is applied to fill the close contours. To filled image, centroids are calculated to localize the regions as shown beside. The final extracted region is then logically operated for extraction of massive region in given MRI image.

5. Feature Extraction

The feature extraction extracts the features of importance for image classification [11]. The following features are extracted from MRI images. Table 1 shows features and its formula.

Table 1. Features and its formula.

Features	Formula
Contrast	$f1 = \sum_{i,j=0}^{N-1} P_{i,j}(i,j)^2$
Angular Second Moment (ASM)	$f2 = \sum_{i,j=0}^{N-1} P_{i,j}$
Inverse Difference Moment (Homogeneity)	$f3 = \sum_{i,j=0}^{N-1} \dfrac{P_{i,j}}{i-(i*j)}$
Entropy	$f4 = \sum_{i,j=0}^{N-1} P_{i,j}(-\ln P_{i,j})$
Maximum Probability	$f5 = \max_{i,j} p(i,j)$

6. MRI Image Classification

Classification is a computational procedure that sorts images into groups ("classes") according to their similarities [2, 15]. The classification process is to categorize all pixels in a digital image into one of several land cover classes or "themes". This categorized data may then be used to produce thematic maps of the land cover present in an image. Normally multispectral data are used to perform the classification and indeed, the spectral pattern present within the data for each pixel is used as the numerical basis for categorization. The objective of image classification is to identify and portray, as a unique gray level (or color), the features occurring in an image in terms of the object or type of land cover these features actually represent on the ground [8].

6.1. Proposed Methods

In this thesis, supervised classifier Neural Network (NN) with BCO is proposed for classifying MRI images with regards to brain tumour. Then, it is compared with.

In machine learning and pattern recognition, classification refers to an algorithmic procedure for assigning a given piece of input data into one of a given number of categories. An algorithm that implements classification, especially in a concrete implementation, is known as a classifier. The term "classifier" sometimes also refers to the mathematical function, implemented by a classification algorithm that maps input data to a category. In this paper, machine learning algorithm is implemented such as NN and NN with BCO classifier. Figure 1 shows proposed methodology for brain image classification.
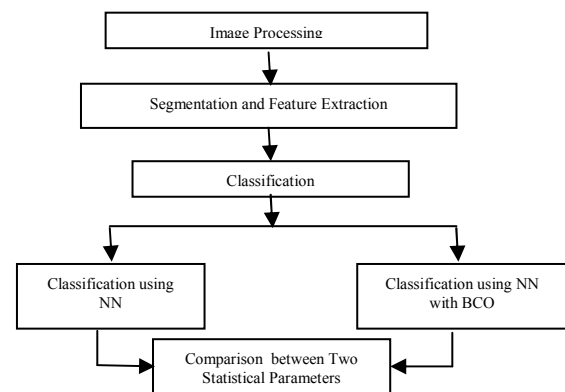

Figure 1. Proposed methodology for brain image classification.

6.1.1. NN Classifier

NNs are a proven, widely used technology to solve such complex classification problems. Loosely modelled after the human brain, NNs are interconnected networks of independent processors that, by changing their connections (known as training), learn the solution to a problem. The architecture of NN is shown in Figure 2.
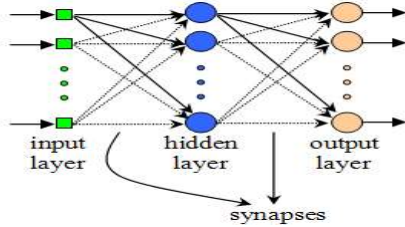

Figure 2. Architecture of NN.

An Artificial NN (ANN), usually called NN, is a mathematical model or computational model that is inspired by the structure and/or functional aspects of biological NNs. ANN consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation.

A feed Forward NN (FNN) is an ANN where connections between the units do not form a directed cycle. This is different from recurrent NNs. In this network, the information moves in only one direction, forward, from the input nodes, through the hidden nodes (if any) and to the output nodes. There are no cycles or loops in the network.

Algorithm 2: Neural network.

Initialization of weights
Step 1: *Initialize weight to small random values.*
Step 2: *While stopping condition is false, do steps 3-10*
Step 3: *For each training pair do steps 4-9. Feed forward.*
Step 4: *Each input unit receives the input signal x_i and transmits this signals to all units in the layer above i.e., hidden units.*
Step 5: *Each hidden unit (Z_j, j=1,…p) sums its weighted input signals*

$$Z_{-inj} = V_{oj} + \sum_{i=1}^{n} x_i V_{ij} \qquad (2)$$

Applying activation function:

$$Z_j = f(Z_{inj}) \qquad (3)$$

And sends this signal to all units in the layer above i.e., output units.

Step 6: *Each output unit (Y_k, k=1,…, m) sums its weighted input signals.*

$$y_{-ink} = w_{ok} + \sum_{j=1}^{p} z_j w_{jk} \qquad (4)$$

And applies its activation function to calculate the output signals.

$$Y_k = f(y_{-ink}) \qquad (5)$$

Back Propagation of Errors.

Step 7: *Each output unit (Y_k, k=1,…, m) receives a target pattern corresponding to an input pattern error information term is calculated as*:

$$\delta_k = (t_k - y_k) f(y_{-ink}) \qquad (6)$$

Step 8: *Each hidden unit (Z_j, j=1,…, n) sums its delta inputs from units in the layer above.*

$$\delta_{-inj} = \sum_{k=1}^{m} \delta_j w_{jk} \qquad (7)$$

The error information term is calculated as:

$$\delta_j = \delta_{-inj} f(z_{-inj}) \qquad (8)$$

Updating of Weight and biases.

Step 9: *Each output unit (Y_k, k=1,…, m) updates its bias and weights (j=0, …, p)*
The weight correction term is given b:

$$\Delta W_{jk} = \alpha \delta_k z_j \qquad (9)$$

And the bias correction term is given by:

$$\Delta W_{ok} = \alpha \delta_k \qquad (10)$$

Therefore,

$$W_{jk}(new) = W_{jk}(old) + \Delta W_{jk} \qquad (11)$$

$$W_{ok}(new) = W_{ok}(old) + \Delta W_5 \qquad (12)$$

Each hidden unit (Z_j, j=1, …, p) updates its bias and weights (i=0,…, n). The weight correction term:

$$\Delta V_{ij} = \alpha \delta_j x_i \qquad (13)$$

The bias correction term:

$$\Delta V_{oj} = \alpha \delta_j \qquad (14)$$

Therefore,

$$V_{ij}(new) = V_{ij}(old) + \Delta V_{ij} \qquad (15)$$

$$V_{oj}(new) = V_{oj}(old) + \Delta V_{oj} \qquad (16)$$

Stop 10: *Test the stopping condition.*

The stopping condition may be the minimization of the errors, number of epochs etc.

In this research work, classification techniques based on feed forward NN are proposed and applied to brain image slices classification using features derived from slices. This is for separating abnormal slices from the data collection containing both normal and abnormal slices. The purpose is to perform segmentation process for tumor calculation only on abnormal slices. The categorization of slices into normal and abnormal is done using statistical features of images such as mean, variance and co-occurrence based textural features of images such as energy, entropy, difference moment, inverse difference moment and correlation.

6.1.2. Neural Network with BCO Classifier

The FNN was chosen as the classifier because it is a powerful tool among supervised classifiers and it can

classify nonlinear separable patterns and approximate an arbitrary continuous function [3]. However, to find the optimal parameters of FNN is a difficult task because the search algorithms are easily trapped in local extreme. Recently, there have been many algorithms available to train the FNN, such as Back-Propagation (BP) algorithm, Genetic Algorithm (GA) [12], Elite Genetic Algorithm with Migration (EGAM) [6], Simulating Annealing (SA) algorithm and Particle Swarm Optimization (PSO) [7]. Unfortunately, the BP, GA, SA and PSO algorithms all demand expensive computational costs and can still be easily trapped into the local best, hence, would probably end up without finding the optimal weights of the FNN. In this paper, we use the NN with BCO Algorithm 3 to find its optimal weights.

Algorithm 3: NN with BCO classifier.

Step 1: *Initialize the population of solutions xij (here I denotes the ith solution and j denotes the jth epoch, i = 1, . . ., SN , here SN denotes the number of solutions with j=0.*

$$x_{i0} = LB + rand(\cdot) \times (UB - LB)(i=1...SN) \qquad (17)$$

Here, LB and UB represents the lower and upper bounds, which can be infinity if not specified. Then, evaluate the population via the specified optimization function.

Step 2: *Repeat and let j=j+1.*
Step 3: *Produce new solutions (food source positions) υij in the neighborhood of xij for the employed bees using the formula:*

$$\upsilon_{ij} = x_{ij} + \Phi_{ij}(x_{ij} - x_{kj}) \qquad (18)$$

Here, k is a solution in the neighborhood of i, Φij is a chaotic random number in the range [−1, 1] calculated·
Step 4: *Apply the greedy selection process between x$_{ij}$ and υ$_{ij}$.*
Step 5: *Calculate the probability values P$_{ij}$ for the solutions x$_{ij}$ by means of their fitness values using the equation:*

$$P_{ij} = \frac{fit_{ij}}{\sum_{i=1}^{SN} fit_{ij}} \qquad (19)$$

Here, fit denotes the scaled fitness as shown in equation.

Step 6: *Normalize P$_{ij}$ values into [0, 1].*
Step 7: *Produce the new solutions (new positions) υ$_{ij}$ for the onlookers from the solutions x$_{ij}$ as Step 3, selected depending on P$_{ij}$ and evaluate them.*
Step 8: *Apply the greedy selection process for the onlookers between xij and υij.*
Step 9: *Determine the abandoned solution (source), if exists and replace it with a new randomly produced solution xi for the scout using the equation:*

$$x_{ij} = \min_i(x_{ij}) + \phi_{ij} * (\max_i(x_{ij}) - \min_i(x_{ij})) \qquad (20)$$

Here, φ$_{ij}$ is a random number in [0, 1].

Step 10: *Memorize the best food source position (solution) achieved so far.*
Step 11: *Go to Step 2 until termination criteria met.*

Karaboga *et al*. [5] under the inspiration of collective behavior on honey bees with better performance in function optimization problems compared with GA, Differential Evolution (DE) and PSO [9]. As is known, normal global optimization techniques conduct only one search operation in iteration. For example, the PSO carries out a global search at the beginning stage and local search in the ending stage [14] nevertheless, the NN with BCO features in the following advantage in that it conducts both a global search and local search in each iteration and as a result the probability of finding the optimal is significantly increased, which effectively avoids local optima to a large extent.

7. Experimental Result Analysis and Discussion

The proposed methods for classify the enhanced brain image classification have been implemented using MATLAB. The results of each method are analyzed by using statistical parameters such as *Sensitivity*, *Specificity* and *Accuracy*. The detailed descriptions of these Statistical parameters are given bellow.

7.1. Confusion Matrix

A confusion matrix contains information about actual and predicted classifications done by a classification system. Performance of such systems is commonly evaluated using the data in the matrix. The following table shows the confusion matrix for a two class classifier. Table 2 shows a confusion matrix for a classification problem with positive and negative class values.

Table 2. Confusion matrix classification problem.

	Positive Prediction	**Negative Prediction**
Positive Class	True Positive (TP)	False Negative (FN)
Negative Class	False Positive (FP)	True Negative (TN)

- TP: Abnormal brain correctly diagnosed as abnormal.
- FP: Normal brain incorrectly identified as abnormal.
- TN: Normal brain correctly identified as normal.
- FN: Abnormal brain incorrectly identified as normal.
- *Sensitivity*: Relates to the test's ability to identify positive results. *Sensitivity* measures the proportion of actual positives which are correctly.

$$Sensitivity = \frac{no.oftruepositives}{no.offalsepostives + no.offalsenegatives} \qquad (21)$$

If a test has high *Sensitivity* then a negative result would suggest the absence of disease.

- *Specificity*: Specificity relates to the test's ability to identify negative results. Specificity measures the proportion of negatives which are correctly identified.

$$Specificity = \frac{no.oftruenegatives}{no.offalsenegatives + no.offalsepositives} \qquad (22)$$

If a test has high specificity, a positive result from the test means a high probability of the presence of disease.

- *Accuracy*: *Accuracy* is also used as a statistical measure of how well a binary classification test correctly identifies or excludes a condition. That is, *Accuracy* is the proportion of true results (both true positives and true negatives) in the population. It is a parameter of the test.

$$Accuracy = \frac{no.oftruepositives + no.oftruenegatives}{no.oftruepos + falsepos + falseneg + trueneg} \quad (23)$$

Computation of NN method for image classification is presented in the Table 3 and computation of NN with BCO method is presented in the Table 4.

Table 3. NN performance analysis value.

No. Of Images	Sensitivity	Specificity	Accuracy
Ten	0.5	0.83	0.7
Twenty	0.85	0.666	0.8
Thirty	0.5	0.833	0.7
Forty	0.8	0.8	0.8
Fifty	0.75	0.833	0.8
Sixty	0.8	0.6	0.7
Seventy	0.66	0.714	0.7
Eighty	0.8	0.8	0.8
Ninety	0.5	0.666	0.6
Hundred	0.75	0.83	0.8
Average	**0.691**	**0.7572**	**0.74**

Table 4. NN with BCO performance analysis value.

No. Of Images	Sensitivity	Specificity	Accuracy
Ten	0.6	0.83	0.7
Twenty	0.5	0.65	0.7
Thirty	0.75	0.833	0.8
Forty	0.88	0.8	0.8
Fifty	0.75	0.66	0.7
Sixty	0.75	0.857	0.8
Seventy	0.63	0.85	0.7
Eighty	0.6	0.87	0.8
Ninety	0.8	0.8	0.8
Hundred	0.75	0.83	0.8
Average	**0.701**	**0.798**	**0.76**

The experiment is conducted over the different MRI images with required statistical parameters and their average result is shown in Table 5.

Table 5. Average values of performance analysis.

Methods	Sensitivity	Specificity	Accuracy
NN	0.69	0.75	0.74
NN+BCO	0.70	0.79	0.76

In this paper, MRI image should be segmented before feature extraction and classification. This segmentation process is done by morphological operations. The classification result is shown in Figure 3.

The detailed analysis of NN and NN with BCO individual measure and the performance analysis chart of NN and NN with BCO are presented among various parameters in Figure 4.

The computational results clearly explain that the NN with BCO performing better than NN. Since, *sensitivity* value, *specificity* and *accuracy* values are high. The performance analysis chart of NN and NN with BCO is presented in Figure 5.
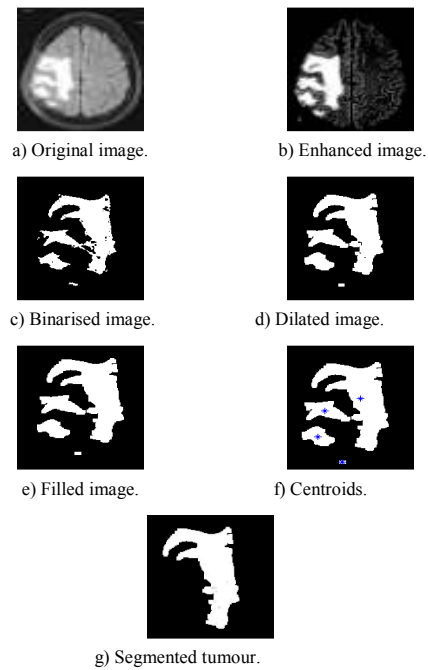

a) Original image.   b) Enhanced image.

c) Binarised image.   d) Dilated image.

e) Filled image.   f) Centroids.

g) Segmented tumour.

Figure 3. Classification of brain image.

a) Sensitivity of NN.   b) Specificity of NN.

c) Accuracy of NN.   d) Sensitivity of NN with BCO.

e) Specificity of NN with BCO.   f) Accuracy of NN with BCO.

g) Performance analysis of NN.   h) Performance analysis of N with BCO.
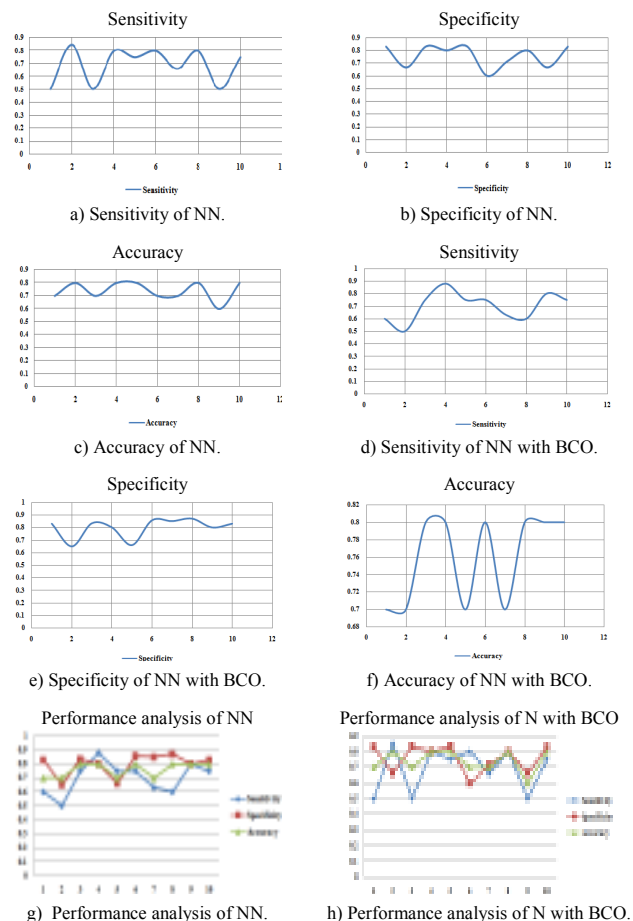
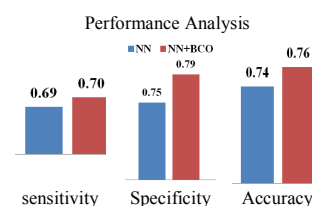Figure 4. Statistical measures of NN and NN with BCO.

Figure 5. Performance analysis.

7. Conclusions

MRI used for brain scans is very useful and effective technique to detect the brain tumor. It is a non-invasive method to take picture of brain and the surrounding images. MRI image quality is enhanced with the help of PDE enhancement method and compared with others. The experiment results clearly show that PDE method is best for enhancing brain MRI image with respect to statistical measures. The enhanced image is converted into bi-level image, which is utilized to Morphological operators for sharpening the regions and filling the gaps for binarized image. ROI is identified effectively using region growing method while almost no loss of information in the actual brain areas of the pre-processed MRI images. From the ROI, five features are extracted. The classification is performed based on the extracted image feature to determine whether the brain image is normal or abnormal using supervised classifier algorithms. The performance of classification algorithm is measured using classification parameters such as *Sensitivity*, *Specificity* and *accuracy*. The experiment conducted over 100 MRI brain images with normal and abnormal cases. NN with BCO achieves 76% *accuracy* while *accuracy* of NN is 74%. The computational results conclude that the proposed classification algorithm NN with BCO realize 2% higher than NN in terms *accuracy*.

References

[1] Cline H., Lorensen E., Kikinis R., and Jolesz F., "Three Dimensional Segmentation of MR Images of the Head using," *Journal of Computer Assisted Tomography*, vol. 14, no. 6, pp. 1037-1045, 1990.

[2] Cocosco A., Zijdenbos P., and Evans C., "A Fully Automatic and Robust Brain MRI Tissue Classification Method," *Medical Image Analysis*, vol. 7, no. 4, pp. 513-527, 2003.

[3] Coulibaly P. and Evora D., "Comparison of Neural Network Methods for Infilling Missing Daily Weather Records," *Journal of Hydrology*, vol. 341, no. 2, pp. 27-41, 2007.

[4] Gonzalez C., Woods E., and Masters R., "Digital Image Processing," *Journal of Biomedical Optics*, vol. 14, no. 2, pp. 029901-029901, 2009.

[5] Karaboga N., Kalinli A., and Karaboga D., "Designing Digital IIR Filters using Ant Colony Optimisation Algorithm," *Engineering Applications of Artificial Intelligence*, vol. 17, no. 3, pp. 301-309, 2004.

[6] KellegÄoz T., Toklu B., and Wilson J., "Elite Guided Steady-State Genetic Algorithm for Minimizing Total Tardiness in Flowshops," *Computers and Industrial Engineering*, vol. 58, no. 2, pp. 300-306, 2010.

[7] Kiranyaz S., Ince T., Yildirim A., and Gabbouj M., "Evolutionary Artificial Neural Networks by Multi-Dimensional Particle Swarm Optimization," *Neural Networks*, vol. 22, no. 10, pp. 1448-1462, 2009.

[8] Mahsa C. and Woo C., "Medical Image Segmentation using a Multi-Agent System Approach," *The International Arab Journal of Information Technology,* vol. 10, no. 3, pp. 222-229, 2013.

[9] Martin H., Howard D., Mark B., Orlando D., *Neural Network Design*, Vikas Publishing House, New Delhi, India, 1996.

[10] Perona P. and Malik, J., "Scale Space and Edge Detection using Anisotropic Diffusion," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629-639, 1990.

[11] Primary Brain Tumors in the United States Statistical Report 2002-2004. available at: http://www.cbtrus.org/reports/2007-2008/ 2007 report.pdf, last visited 2007,

[12] Robotham H., Bosch P., Gutiérrez-Estrada J., Castillo J., Pulido-Calvo I., "Acoustic Identification of Small Pelagic Fish Species in Chile using Support Vector Machines and Neural Networks," *Fisheries Research*, vol. 102, no. 2, pp. 115-122, 2010.

[13] Sathya S. and Manavalan R., "Analysis of Background Detection and Contrast Enhancement of MRI Images," *International Journal of Computer Applications*, vol. 36, no. 12, 2011.

[14] Shujun F., Qiuqi R., Chengpo M., and Wenqia W., "Edge Detection for Image Measurement based on Nonlinear Diffusion Filterin," *in Optics and Precision Engineering*, vol. 15, no. 2, pp. 289-293, 2007

[15] Zhang S., Gong S., Guan Y., Zhang P., and Gong Q., "A Novel Iga-Edspso Hybrid Algorithm for the Synthesis of Sparse Arrays," *Progress in Electromagnetics Research*, vol. 89, pp. 121-134, 2009.

[16] Zulaikha B. and Mohamed S., "A Robust Segmentation Approach for Noisy Medical Images using Fuzzy Clustering With Spatial Probability," *the International Arab Journal of Information Technology*, vol. 9, no. 1, pp. 74-83, 2012.

Sathya Subramaniam Obtained MSc, computer science from Vivekananda College of Arts and Science, Periyar University, India, in 2008, MPhil in computer science from KSR College of Arts and Science, Periyar University, India in 2011. She worked as Asst Prof, Department of Computer Science and Applications, Kandaswami kandar's College, India.

Manavalan   Radhakrishnan obtained MSc, computer science from St.Joseph's College of Bharathidasan University, India in 1999 and MPhil in computer science from Manonmaniam Sundaranar University, India in 2002. He works as Asst Prof and Head, Department of Computer Science and Applications, KSR College of Arts and Science, India. He pursues PhD in Medical Image Processing. His areas of interest are medical image processing and analysis, soft computing, pattern recognition and theory of computation.