Multimodal Brain Tumor Detection and Classification in MRI Images Using Convolutional Neural Network

Mr.D.Balaji [1],Ms.S.Charulatha[2], Ms.B.Susmitha[3], Ms.M.Vineka[4]
[1] *Assistant Professor Department of Electronics and Communication Engineering, Mahendra College of Engineering, Salem (Dt), TamilNadu-636 106, India.*
[2,3,4] *Students of Department of Electronics and Communication Engineering, Mahendra College of Engineering, Salem (Dt), TamilNadu-636 106, India.*

Abstract A groundbreaking advancement in clinical practice is the automated classification and recognition of brain tumors. This study presents an advanced clinical support system designed to automatically discern the type and progression of brain tumors using finely tuned neural networks. Initial data processing involves Ensemble Genetic filtering and an enhanced graph cut method, succeeded by segmentation through Fuzzy clustering using Local Approximation of Memberships (Segmentation clustering). Feature extraction is then conducted using Gray Level Co-occurrence Matrix (GLCM) and Histogram of oriented gradients (HOG). The model further employs Firefly-optimized Probabilistic Neural Networks for classification and Cuckoo Search-optimized Adaptive Neuro-Fuzzy Inference System (NFS) for tumor identification, resulting in highly precise tumor size and stage determination. The system's efficacy was assessed, achieving an impressive 98.3% accuracy rate. Physicians can leverage the outcomes of this system for informed clinical interventions.

Keywords Automated classification, brain tumors, neural networks, Ensemble Genetic filtering, Segmentation clustering, feature extraction, clinical interventions.

INTRODUCTION

A significant advancement in medical practice involves the automated identification and categorization of brain tumors. These aberrant tissue masses, when situated in delicate brain regions, can impair bodily functions. Brain tumors are typically classified as either Benignant, which have limited capacity for spreading but can still affect surrounding healthy tissue, or Neoplastic., which are more aggressive and commonly referred to as brain cancer. Imaging techniques such as Computed Tomography (CT), Ultrasound, and magnetic resonance imaging (MRI) play a pivotal role in detecting and treating these tumors [1-2]. Manual examination by radiologists and other specialists is time-consuming and prone to variability in tumor detection. Computer-aided diagnosis (CAD) systems aim to enhance diagnostic accuracy and have been particularly useful for remote treatments and telemedicine applications [3], facilitated by wireless body area networks and wireless multimedia sensor networks for transmitting patient information [4]. However, automatic segmentation and characterization of brain tumors from MRI images remain challenging due to the diverse nature of tumor types, shapes, and appearances. Thus, there is a pressing need for robust and automated diagnostic tools for tumor identification, segmentation, and classification [5].

Several methodologies for brain tumor detection and segmentation are documented in the literature [6], [7], [8], [9]. For instance, Devasena and Hemalatha proposed a CAD framework based on a hybrid anomaly detection algorithm (HADA) for localizing anomalous regions in MRI scans [10]. The framework includes noise reduction, smoothing, feature extraction, feature reduction, and classification steps. In another study, Othman and Basri introduced an effective computer-aided detection system for brain tumor classification, employing square-based segmentation and component labeling for tumor segmentation, and texture features extracted using GLCM [12] for feature extraction [11]. Subsequently, a probabilistic neural network (PNN) technique was employed for tumor classification. Arakeri and Reddy developed a software framework for tissue characterization of brain tumors, utilizing multi-channel preprocessing to enhance input images and a combination of modified FCM clustering and wavelet decomposition for segmentation [13]. Three different

classifiers, namely support vector machine (SVM), artificial neural network (ANN), and k-nearest neighbor (k-NN), were utilized for tumor stage detection. Dandıl et al. developed a CAD system for brain tumor classification into Benignant and Neoplastic. categories, employing Spatial-Fuzzy C-Means (FCM) segmentation and SVM for classification, achieving a classification accuracy of 91.49% [14].

Despite the plethora of tools developed for automatic brain tumor detection and classification, identifying the most effective one has become a cumbersome task. Many tools with similar methodologies and limited precision have compounded the issue. However, thorough exploration has revealed several limitations in existing mechanisms across various phases of brain tumor detection, including pre-processing, segmentation, feature extraction, and classification. Analyzing these challenges, this paper presents a computer-aided decision support system designed to classify and identify brain tumor stages with improved accuracy. Advanced techniques are employed to enhance brain tumor detection in this developed system.

Implementation of proposed system

The functional block diagram of the proposed brain tumor detection and classification decision support system is given in Fig. 1. Initially the MRI brain images are collected and processed using the noise removal and Brain extraction techniques. Then the features are extracted and are used to analyse and classify the tumor for further diagnosis by the clinical experts or physicians.

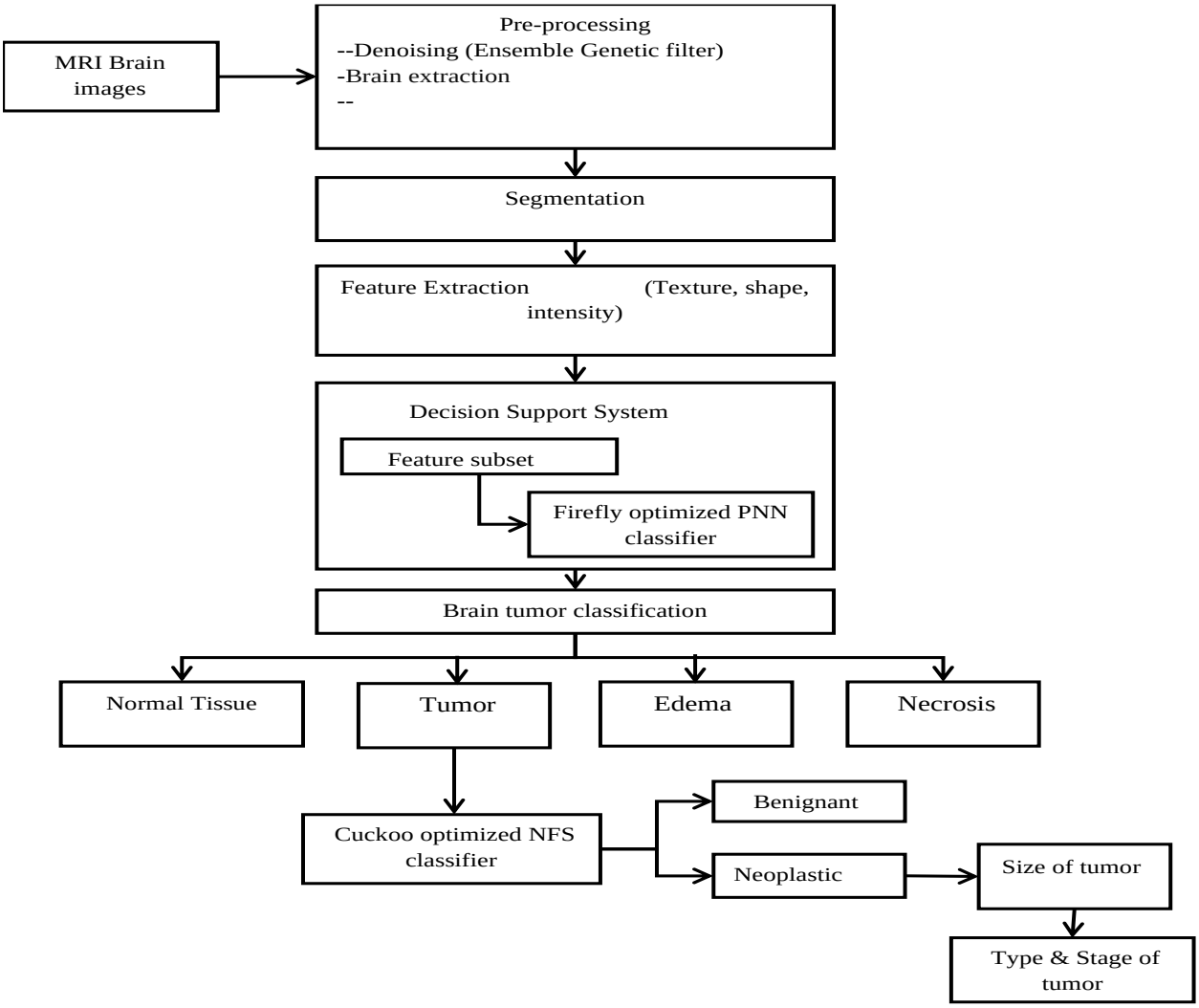

Fig.1 Functional Block diagram of proposed system

Step 1: Data Acquisition and pre-processing

The input MRI brain images to be utilized in this study are collected from patients with symptoms of brain tumor. The MRI dedicated systems are utilized to capture these images which are visualized using a computer. The collected images are first pre-processed to remove the noises, Brain extraction and enhancing the quality. Thus the pre-processing system of the proposed diagnosis tool consists of three stages: Noise removal, Brain extraction and Image enhancement.

Noise removal is accomplished through a novel ensemble genetic filter which combines the median and Gaussian filters through genetic programming (GP). To prepare the GP, first features are removed from the filters to frame the component vector. For this reason, first features are extracted by proposing the use of Median and Gaussian filters. In Median filter [15], a neighbourhood window is built up around a focal pixel. The focal pixel is supplanted with median value. To choose the median value, all pixels are requested in an sorted way and the median is chosen. The median filter is characterized as [31]

$$u[m,n] = median\{v[i,j], \quad i,j \in NW\} \quad (1)$$

Where *NW* represents squared neighbourhood window centered at pixel position $[m,n]$in the image.

The Gaussian filter [16]accomplishes de-convolution by inverse filtering process by means of high-pass filter and eradicates noise by a low-pass filter. The Gaussian filter in Fourier form is said as:[31]

$$W(f_1,f_2) = \frac{H^*(f_1,f_2)S_{xx}(f_1,f_2)}{|H(f_1,f_2)|^2 S_{xx}(f_1,f_2)+S_{nn}(f_1,f_2)} \quad (2)$$

Where $S_{xx}(f_1,f_2)$ and $S_{nn}(f_1,f_2)$ are power spectrum of original image and noise respectively, $H(f_1,f_2)$ represents the blurring filter.

After extracting the features using the median and Gaussian filters, the information of these features is combined and a feature vector is formed which contains the features of both filters Median and Wiener. The features are then used for the GP phase.Fig.2 shows the proposed ensemble genetic filter based noise removal system for MRI brain images.

In this stage, the feature vector is passed to Genetic Programming (GP) module to develop an optimal expression for the restoration of actual values of noisy pixels [17]. To begin the training of GP and develop the optimal numerical expression, it is essential to produce an underlying population to fill in as the beginning stage for the GP. So in the proposed strategy, an underlying population comprising of 100 individuals is made arbitrarily by utilizing the ramped half and half [18] technique. All factors, constants and capacities utilized have been initialized randomly at beginning.

After producing the population, the fitness of every solution has been computed by utilizing fitness measure. In the proposed strategy, Mean Squared Error (MSE) is utilized as fitness function which checks the general fitness of every individual to serve for the following generations. Individuals having least MSE values indicate great solution when contrasted with those solutions which have vast MSE values. Least MSE additionally demonstrates how much GP advanced while GP scientific expression has been instated arbitrarily. Amid advancement procedure of GP, there are distinctive techniques to choose best solutions from starting populace, yet competition choice has been performed for choice of solution for generation executions as it performs better among other choice models. For this situation, ten solutions have been chosen through tournament and utilized for generation measure. The tournament size assumes an essential part in the determination of individuals and on the off chance that it is bigger, powerless individuals have a little opportunity to be chosen. These solutions have been utilized later amid crossover and mutations which increment the development procedure of GP.

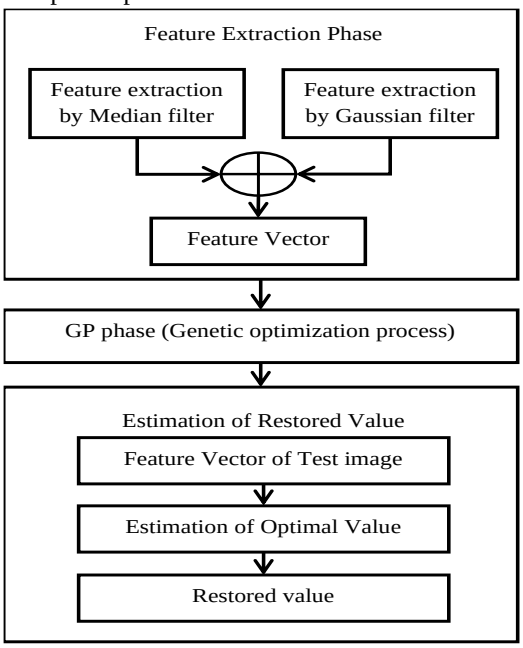

Fig. 2 Noise removal method

In the proposed strategy, one point crossover with likelihood 0.7 and point mutation with likelihood 0.1 have been utilized. Too high mutation rate diminishes the search capacity of GP while a too little quite often neglects to a nearby ideal. Crossover tries to focalize to a particular point. The proposed technique is tried with various blends of rates in the vicinity of 0 and 1 of crossover and mutation however the above settings are most appropriate for rebuilding of images ruined with impulse noise.

Table 1 Performance of Noise removal filtering techniques

S. No	Filtering Techniques	Parameters			
		MSE	CNR	IQI	MAE
1	Mean Filter	45.712	27.11	0.5412	38.761
2	Median Filter	31.87	38.041	0.5987	26.552
3	Gaussian filter	20.331	51.554	0.6645	21.99
4	Ensemble Genetic Filter	5.2119	70.885	0.7459	17.818

Table 2 Performance of Brain extraction techniques

S. No	Brain extraction techniques	Parameters			
		DS	**JS**	**FP**	**FN**
1	Distance waveform	0.60	0.44	0.0691	0.1188
2	Morphological opening	0.77	0.64	0.0776	0.1045
3	Graph cut	0.89	0.81	0.048	0.0613
4	Improved Graph cut	0.95	0.91	0.0305	0.0312

The proposed strategy is likewise tried by playing out the two point crossover and multipoint crossover however the single point crossover creates great outcomes. After advancing GP for number of times, an optimal scientific expression has been returned by the GP in a mechanized way. The best optimal expression returned by GP has been produced after total training of the GP by utilizing feature set created in the feature extraction step. The optimal scientific expression is a mix of every one of those features separated with median and Gaussian filters, capacities like trigonometric and a few constants. This best optimal developed expression has been created after more than one thousand generations. Thus, it can be concluded that testing to estimate the true value takes very less time as compared to training once GP has been completely trained. In order to de-noise the image, features have been extracted from testing noisy image and replaced in the optimal evolved expression. After re-establishing the true values of noisy pixels, these values are replaced in the noisy at their individual positions. In this way, the MRI brain image is de-noised.

For Brain extraction of the brain images, the improved graph cut method is utilized. The graph cut, which has been utilized commonly [19], was undesirable to segment skull regions due to the poor quality of images with speckles and low contrast. Hence to suit all images, the improved graph cut is proposed by the combination of energy function with elliptical shape constraint, which is given as: [31]

$$E = (1 - \mu)\{\lambda \sum_{p \in P}[w(p,S) + w(p,T)] + \sum_{p \in P,\{p,q\} \in N} w(p,q)\} + \mu \sum_{p \in P,\{p,q\} \in N} D\left(\frac{p+q}{2}\right) \quad (3)$$

Where S, T are two disjoint set; (p, q) are the pixel vertex; μ and λ are the weighting factor between two terms; $D(.)$ is defined as the distance from the pixel to the ellipse.

The proposed technique for Brain extraction utilizes three groups of operations. These are: thresholding to obtain preliminary mask, removal of narrow connections using graph cuts and post-processing. The objective of the last step is to reinstate partial volume gray matter (GM) voxels inadvertently uninvolved following thresholding. The improved graph cut algorithm can be applied to the removal of narrow connection as follows. Given the initial mask (foreground F and background B), the graph is defined on the whole image. The goal is then to choose appropriate seed regions and edge weights, so that the desired cut has the minimum value of all admissible cuts, i.e. cuts that separate the seed regions. For example, assume that (g1, g2) is the desired cut. First, the foreground and background seeds are needed to be

fully contained inside g2 and B respectively, since the cut cannot be made through the seeds. To make the value of the cut (g2,g2) smaller than that of (g1,g1) or any other admissible cut, the weights of the edges surrounding g2 should be made small while the weights of the edges surrounding other cuts large. While this is not an easy task in general, in the improved graph cuts helps this to be achieved in the case of separation of brain from non-brain tissues. Thus the correct separation depends on the cut value constraint.[31]

$$\hat{x} = \arg\inf_{g} |\delta g|, f \subseteq g, b \subseteq \bar{g} \qquad (4)$$

Where $|\delta g|$ is the ratio of cut value; f & b are the foreground and background seeds. Fig 3 shows the noise removed output of ensemble genetic filter and the skull stripped image using improved graph cut.

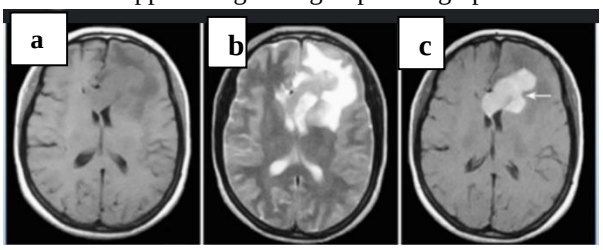

Fig.3 a) Input sample b) Noise removal output **c)** Brain extraction

Following the Brain extraction, the image enhancement is carried out. Image enhancement can be achieved by employing filter based techniques. In this model, the simpler median filter is utilized as the image is already pre-processed more efficiently during the noise removal and Brain extraction stages.

Step 2: Segmentation

The segmentation is to recognize at least one regions of interest (ROIs) from the chosen MRI image after pre-processing. The objective of image segmentation is to partition a brain image into an arrangement of semantically important, homogeneous, and non-covering regions of comparative attributes, for example, intensity, depth, colour, or texture. The segmentation result is either an image of labels recognizing each homogeneous region or set of contours which portray the region limits. Crucial segments of auxiliary brain MRI examination incorporate the arrangement of MRI information into particular tissue sorts and the distinguishing proof and depiction of particular anatomical structures. Arrangement intends to allocate to every component in the image a tissue class, where the classes are characterized ahead of time. The issues of segmentation and classification are interlinked because segmentation implies a classification, while a classifier implicitly segments an image. In the proposed decision support tool, the brain image segmentation is performed using Segmentation clustering.

SeSegmentation clustering [20] identifies clusters within dense regions of an image and organizes pixels into groups based on spatial connections. It constructs a k-Nearest Neighbors (KNN) graph to differentiate cluster centers and outliers. Regions of the brain with the highest local density are designated as Cluster Supporting Pixels (CSP), while areas with lower density than a specified threshold are labeled as outliers. CSPs are assigned full membership to represent them as cluster centroids, while outliers receive full membership to the outlier group. Remaining brain regions are then assigned fuzzy memberships, varying in degrees, to the cluster-supporting elements. This method autonomously determines the number of clusters and outliers without requiring pre-defined values. Segmentation clustering necessitates initial parameters such as the number of KNN and the threshold value for outliers. The segmented output generated by Segmentation clustering is illustrated in Fig 4.

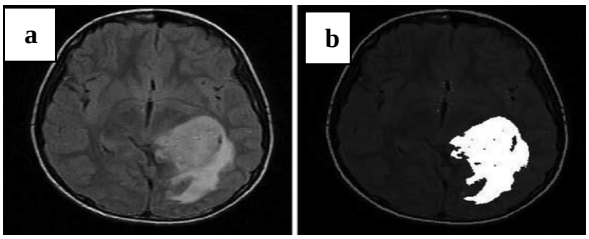

Fig.4 a) Clustering Process b) Segmented Output

The operation Segmentation for brain image segmentation can be summarized as follows.

1. Extraction of the structure information from the image

i) Construct a neighbourhood graph to connect each pixel to its KNN

ii) Estimate a density for each pixel based on its proximities to its KNN

iii) Pixels are classified into

a) CSP: pixel with density higher than all its neighbours

b) Cluster Outliers: pixel with density lower than all its neighbours and lower than a predefined threshold

2. Local/Neighbourhood approximation of fuzzy memberships

i) Initialization of fuzzy membership

a) Each CSP is assigned with fixed and full membership to itself to represent one cluster

b) All outliers are assigned with fixed and full membership to the outlier group

c) The remaining pixels are assigned with equal memberships to all clusters and the outlier group

ii) Then the fuzzy memberships of a wide range of pixels are refreshed by a uniting iterative system called Local/Neighbourhood approximation of Fuzzy Memberships.

iii) Cluster construction from fuzzy memberships in two possible ways

a) One-to-one pixel-cluster assignment, to assign each pixel to the cluster in which it has the maximum membership

b) One-to-multiple pixel-clusters assignment, to assign each pixel to the cluster in which exceeds a specified threshold.

Table 3 Performance of Brain image segmentation

S. No	Segmentation techniques	Parameters		
		Accuracy	Precision	Recall
1	K-means	0.821	0.765	0.778
2	Fuzzy C-means	0.856	0.7865	0.7932
3	Active contour model	0.8971	0.8123	0.8212
4	Cluster based Segmentation	0.987	0.869	0.8439

Step 3: Feature Extraction

Features, the characteristics of the objects of interest, if selected deliberately are representative of the maximum relevant information that the image has to offer for a complete characterization of a lesion. The motivation behind feature extraction is to minimize the tedious work of differentiating one pattern from another in order to provide qualified information to the classifier as relevant feature vectors. In this proposed technique, texture, shape and intensity features are extricated. For feature extraction, GLCM and HOG descriptors are utilized.

GLCM [21] is described for an image by the technique for dividing of co-occurring standards at a given offset. In the case of considering the gray-scale values of the image or different measures of shading, the GLCM is utilized for the estimation of texture of the image. Since GLCMs are reliably vast and intermittent, Features produced utilizing this technique is generally characterized as Haralick features. A mathematical definition of the GLCM for the brain image is as per the following:

- P(I,j), is the given a position operator
- Assume A is an n x n matrix whose element A[i][j] is the number of times that points with grey level value g[i] occur, in the position specified by P, relative to points with grey level value g[j].
- Then C be the n x n matrix that is formed by subdividing A with the entire number of point pairs that fascinate P. C[i][j] is a portion of the combined probability that two of points pleasing P with values g[i], g[j].C is called as co-occurrence matrix defined by P.

Assume t is a transcription, then a co-occurrence matrix C_t of a domain is defined for every grey-level (a, b) by: [31]

$$C_t(a,b) = card\{(p, p+t) \in R^2 | A[p] | a, A[p+t] = b \quad (5)$$

Here, $C_t(a,b)$ is the number of site-couples, denoted by $(p, p+t)$ that are separated by a transcription vector t, with a being the grey-level of p, and b being the grey-level of $p+t$ .Hence, for each Haralick texture features, a co-occurrence matrix is acquired. These co-occurrence matrices serve as the graphical sharing and the dependence of the grey levels within a local area. Each (i,j)th entry in the matrix representing the probability of transferring from one pixel of brain image with a grey level of 'i' to another with a grey level of 'j' under a preordained distance and angle. From these matrices, sets of analytical measures are calculated, called feature vectors.

Similarly, HOG feature vectors [22] also extracted for gradient character of brain images. A local feature similar with HOG descriptor is used to depict the gradient character of a block around the key point. For a matching point $f_i = \{d_i, l_i, x_i, o_i, a_i\}$, the extraction of local feature can be described as:

1. To make the local feature (HOG) be rotationally invariant, the image from which key point f_i is mined should be rotated relative to the overriding orientation o_i.
2. The image gradient magnitudes and orientations are sampled in a 16×16 cell block around the key point location x_i.
3. Gradient magnitudes are amassed into orientation histograms crisp the contents over 4×4 sub-regions with 8 histogram channels.
4. The description matrix is transformed to a local feature (HOG) vector, and the vector is then normalized to unit length. The final local feature (HOG) vector is a geometric description of key point.

As the initial step, the gradient image is processed by convolving the input image with a appropriate filter mask. A grid of histograms is then developed, where every histogram arranges the individual inclinations into containers as per their introduction. To preserve locality, a histogram is registered for every cell in an equally separated lattice. Accordingly, every cell contains a similar number of inclinations (contingent upon the cell estimate) and gets appointed one histogram. The cells themselves are then sorted out in rectangular blocks, which may cover. The histogram values of all cells inside one block are connected to frame a vector. The vector of each block is then standardized and in this way, the connection of every one of those block vectors yields the last feature vector.

Table 4 Quantitative measures of extracted features

S. No	Features	Estimated values	
		Benignant	Neoplastic.
1	Self-correlation	2.123	4.4729
2	Contrast	4.5678	1.0395
3	Correlation 1	-1.897	0.09941
4	Correlation 2	-1.765	0.0993
5	Cluster Salience	1.345	2.778
6	Cluster shade	-8.2426	4.6138
7	Dissimilarity	23.0755	31.785
8	Energy	1.878	0.0045
9	Entropy	8.765	5.4275
10	Homogeneity 1	8.62322	0.03138
11	Homogeneity 2	3.3342	0.00108
12	Maximum probability	1.9876	0.0064
13	Information coefficient of correlation 1	-4.876	-0.0322
14	Information coefficient of correlation 2	-5.32	0.40386
15	Sum of squares	2.6754	2.8082

Step 4: Categorisation of Tumor

i) Tumor Spotting

The first stage of brain tumor classification is the detection and diversification of the tumor tissues from edema, necrosis and normal tissues. For detecting the presence of tumor, the Firefly optimized PNN (F-PNN) classifier is employed. The mechanism of firefly optimization with PNN is shown in Fig 5.

Probabilistic neural network (PNN) [11] is developed based on the Bayes classification rules and the probability density function estimation method of Parzen window. PNN is a supervised neural network that is widely used in the area of pattern recognition and it has infinite potential in tumor diagnosis for its parallel-distributed processing, self-organization and self-study ability. A PNN consists of four layers: input, pattern, summation and output training/testing datasets (for example, from x_1 to x_n). The number of neurons in input layer is equal to the number of attributes in the dataset. The second layer contains m neurons, which is equal to the number of the training samples of all classes k, that is $m = N_1 + N_2 + .. + N_k$. The Euclidean distances between the input sample and training samples are calculated to acquire the similarity according to the Gaussian probability density function [31]

$$y_{ij} = \exp\left(-\frac{\left\|X-X_j^{(i)}\right\|^2}{2\sigma^2}\right) \quad i = 1,2,\dots,k \qquad (6)$$

Here $X = [x_1, x_2, .., x_n]$ is the argument vector to be classified, N_i is the number of training samples of ith class, and σ denotes the smoothing parameter.

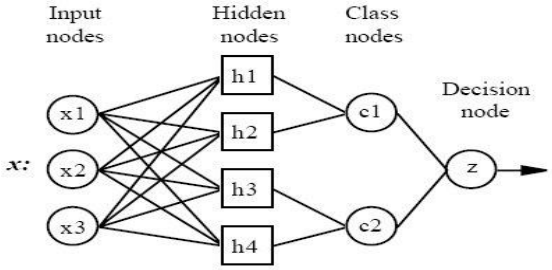

Fig. 5 PNN Architecture

The third layer contains summation units needed to complete the probability estimation. There are as many summation units as classes. Each summation unit receives input data only from those pattern units belonging to its respective class. Using the Parzen method, the probability density function for multiple variables can then be expressed as follows: [31]

$$pdf_i = \frac{1}{(\sigma\sqrt{2\pi})^n N_i} \sum_{j=1}^{N_i} y_{ij} \quad i = 1,2,\dots,k \quad (7)$$

Finally, the output layer can determine the decision category for input vector X. In this layer, the maximum of the summation node outputs can be found. Through a simple comparison, only the unit corresponding to the class with the highest summation unit value produces an output of one, while others generate a value of zero thereby indicating the classification decision for the input vector.[31]

$$O(X) = argmax\{pdf_1, pdf_2, .., pdf_k \qquad (8)$$

In the training process of PNN, the most important aspect is the selection of smoothing parameter σ. A proper choice of this parameter has a major influence on the classification ability of the network. Conventionally, the PNN only need to set the same smoothing parameter based on the experience, which cannot fully reflect the correlation degree between samples. Here, the smoothing parameter is computed separately for each attribute of the samples. This type of model is a more elastic classifier, since in such a case; the influence of each variable on neighbouring points differs. Then, the ith summation neuron provides the following output: [31]

$$pdf_i = \frac{1}{(\sqrt{2\pi})^n N_i \prod_{i=1}^{n} \sigma_j} \sum_{j=1}^{N_i} \exp\left(-\frac{\left\|X-X_j^{(i)}\right\|^2}{2\sigma_j^2}\right) \quad (9)$$

The firefly optimization algorithm [23] is used to optimize the smoothing parameters (the total number of the parameters is equal to n), in order to achieve the optimal classification effect. The flow diagram of the proposed optimized PNN using Firefly algorithm (FA) is shown in Fig 6.

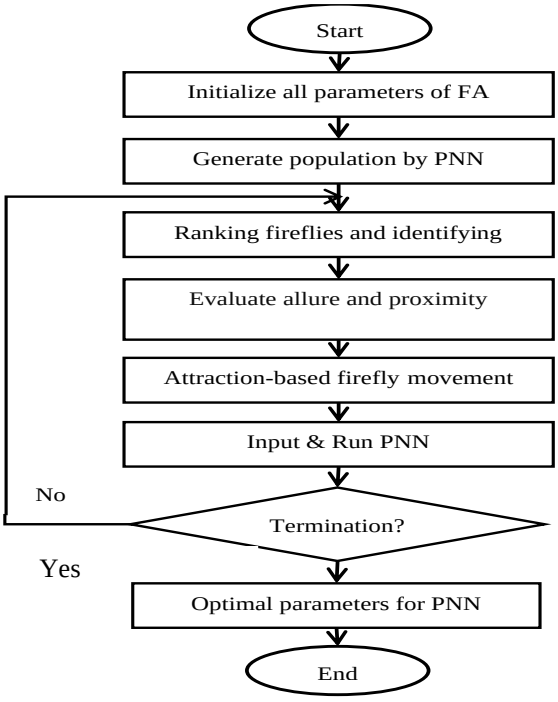

Fig. 6 Proposed F-PNN flowchart

In the proposed F-PNN, in order to achieve the optimal classification effect the PNN parameters are optimized. As the number of smoothing parameters is equal to the dimensionality of the samples, the firefly population is set by PNN. The attractiveness $\beta(r)$ determines the attractive property between the fireflies (in this case the smoothing of parameters). The distance r_{ij} and movement of firefly towards brighter one (defined as ff_i) are given as [31]

$$r_{ij} = \|ff_i - ff_j\| = \sqrt{\sum_{k=1}^{d}(ff_{i,k} - ff_{j,k})^2} \qquad (10)$$

$$ff_i = ff_j = \beta * \exp(-\gamma {r_{ij}}^2) * (ff_j - ff_i) + \alpha(rand - \tfrac{1}{2}) \qquad (11)$$

Table 5 Performance comparison of Feature Subset selection

Feature set	Specificity (%)	Sensitivity (%)	Overall Accuracy (%)
Texture	85	88	94.2
Shape	68	86	93.8
Intensity	79	83	94.1
Texture +Shape	94	89	97.2
Texture +Intensity	90	90	96.7
Shape+Intensity	93	88	97.4
Texture +Shape +Intensity	97	93	99.14

Here the first term is the current position of a firefly, the second term is used to consider the attraction of a firefly towards the intensity of the light of neighbouring fireflies, and the third term is used for the random movement of a firefly when it cannot 'see' any brighter ones. The coefficient α ، is a randomization parameter determined by the problem of interest, while rand is a random number generator consistently distributed in the space [0, 1].

Table 6 Details of Firefly Optimized Probabilistic Neural Network

S. No.	Functions used for F-PNN	Design parameters
1	Learning Rate	1.5
2	Momentum	2.5
3	Threshold value	0.0001
4	Activation	Sigmoid
5	Hidden Layer	2
6	Number of Hidden Units	40
7	Input neurons	28
8	Output Neuron	1
9	Maximum MSE	0.025
10	Number of iterations	500

In the proposed model, the smoothing parameters of PNN can be represented by $ff(i,j)$. Then, the intensity of light γ(also called the fitness value of firefly i) should be calculated. The classification accuracy is taken as the fitness function to represent the classification performance of the F-PNN model. The fireflies are operated and the sub-swarms are updated. When the iterations reach the max iterative number, the termination criterion satisfies and the optimal smoothing parameters of PNN model can be obtained. The procedure structure of the classification model is illustrated as Fig 6. This classification results in the detection of the brain tissues into tumor, normal, edema and necrosis categories. In the next stages, the tumor identified tissues are again classified to determine its type and stage.

ii) Neoplasm Classification

In this phase, the Cuckoo Search optimized Adaptive Neuro-Fuzzy Inference System (CS-NFS) [24] is utilized to classify the tumorous tissues into Benignant and Neoplastic. tumors and also determine the stage of tumor. Cuckoo search optimization [25] is based on the cuckoo egg laying procedure. For a maximization problem, the quality or fitness of a solution is directly proportional to the objective function. In CS algorithm the following simple representations that each egg in a nest represents a solution, and a cuckoo egg represent a new solution, the aim is to use the new and potentially better solutions (cuckoos) to replace a not-so-good solution in the nests. Of course, this algorithm can be extended to the more complicated case where each nest has multiple eggs representing a set of solutions. For this present work, the simplest approach is used where each nest has only a single egg. When generating new solutions, $CS_i(t+1)$ for a cuckoo i, a Levy flight is performed. [31]

$$CS_i(t+1) = CS_i(t) + \alpha \oplus Levy(\lambda) \qquad (12)$$

Here $\alpha > 0$ is the step size. It is associated to the scales of the ROI. The Levy flight essentially provides a random walk while the random step length is drawn from a Levy distribution [31]

$$Levy \sim u = t^{-\lambda}, \qquad (1 < \lambda \leq 3) \qquad (13)$$

Which has an infinite variance with an infinite mean. Here the steps essentially form a random walk process with a power law step-length distribution with a heavy tail. Some of the new solutions should be generated by Levy walk around the best solution obtained so far, this will speed up the local search. However, a substantial fraction of the new solutions should be generated by far field randomization and whose locations should be far enough from the current best solution, this will make sure the system will not be trapped in a local optimum.

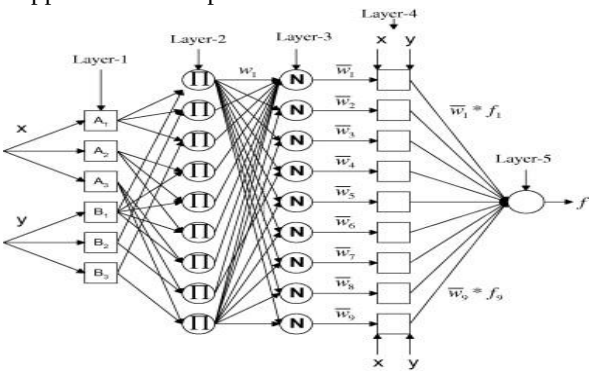

Fig.7 NFS Architecture

In the architecture of NFS process in Fig. 7, there are three learning parameters in NFS: the linguistic hedge (p_c) that influences the membership function value of the images, the consequent parameter (k_c) that enhances the performances value and the parameter (O_c) that represents the optimal rule in the NFS structure [26]. NFS is one nest that represented a potential solution where NFS has some parameters as a length of dimension. Denote

C_c is a nest, and let p_c denote the number of linguistic hedge parameters of each nest. Then an

algebraic representation of the length of dimension formed byp_c is illustrated below:[31]

$C_c = \{p_{cji} | i \in I; j \in J\}$ (14)

where I is a set of fuzzy sets and J is a set of inputs of each particle. Further, an algebraic representation of the length of the dimension formed by k_c is as: [31]

$C_c = \{k_{crj} | j \in J; r \in R\}$ (15)

where R is a set of rules that form CS-NFS. Furthermore, an algebraic representation of the $O_c$parameter, is as follows: [31]

$C_c = \{O_{cr} | r \in R\}$ (16)

Therefore the algebraic representation of each parameter in CS-NFS is given as [31]

$C_{cn} = \{(p_{cnji}, k_{cnrj}, O_{cnr}) | i \in I; j \in J; r \in R\}$(17)

where n = 1, 2,..., N and N are a number of nests. The proposed model of CS-NFS is given in the flowchart in Fig 8.

The membership function for the fuzzy part was been the bell function, with the most extreme value equivalent to one, and the base value equivalent to zero. The proposed CS-NFS hybrid algorithm can be utilized to decrease the likelihood of being trapped in local minima circumstance and furthermore improve accuracy and global search capacity of NFS training. In this manner the final tumor grouping can accomplished with greatest accuracy because of the utilization of enhanced NFS classifier. This results in the accurate detection of the tumor type (Benignant or Neoplastic.).

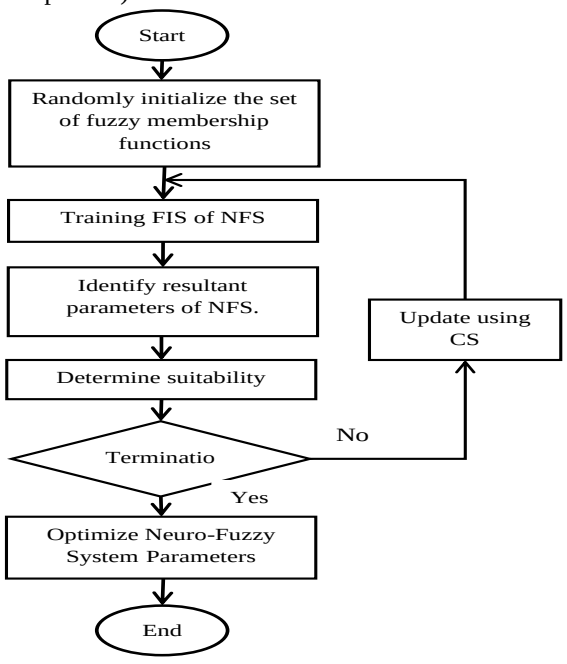

Fig.8 CS -NFS Flowchart

The total procedure engaged with CS-NFS can be outlined as takes after:

Step-1 Randomly generated the initial population of host nests in the search space, each representing a set of 36 numbers of antecedent parameters of NFS in sequence.

Step-2 A randomly nest (sequence of antecedent parameters) is generated; its objective function (classification accuracy) is determined and compared with that of randomly chosen nest, if the nest is better, its replace the chosen nest, otherwise it leaves intact.

Step-3 The worst nests are removed and replace them with new nests by Levy flight technique as well as by random generation.

Step-4 Solutions (sequence of antecedent parameters) are ranked and then the nest is updated again via Levy flight method and steps 2 to 4 are repeated until the CS algorithm cannot find any better parameters to reduce the error value.

Step-5.Train the optimal fuzzy parameters of NFS to calculate the suitable parameters. The parameters obtained from the CS algorithm are used to train the NFS.

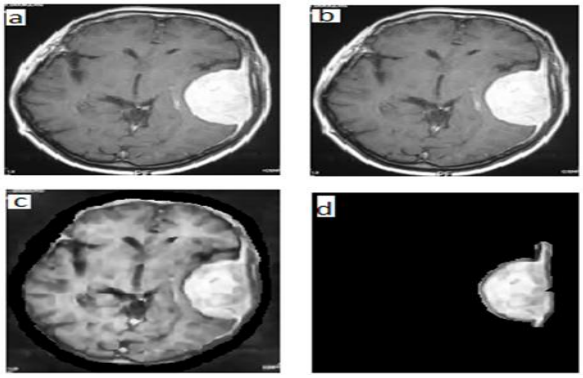

Fig. 9 a) Input (Neoplastic sample) b)Denoised output c) Brain Extraction d) Segmented output

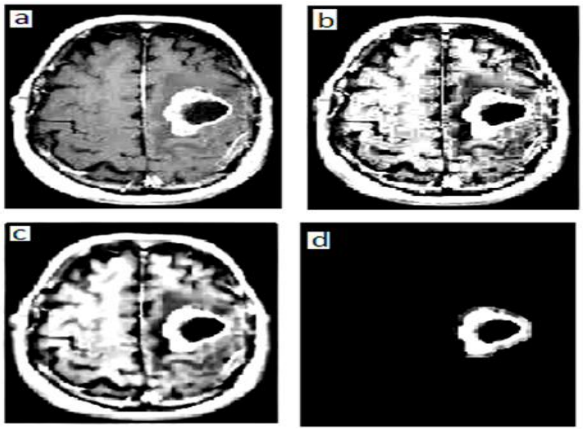

Fig. 10 a) Input (Benignant sample) b)Denoised output c) Brain Extraction d) Segmented output

Evaluation of Tumor Stage

Stage 0.This stage indicates the size of suspected tissue is normal and there is no tumor.

Stage 1. Size<0.5mm/inch.This is a separate group of tumors called juvenile pilocytic astrocytoma (JPA) and subependymoma. These are non-invasive and slow growing and can often be cured with surgery.

Stage 2. Size 1-4mm/inch.These tumors do not have actively dividing cells or dead cells in the tumor, called necrosis, but show many abnormal cells. A stage II tumor can be an astrocytoma, ependymoma, or oligodendroglioma.

Stage 3. Size 5-10mm/inch.These tumors are given a grade based on the cell type. For example, anaplastic astrocytoma is a stage III tumor that contains dividing cells but no dead cells. In contrast, anaplastic oligodendroglioma and anaplastic ependymoma are stage III tumors that do have dead cells.

Stage 4. Size>10mm/inch.A stage IV tumor is usually glioblastoma. Cells in the tumor are actively dividing. In addition, the tumor has blood vessel growth and areas of dead tissue.

RESULTS AND DISCUSSION

The input images are sourced from diverse platforms such as MR-TIP, NCIGT, BraTS, BITE, and TCIA, contributing to the evaluation of the proposed decision support system's efficacy in automatic brain tumor detection and classification. Utilizing MATLAB, simulations and feature extraction are conducted, preceded by pre-processing through ensemble genetic filtering and enhanced graph cut-based Brain extraction of sample images. Comparative analysis, depicted in Table 1, assesses the performance of the proposed filter against alternative filters based on metrics like MSE, Mean Absolute Error, PSNR, CNR, Image Quality Index (IQI), and Structured Similarity Index Measure (SSIM), demonstrating superior noise removal capabilities of the ensemble genetic filter. Similarly, Table 2 showcases the efficacy of improved graph cut-based Brain extraction compared to other methods, gauged through Dice similarity (DS) and Jaccard similarity (JS). The outcomes, exemplified in Fig 9 and Fig 10 for Neoplastic. and Benignant tumors respectively, highlight the accuracy of noise filtering, Brain extraction, and tumor segmentation. Segmentation clustering-based segmentation outperforms other methods, as evidenced in Table 3, contributing to enhanced tumor delineation.

Feature extraction, focusing on shape, texture, and intensity attributes via quantitative analysis techniques such as GLCM and HOG descriptors, is instrumental in achieving optimal results, as detailed in Table 4. The model's decision-making prowess, validated through expert medical interpretation, underscores the efficacy of single and multiple feature subsets, culminating in a 99.14% accuracy when incorporating Texture + Shape + Intensity features, as shown in Table 5.

Table 7 Implementation of the Classifier

Classes	No. Of data for training / testing	No. Of correctly classified data for NFS			Percentage of Correct Classification		
		Without Optimal Feature Subset Selection	With Optimal Feature Subset Selection	With Optimal Feature Subset Selection using CS-NFS	Without Optimal Feature Subset Selection	With Optimal Feature Subset Selection	With Optimal Feature Subset Selection using CS-NFS
Benignant	80/65	56	60	63	86.15	92.3	98.4
Neoplastic.	80/65	57	59	64	87.69	90.76	98.27
Average	-	-	-	-	87.03	91.55	98.3

Table 8 Accuracy Comparison of Recent Brain Tumor Analysis Methods

Work	Methods					Classification Accuracy (%)
	Denoising	Segmentation	Feature Extraction	Feature Selection	Classifier	
Rajalekshmi et at [27]	Weiner Filter	Colour convertedK-Means	Wrapper Approach	Principal component analysis	SVM	92.6
Wang et al [28]	Median Filter	-	Stationary Wavelet transform	Principal component analysis	HPA-FNN	96.43
Nazir et al [29]	Denoising Filters	-	Colour moment extraction	-	Feed forward ANN	97.65

Zhang et al [30]	-	-	Discrete wavelet packet transform	Tsallis Entropy	Generalized Eigenvalue Proximal SVM	97.77
Proposed Work	Ensemble Genetic Filter	Segmentation clustering	GLCM	-	F-PNN &CS-NFS	98.3

The parameters for F-PNN are outlined in Table 6, facilitating the classification system's robustness, evidenced by the high accuracy of 98.3% showcased in Table 7, surpassing many existing methods. The system's efficiency is further underscored by its reduced execution time, especially with the utilization of multiple feature subsets. Comparative analysis with recent brain tumor detection and classification methods, as presented in Table 8, reaffirms the superiority of the proposed model, boasting a 98.3% accuracy rate.

Upon identifying tumor malignancy, the system proceeds to determine tumor size based on the number of tumorous cells, aiding in staging and approximating tumor type, thus enhancing clinical decision-making.

CONCLUSION

The effectiveness of the proposed clinical decision-making system for automated brain tumor detection and classification has been confirmed through rigorous validation with diverse samples. Analysis of the results indicates that integrating ensemble genetic filtering and multiple feature subsets notably enhances tumor classification. The findings demonstrate the superior performance of the proposed model, achieving an accuracy of 98.3%. Moreover, this decision system not only assesses tumor stages but also predicts tumor types, thereby offering comprehensive diagnostic insights.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Human and animal rights and informed consent This article does not contain any studies with human participants or animals performed by any of the authors.

REFERENCE

[1] Iftekharuddin, K. M., Zheng, J., Islam, M. A., Ogg, R. J., andLanningham, F., Brain tumor detection in MRI: technique and statistical validation. Proceedings of Fortieth Asilomar Conference on Signals, Systems and Computers,pp. 1983-1987, 2006.

[2] Kharrat, A., Benamrane, N., Messaoud, M. B., and Abid, M., Detection of brain tumor in medical images. Proceedings of 3rd International Conference on Signals, Circuits and Systems (SCS). pp. 1-6,2009.

[3] Wei, J., Hu, X. and Liu, W., An improved authentication scheme for telecare medicine information systems. *Journal of medical systems*, *36*(6):3597-3604, 2012.

[4] Al-Ariki, H. D. E., and Swamy, M. S., A survey and analysis of multipath routing protocols in wireless multimedia sensor networks. *Wireless Networks*, *23*(6):1823-1835, 2017.

[5] El-Dahshan, E.S.A., Mohsen, H.M., Revett, K. and Salem, A.B.M., Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm. *Expert systems with Applications*, *41*(11):5526-5545 2014.

[6] Ulku, E.E. and Camurcu, A.Y., November. Computer aided brain tumor detection with histogram equalization and morphological image processing techniques. Proceedings ofInternational Conference on Electronics, Computer and Computation (ICECCO), pp. 48-51, 2013.

[7] Gordillo, N., Montseny, E. and Sobrevilla, P., State of the art survey on MRI brain tumor segmentation. *Magnetic resonance imaging*, *31*(8):1426-1438, 2013.

[8] Abdel-Maksoud, E., Elmogy, M. and Al-Awadi, R., Brain tumor segmentation based on a hybrid clustering technique. *Egyptian Informatics Journal*, *16*(1):71-81, 2015.

[9] Amien, M.B., Abd-elrehman, A. and Ibrahim, W., An Intelligent-Model for Automatic Brain-Tumor Diagnosis based-on MRI Images. *International Journal of Computer Applications*, *72*(23):21-24, 2013.

[10] Devasena, C.L. and Hemalatha, M., Efficient computer aided diagnosis of abnormal parts detection in magnetic resonance images using hybrid abnormality detection algorithm. *Central European Journal of Computer Science*, *3*(3):117-128, 2013.

[11] Othman, M.F. and Basri, M.A.M., 2011, January. Probabilistic neural network for brain tumor classification. Proceedings of Second International Conference on Intelligent Systems, Modelling and Simulation (ISMS), pp. 136-138, 2011.

[12] Latif-Amet, A., Ertüzün, A. and Erçil, A., An efficient method for texture defect detection: sub-band domain co-occurrence matrices. *Image and Vision computing*, *18*(6):543-553, 2000.

[13] Arakeri, M.P. and Reddy, G.R.M., Computer-aided diagnosis system for tissue characterization of brain tumor on magnetic resonance images. *Signal, Image and Video Processing*, *9*(2):409-425, 2015.

[14] Dandıl, E., Çakıroğlu, M. and Ekşi, Z., Computer-aided diagnosis of malign and Benignant brain tumors on MR images. In ICT Innovations 2014 (pp. 157-166). Springer, Cham, 2015.

[15] Patil, R.C. and Bhalchandra, A.S., Brain tumour extraction from MRI images using MATLAB. *International Journal of Electronics, Communication and Soft Computing Science & Engineering (IJECSCSE)*, *2*(1): 1-4, 2012.

[16] Motwani, M.C., Gadiya, M.C., Motwani, R.C. and Harris, F.C., Survey of image denoising techniques. Proceedings of GSPX. pp. 27-30, 2004.

[17] Park, S., Hwang, J., Rou, K. and Kim, E., A new particle filter inspired by biological evolution: Genetic filter. *World Academy of Science, Engineering and Technology*, *33*:83-87, 2007.

[18] Wang, T. and Li, X., An efficient impulse noise reduction algorithm. Proceedings of International Conference on Multimedia Technology (ICMT), pp. 164-167, 2011.

[19] Sadananthan, S.A., Zheng, W., Chee, M.W. and Zagorodnov, V., Brain extraction using graph cuts. *NeuroImage*, *49*(1):225-239, 2010.

[20] Fu, L. and Medico, E., Segmentation, a novel fuzzy clustering method for the analysis of DNA microarray data. *BMC bioinformatics*, *8*(1):3-18, 2007.

[21] Jafarpour, S., Sedghi, Z. and Amirani, M.C., A robust brain MRI classification with GLCM features. *Int. J. Comput. Appl*, *37*(12):1-5, 2012.

[22] Dalal, N. and Triggs, B., Histograms of oriented gradients for human detection. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1, pp. 886-893, 2005.

[23] Yang, X.S., Firefly algorithms for multimodal optimization. Proceedings of International symposium on stochastic algorithms, pp. 169-178, 2009.

[24] Mohanty, P.K. and Parhi, D.R., 2015. A new hybrid optimization algorithm for multiple mobile robots navigation based on the CS-NFS approach. *Memetic Computing*, *7*(4), pp.255-273.

[25] Yang, X.S. and Deb, S., Cuckoo search via Lévy flights. Proceedings of *World Congress on Nature & Biologically Inspired Computing*, pp. 210-214, 2009.

[26] Çaydaş, U., Hasçalık, A. and Ekici, S., An adaptive neuro-fuzzy inference system (NFS) model for wire-EDM. *Expert Systems with Applications*, *36*(3):6135-6139, 2009.

[27] Rajalakshmi, N. and Prabha, V.L., Automated Classification Of Brain MRI Using Color converted K-Means Clustering Segmentation And Application Of Different Kernel Functions With Multi-Class SVM. *European Scientific Journal, ESJ*, *9*(21), 2013.

[28] Wang, S., Zhang, Y., Dong, Z., Du, S., Ji, G., Yan, J., Yang, J., Wang, Q., Feng, C. and Phillips, P., Feed-forward neural network optimized by hybridization of PSO and ABC for abnormal brain detection. *International Journal of Imaging Systems and Technology*, *25*(2):153-164, 2015.

[29] Nazir, M., Wahid, F. and Ali Khan, S., A simple and intelligent approach for brain MRI classification. *Journal of Intelligent & Fuzzy Systems*, *28*(3):1127-1135, 2015.

[30] Zhang, Y., Dong, Z., Wang, S., Ji, G. and Yang, J., Preclinical diagnosis of magnetic resonance (MR) brain images via discrete wavelet packet transform with Tsallis entropy and generalized eigen value proximal support vectormachine (GEPSVM). *Entropy*, *17*(4):1795-1813, 2015

[31] Dr.M.suganthi, Dr.Rupa Ezhilarasi,'A Clinical Support System for Brain Tumor Classification Using Soft Computing Techniques', Journal of

Medical Systems, Springer, vol.43, no.5, May 2019, DOI 10.1007/s10916-019-1266-9