IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 4 September 2023, accepted 30 September 2023, date of publication 6 October 2023, date of current version 18 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3322450

RESEARCH ARTICLE

MRI Brain Tumor Detection Methods Using Contourlet Transform Based on Time Adaptive Self-Organizing Map

ALI FARZAMNIA[1], (Senior Member, IEEE), SEYED HAMIDREZA HAZAVEH[2], SEYEDE SAFIEH SIADAT[3], AND ERVIN GUBIN MOUNG[4]

[1]Faculty of Engineering, Universiti Malaysia Sabah, Kota Kinabalu, Sabah 88400, Malaysia
[2]Faculty of Mechanical, Electrical Power and Computer, Science and Research Branch, Islamic Azad University (IAU), Tehran 1477893855, Iran
[3]Faculty of Computer Science, Payame Noor University (PNU), Tehran 193954697, Iran
[4]Faculty of Computing and Informatics, Universiti Malaysia Sabah, Kota Kinabalu 88400, Malaysia

Corresponding authors: Seyed Hamidreza Hazaveh (r.hazaveh@gmail.com) and Ali Farzamnia (alifarzamnia@ums.edu.my)

This work was supported by the Penerbit Universiti Malaysia Sabah (UMS); and the Faculty of Engineering, UMS.

ABSTRACT The brain is one of the most complex organs in the body, composed of billions of cells that work together to ensure proper functioning. However, when cells divide in a disorderly manner, abnormal growths can occur, forming colonies that can disrupt the normal functioning of the brain and damage healthy cells. Brain tumors can be classified as either benign or low-grade (grade 1 and 2), or malignant or high-grade (grade 3 and 4). In this article, we propose a novel method that uses contourlet transform and time adaptive self-organizing map, optimized by the whale optimization algorithm, in order to distinguish between benign and malignant brain tumors in MRI images. Accurate classification of these images is critical for medical diagnosis and treatment. Our method is compared to other methods used in past research and shows promising results for the precise classification of MRI brain images. Through conducting experiments on different test samples, our system has successfully attained a classification accuracy exceeding 98.5%. Furthermore, it has managed to maintain a satisfactory level of efficiency in terms of run-time.

INDEX TERMS Brain, classification, contourlet transform, detection, neural networks, time adaptive self-organizing map (TASOM), tumor, whale optimization algorithm (WOA).

NOMENCLATURE

MRI	Magnetic Resonance Imaging.
GLCM	Gray-Level Co-occurrence Matrix.
PCA	Principal Component Analysis.
GA	Genetic Algorithm.
ROC	Receiver Operating Characteristic.
SD	Standard Deviation.
W	Weight vector.
N	Number of Neurons in the network.
α, β	Controlling local neighborhood errors.
S_f, S_g	Controlling compression and topological neuron orders.
DUW, *DWD*	Controlling classification's accuracy and speed.
RBF	Radial Basis Function.
IDM	Inverse Difference Moment.
FCM	Fuzzy C-Means.
GNB	Gaussian Naïve Bayes.

The associate editor coordinating the review of this manuscript and approving it for publication was Zhen Ren.

I. INTRODUCTION

The brain is a highly complex organ that comprises billions of cells, and abnormal growth occurs when cells divide uncontrollably, forming a group of abnormal cells within or around the brain. This can impair normal brain function and destroy healthy cells. Brain tumors are classified as benign or low-grade (grade I and II) and malignant or high-grade (grade III and IV). While benign tumors are considered less

aggressive and grow slowly, malignant tumors are cancerous and grow rapidly with undefined boundaries [1]. Magnetic resonance imaging (MRI) is a powerful tool for detecting brain tumors and modeling their progress [2]. MRI images provide helpful information about brain structure and abnormalities, and have been widely used in automated approaches for brain tumor detection and classification. Artificial Neural Networks have been the most popular classification approach in this field for several decades. The time adaptive self-organizing map (TASOM) network is a modified version of the self-organizing map (SOM) network, which incorporates adaptive learning rates and neighborhood sizes as its learning parameters. In TASOM, each neuron has its own unique learning rate and neighborhood size. When presented with a new input vector, the learning rate and neighborhood size of the winning neuron, along with the learning rates of its neighboring neurons, are updated accordingly. The TASOM algorithm also employs a scaling vector to compensate for scaling transformations.

An examination of the updating rules in the algorithm reveals that the learning parameters can dynamically increase or decrease to adapt to a changing environment. This adaptation process aims to achieve the minimum increase or decrease according to a specific measure. This network tailored for different applications, such as the classification of MRI brain Tumor images. It is expected to become the state of the art in various health informatics areas such as bioinformatics, medical informatics, and medical image analysis.

The time adaptive self-organizing map (TASOM) model, optimized by the whale algorithm, is an exciting trend in machine learning for the efficient classification of brain tumors. This method has the potential to assist the medical sector in promptly diagnosing serious diseases. In image classification, classical methods are still important compared to modern methods like Convolutional Neural Networks (CNN) for several reasons:

1) **Suitability for Small Datasets:** Classical methods can be more effective when dealing with small datasets. Deep learning models like CNNs usually require large amounts of data for training, while classical methods can yield satisfactory results with smaller datasets.
2) **Domain-Specific Applications**: In some domain-specific tasks or specific scenarios, classical methods may outperform CNNs due to their tailored design for the particular problem at hand.
3) **Pre-trained Models Not Always Available:** Training deep learning models like CNNs from scratch requires extensive labeled data and substantial computational resources. In cases where pre-trained models are not available, classical methods can serve as viable alternatives.
4) **Reduced Overfitting Risk**: Classical methods may be less prone to overfitting, especially when dealing with limited training data. CNNs, with their large number of parameters, have a higher risk of overfitting when the training data is insufficient

This paper proposes a methodology using TASOM neural networks, our work involves identifying specific features or patterns in brain images that can aid in the detection and classification of tumors [3]. These features are typically extracted from the brain images using various techniques, such as contourlet transform and the Otsu method. The extracted features are then used to train TASOM classifier in order to detect cancerous brain tumors. This neural network is trained on a dataset of brain images that have been labeled as either benign or malignant. Once the neural network has been trained, it can be optimized using the Whale Algorithm to classify new brain images into these categories based on the identified features. This can be done in real-time, making it a valuable tool for early-stage abnormality detection and tumor diagnosis [4], [5], [6].

Image segmentation involves dividing an image into several distinct segments or regions of interest. The primary aim of image segmentation is to streamline or alter the way an image is represented, making it more interpretable and conducive to analysis. In this work, we utilized the Otsu method for image segmentation. The Otsu method is a widely used algorithm employed in this field.

In this paper, we aim to address the following contributions:

- Development of a novel tumor detection method.
- Improved accuracy in tumor classification.
- Efficient real-time diagnosis.
- Applicability to health informatics.

Over the last few years, numerous researchers have contributed to the field by publishing their works, presenting diverse methodologies, and exploring various approaches for the purposes of detection and classification in different domains. Table 1 demonstrates a wide array of techniques that have been suggested for automatic brain MRI classification, encompassing both classical machine learning and deep learning methods.

Our dataset was downloaded from the Kaggle website, namely Brain Tumor Detection. The database comprises 1000 images as shown some samples in Fig.1, out of which 500 images contain tumors including benign while the remaining 500 images are malign. In this work, we assign 70% of the dataset as training data and the remaining 30% as test data. As indicated in Table 2, all the tests were conducted under the mentioned conditions.

II. TASOM AND ITS CLASSIFICATION ALGORITHM

TASOM is one of the most robust neural networks for data classification applied in our research. In this study, the capability of this tool in data classification is explicitly assessed compared to other machine learning tools. SOM networks are based on competition between neurons, where neurons compete with each other in processing input data. The winning neurons and their neighboring peers are then activated by the input. The neighborhood function of the neurons is defined by a set or function that specifies their connections [17], [30].

TABLE 1. Recent researches focused on brain tumor classification.

No.	Author	Year	Type of Solution	Classification Method	Feature Extraction Method	Accuracy
1	Rajan and Sundar [7]	2019	Classical Machine Learning	Support Vector Machine (SVM)	Adaptive Gray-Level Co-Occurrence Matrix (AGLCM)	98%
2	Shree and Kumar [8]	2018	Classical Machine Learning	Probabilistic Neural network (PNN)	Gray level Co-Occurrence Matrix	95%
3	Z. Ullah [9]	2020	Classical Machine Learning	Feed-forward Neural Network	DWT	95.8%
4	B. Ural [10]	2018	Classical Machine Learning	Probabilistic Neural network (PNN)	K-mean with Fuzzy c-mean	90%
5	J. Francisco [11]	2021	Advanced Deep Learning	Multi-pathway Convolutional Neural Network (CNN)	CNN	97.3%
6	Ahmet and Mohammad [12]	2020	Advanced Deep Learning	CNN	CNN	97.2%
7	Hemanth [13]	2019	Advanced Deep Learning	CNN	CNN	94.5%
8	Saxena [14]	2019	Advanced Deep Learning	CNN Networks with transfer learning	CNN	95%
9	N. M. Dipu [15]	2021	Advanced Deep Learning	YOLOv5 model (CNN)	cutting-edge object detection	96%
10	Alok Sarkar [16]	2022	Advanced Deep Learning	Alex Net CNN	CNN	98%

TABLE 2. Experimental conditions.

Specification	VALUES
CPU	Intel Core i7
Speed	3.0 GHZ
RAM	16 GB
Image Size	256 × 256
Image Type	Grey Levels
Technology	MRI

During network training, the neighborhood gradually shrinks, and the neurons' learning rate decreases to prepare the network for convergence and to minimize error. However, these small parameters at the end of network training may reduce the network's ability to handle new data in variable and dynamic environments.

TASOM modifies the network's parameters for each input data to overcome this problem, and each neuron has a different learning rate, which is dynamically updated according to a specific criterion. The neurons' neighborhood functions are also automatically updated, and a winner neuron is identified for each new input, and only the winner neurons' neighborhood function is updated. In the TASOM network, a scale vector is used if the input vector distribution is asymmetric to maintain topological order and minimize quantization error. The uniform vector is used if consistency is essential, and topological order has less significance, allowing the network to adapt rapidly to changes in input vectors.

A. TASOM ALGORITHM FOR LEARNING THE RATES AND SIZE OF ADAPTIVE NEIGHBORHOOD

TASOM algorithm, with its consistent learning rates and neighborhood sets, can be summarized in the following seven steps:

Step (1) Initialization: Choose a value for initial weight vectors $\mathbf{w_j(0)}$ for $j = 1, 2, \ldots, N$ where "N" is the number of neurons in the network. A value close to 1 must be considered as parameter $\eta_{\mathbf{j}}(\mathbf{n})$. α, β, α_s and β_s parameters can have values between 0 and 1.

The constant parameters S_f and S_g should be set to satisfy the application's needs. "$\mathbf{R_i(0)}$" should be set to include

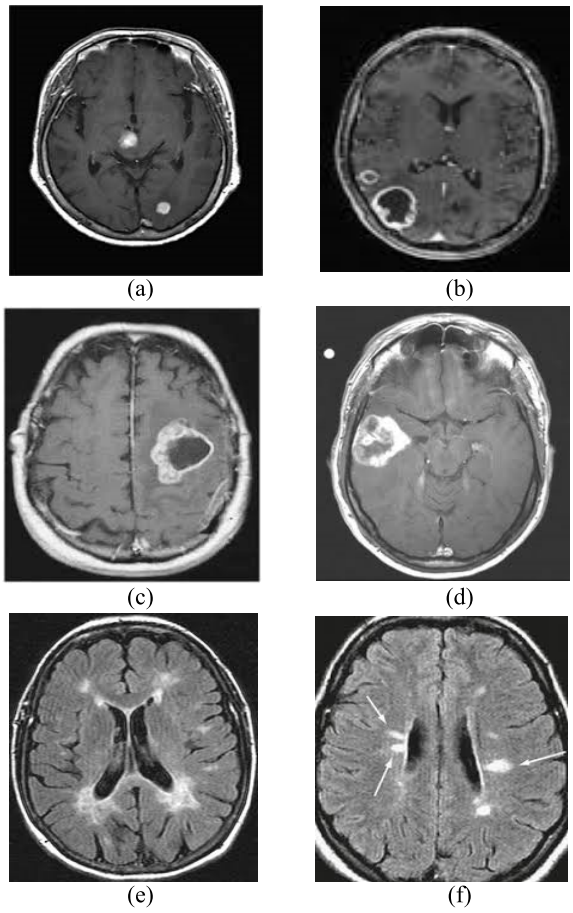


FIGURE 1. The types of brain MRI images in Kaggle dataset: benign tumor (a) and (b), malign tumor (c) and (d), normal brain (e) and (f).

all the neurons. The components $S_k(0)$ of the scaling vector $\mathbf{S(0)= [\ s_1(0), \ldots, s_p(0)]^T}$ should be set to small positive values, where p is the dimension of the input and weight vectors. The parameters $\mathbf{E_k(0)}$ and $\mathbf{E2_k(0)}$ are initialized with some small random values. "DUW" and "DWD" control each clustering accuracy. These two parameters indicate the maximum and minimum distances that two adjacent neurons can have. Selecting small values for both of these parameters increases clustering accuracy and increases the number of neurons.

In this paper, neighboring neurons of each neuron "i" in a network include "LN_i"set. For each neuron "i" in a one-dimensional network $LN_i = \{i-1, i+1\}$, and similarly for each neuron (i_1, i_2) in a 2-dimensional network.

$$LN_{(i_1,i_2)} = \{(i_1-1,i_2)\,,(i_1+1,i_2)\,,(i_1,i_2-1)\,,(i_1,i_2+1)\} \tag{1}$$

$$i\,(x) = arg_j \min \left|\left|X\,(n) - W_j\,(n)\right|\right|_S \quad j = 1, 2, \ldots, N \tag{2}$$

$$||X(n) - W_j(n)||_S = (\sum_k (x_k(n) - w_{j,k}(n))/s_k(n)^2)^{\frac{1}{2}} \tag{3}$$

Step (2) Sampling: Obtain the next input vector "$\boldsymbol{X}$" from the input distribution.

Step (3) Similarity Matching: Find winner **"i(x)"** neuron at time n using a minimum Euclidian distance function scaled by the **"S(n)**" function.

Step (4) Updating Neighborhood Size: A neighborhood set of winner neuron **"i(x)"** is updated with the following equations:

$$\Lambda_i\,(n+1) = \{j \in N | d(i,j) \le R_i(n+1)\} \tag{4}$$

$$R_i(n+1) = R_i(n) + \beta(g(\frac{1}{|NH_i|\,.S_g} \sum_{j \in NH_i} ||w_i(x) - w_j(x)||_S) - R_i(x))NH_i = \{i-1, i+1\} \tag{5}$$

where "β" is a fixed parameter between "0" and "1" and specifies the speed employed, by which neighborhood size can control local neighborhood errors and "S_g" controls the compression and topological order of neuron weights. The neighborhood set of other neurons does not change. The "g" function is an ascending scalar function used for normalizing the weight distance vector. The weights are updated as follows:

$$W_j\,(n+1) = \begin{cases} W_j(n) + \eta_j(n+1)\left[X\,(n) - W_j\,(n)\right], \\ \qquad\qquad j \in \Lambda_{i(x)}(n+1) \\ W_j(n) \qquad otherwise \end{cases} \tag{6}$$

Step (5) Updating Learning Rate: In the neighborhood set **"$\Lambda_{\mathbf{i(x)}}$(n+1)",** learning rate **"$\boldsymbol{\eta}_\mathbf{j}$(x)"** from the winner neuron **"i(x)"** is updated as the following equation:

$$\eta_j(n+1) = \eta_i(n) + \alpha(f||x(n) - w_j(n)||_S/s_f) - \eta_j(n) \quad j \in \Lambda\_i(x)(n+1) \tag{7}$$

Another neuron's learning rate does not change anymore. "f" function is a monotonous or increasing function used for the normalizing distance between weight and input vectors.

Step (6) Updating the Scale Vector:

Set $\mathbf{S(n)=[S_1(n),\ldots,S_P(n)]^T}$, adjusted according to the following equations, where E represents the expected value:

$$S_k\,(n+1) = \sqrt{E2_k(n+1) - E_k(n+1)^2} \tag{8}$$

$$E2_k\,(n+1) = E2_k\,(n) + \alpha_s\left(x_k^2\,(n) - E2_k\,(n)\right) \tag{9}$$

$$E_k\,(n+1) = E_k\,(n) + \beta_s\,(x_k\,(n) - E_k\,(n)) \tag{10}$$

This scale vector is non-uniform. The uniform vector is defined as follows:

$$S\,(n+1) = \sqrt{\sum_k S_k^2\,(n+1)} \tag{11}$$

Step (7) Return to Step (2).

B. TASOM ALGORITHM FOR CLASSIFICATION

The data classification algorithm initially works with two neurons [3], [18]. The input of more data from the environment causes neurons to be added to the network, and idle neurons will be removed from the network. The algorithm can be summarized in the following steps:

Step (1) An adaptive network is made in the form of an open linear topology with two neurons.

Step (2) The network is trained with the number of K × C input data vectors, where "C" is the number of neurons present in the network and "K" is a constant number trains this number of input vectors and forms an iteration.

Step (3) Any neuron which does not win after one complete iteration is removed from the network.

Step (4) If the distance between two neighboring neurons $||W_i - W_{i+1}||$ is greater than the "DUW" parameter, then a new neuron j is inserted between the two neurons. The parameters of new neuron equal:

$$\frac{w_i + w_{i+1}}{2}, \quad \eta_j = \frac{\eta_i + \eta_{i+1}}{2} \tag{12}$$

Step (5) Every two neighboring neurons "i" and "$i+1$" where the value $||W_i - W_{i+1}||$ is less than "DWD" parameter are replaced with a new neuron "j", weight vector, and the learning parameters derived from Equation (12).

Step (6) Network training proceeds to Step (2).

Any neuron having no idle neuron next to itself makes up one class. This algorithm utilizes a modular adaptive network with a chain-like topology. The adaptive network starts its work with only two neurons as an initial setup. Then, with the introduction of more data from the environment, additional neurons can be added to the network. If the number of neurons in the network reaches the maximum allowed, the algorithm will stop, and no new neurons will be added to the network. This limitation is in place to control the size and complexity of the network.

One of the most critical applications of this competitive network is data clustering and classification [3]. In this research, as indicated in Table 3, TASOM parameters, include the maximum values for the neuron parameter "K", DUW, DWD, the number of primary neurons, and S_f, S_g are 100, 100, 0.05, 0.0005, 2, 0.5, and 0.5, respectively.

TABLE 3. The tasom parameters for the classification.

Parameters	VALUE
Maximum Neurons	100
Number of Neurons	2
K	100
DUW	0.05
DWD	0.0005
Sf	0.5
Sg	0.5
Epoch	10

III. METHODOLOGY

Methodology in data classification involves the systematic process of categorizing or grouping data into different classes or categories, based on their characteristics or attributes. The goal is to develop models or algorithms that can accurately assign new unseen data instances to the appropriate classes. Fig.2 indicates the proposed technique consists of three main steps:

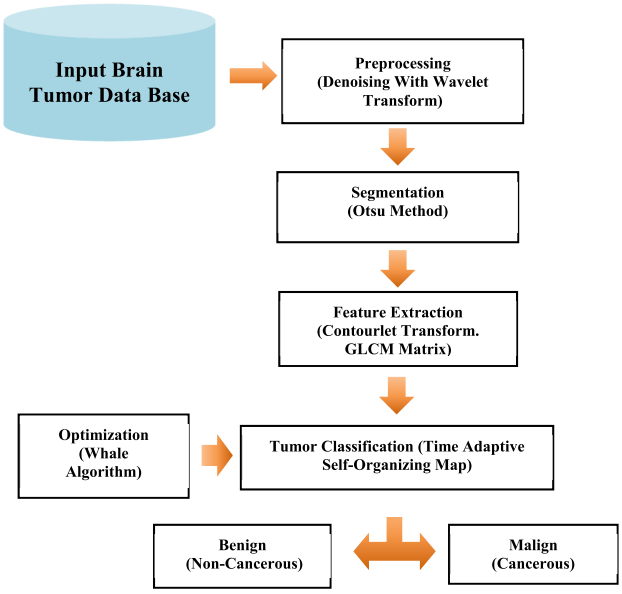


FIGURE 2. The flowchart indicates the process of brain tumor classification.

Step 1: Involves preprocessing, which includes denoising using wavelet transform, followed by feature extraction using contourlet transform, and feature reduction through the use of PCA.

Step 2: The classifiers are trained using various machine-learning methods, such as TASOM.

Step 3: Submitting new MRI brain images (training sets) to the trained classifiers and obtaining the prediction output [19].

A. PREPROCESSING

Image preprocessing is a crucial step in image analysis and computer vision tasks. It involves applying a series of operations and techniques to enhance the quality of images, remove noise or artifacts, and prepare them for subsequent analysis or processing. Image preprocessing plays a vital role in improving the accuracy and reliability of image-based algorithms, such as object detection, image classification, and image segmentation. The first step in our preprocessing chain is the denoising of medical images using wavelets. In such images, edges represent places where the image brightness changes rapidly. Therefore, maintaining edges while denoising an image is crucial for preserving image quality [20]. Traditional low-pass filtering removes noises although it results in the reduction of edge sharpness and has a detrimental effect

on the quality of the image. Wavelets, on the other hand, are capable of removing noise while maintaining important features. In this work, we used "Daubechies" wavelet. This transform is the most popular transform that leads to the foundation of wavelet-based multi-dimensional signal processing. Our input images are decomposed into three levels and we apply a soft thresholding technique to the wavelet coefficients to remove noise.

We have also implemented Principal Component Analysis (PCA) to obtain only the necessary features for optimal processing [21]. PCA reduces the dimensionality of the predictor space, creating classification models that prevent overfitting. PCA linearly transforms predictors to remove redundant dimensions and generates a new set of variables called principal components [22], [23].

B. SEGMENTATION FOR FEATURE EXTRACTION

Segmentation is a process in data analysis and feature extraction that involves partitioning an input signal, image, or dataset into meaningful or homogeneous regions or segments. It aims to identify and extract distinct patterns, objects, or regions of interest from the data. Segmentation for feature extraction is particularly useful when dealing with complex data such as images, where different regions or objects may have varying characteristics and need to be analyzed separately. By segmenting the data, we can isolate and extract relevant features from each segment, which can then be used as inputs for classification or further analysis.

1) OTSU METHOD

The Otsu method is a popular algorithm used for image segmentation, which is the process of partitioning an image into multiple segments or regions of interest. The goal of image segmentation is to simplify or change the representation of an image into something more meaningful and easier to analyze. The Otsu method is a thresholding technique that determines the optimal threshold value that separates the pixels of an image into foreground and background. The algorithm calculates the variance between two classes of pixels, which is the sum of the variances within each class. The threshold value that minimizes the variance between the two classes, is chosen as the optimal threshold value [24]. The Otsu method is a simple and effective technique for thresholding and has been used in various applications, including object recognition, face recognition, and medical image analysis [25], [26].

The definition of an image with threshold "g (x, y)" is given by equation (13):

$$g(x, y) = \begin{cases} 1, & f(x, y) \geq T \\ 0, & f(x, y) \leq T \end{cases} \tag{13}$$

where "f(x,y)" is the intensity of the input image at point "(x,y)", "T" is the threshold value, and "g(x,y)" is the thresholding result.

This result is a binary image, where pixels with an intensity value of "1" correspond to objects, and pixels with a value of "0" correspond to the background. The Fig.3 Shows the value of optimal threshold "T" is 100.

Finally, we applied the Otsu algorithm [27] for image segmentation. The segments, fewer than 80 pixels, were omitted. The results are presented in Fig.4.

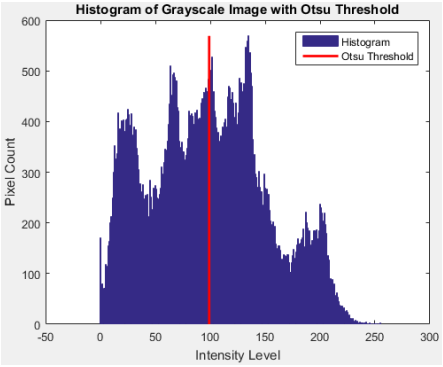


FIGURE 3. Histogram using OTSU threshold.

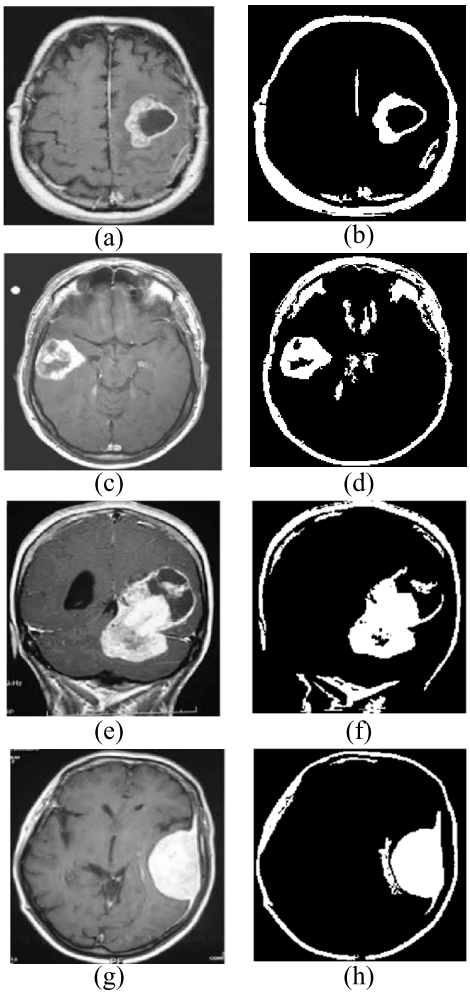


FIGURE 4. Image segmentation by OTSU method, (A), (C), original images of benign tumors, (B), (D) segmented images of benign tumors, (E), (G) original images of malign tumors, (F), (H) segmented images of malign tumors.

2) FEATURE EXTRACTION USING CONTOURLET TRANSFORM

Feature extraction using the contourlet Transform is a technique that combines the benefits of wavelet transform and directional filter banks to extract meaningful features from images or signals. The contourlet transform is specifically designed to capture both the local and directional information present in the data, making it suitable for applications such as image analysis, object recognition, and texture classification. The contourlet transform is composed of basic functions with different directions in multiple scales and flexible aspect ratios. This framework forms a basis with small redundancy, unlike other transforms. The basic elements of the transform are oriented in various directions, much more than the few directions offered by other separable transform techniques [28], [29], [30]. The contourlet transform is a discrete extension of the curvelet transform that aims to capture curves instead of points and provides directionality. contourlets possess not only the main features of wavelets (multiscale and time-frequency localization) but also offer a high degree of directionality and anisotropy. The fundamental difference between contourlets and other multiscale directional systems is that the contourlet transform takes a different and flexible number of directions at each scale while achieving nearly critical sampling. In addition, the iterated filter banks used in this method make it computationally efficient, requiring only ''O(N)'' operations for an N-pixel image.

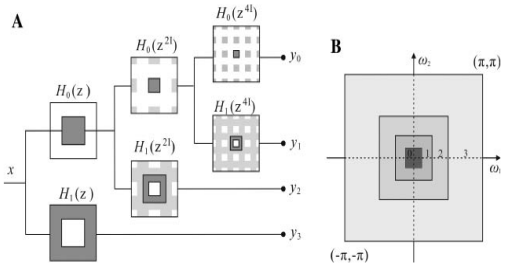


FIGURE 5. (A) Three-stage pyramid decomposition. (B) sub-bands on the 2-D frequency plane.

The Fig.5 illustrates the structural design of the contourlet transform via Laplacian pyramid and directional filter banks which decompose the signal is as follows [17]:

- The input image is decomposed into four frequency components: LL (Low Low), LH (Low High), HL (High Low), and HH (High High).
- At each level, the Laplacian pyramid produces a low-pass output (LL) and a band-pass output (LH, HL, HH).
- The band-pass output is then passed through the directional filter bank, resulting in contourlet coefficients.
- The low-pass output is again passed through the Laplacian pyramid to obtain more coefficients, and this process is repeated until the fine details of the image are retrieved.
- The image is then reformed by applying the inverse contourlet transform.

In our work, we assume that our function ''pdfbdec'' will decompose an image using a pyramidal directional filter bank (PDFB) with a ''7-9'' filter for the pyramidal decomposition stage and a ''pkva'' filter for the direction decomposition stage. There will be four levels of pyramidal decomposition and the numbers of directional decompositions at each pyramidal level (from coarse to fine) are 0, 3, 3, and 4. In Table 4, the results, obtained from applying contourlet transform, are indicated.

TABLE 4. Results from experiments, (1) Applying contourlet in tasom, (2) Without using contourlet.

NO.	Experiment	Accuracy	Recall	MSE	General Time(s)
1	1	0.979	0.983	0.0244	47
	2	0.973	0.967	0.0299	41
2	1	0.988	0.984	0.0218	46
	2	0.971	0.975	0.0414	42
3	1	0.986	1.000	0.0328	49
	2	0.969	0.988	0.0409	43
4	1	0.990	1.000	0.0369	46
	2	0.976	1.000	0.0343	41
5	1	0.992	0.984	0.0235	48
	2	0.985	0.972	0.0322	44

TABLE 5. The extracted statistical feature values from malignant brain tumor images are obtained by applying contourlet transform after 100 iterations.

Feature	VALUE
Mean	0.004
SD	0.089
Entropy	3.6
RMS	0.089
Variance	0.008
Smoothness	0.94
Kurtosis	6.013
Skewness	0.526
IDM	0.38
Contrast	0.225
Correlation	0.134
Energy	0.746
Homogeneity	0.929

After calculating the GLCM matrix, we then extract our features. GLCM approach is a popular method for image applications, particularly for gray level co-occurrence matrix analysis, as it is considered one of the most effective methods available. In the field of object classification, there are a variety of feature extraction techniques used, such as GLCM and K-nearest neighbor (KNN). The GLCM approach is commonly employed for MRI image feature extraction, given its high level of accuracy, and is frequently used in the analysis of texture features [31], [32]. Textural analysis can be useful in enhancing the detection and grading of tumors. Overall, as shown in Table 5, 13 different statistical features can be

extracted from the image using the Contourlet transform: contrast, correlation, energy, homogeneity, mean, standard deviation, entropy, RMS, variance, smoothness, kurtosis, skewness and IDM [33].

3) WHALE OPTIMIZATION ALGORITHM (WOA)

The Whale Optimization Algorithm (WOA) is a nature-inspired metaheuristic optimization algorithm that is based on the social behavior of humpback whales. The Whale Optimization Algorithm has been applied to various optimization problems, including engineering design, data clustering, feature selection, image segmentation, and neural network training. We apply this algorithm to enhance our results in the experiments. In WOA, one individual whale's position is considered as the target, and other whales try to update their positions towards it [18]. This is done by simultaneously swimming around the target within a shrinking circle and along a spiral shaped path, with a 50% probability of choosing between the two models to update their positions during optimization. When the random variable "p" is smaller than "0.5", whale individuals update their positions by the shrinking encircling model (14)-(16):

$$X_{i(t+1)} = X^*(t) - A|CX^*(t) - X_i(t)| \quad (14)$$

$$A = a(2r - 1) \quad (15)$$

$$C = 2r \quad (16)$$

where **"Xi(t)"** represents the "ith" individual in the whale population **{Xi(t)} (i = 1, 2, …, SP), X*(t)** represents the best individual with the smallest fitness observed so far, "SP" represents the population dimension, and "t" indicates the current iteration ranging from "1" to the maximum iteration number "M". "a" is a number that decreases with "t" during iterations, and "r" is a random number between "0" and "1". Furthermore, when $|A|$ is larger than 1, the positions of whale individuals are updated based on a randomly chosen individual **"X_{rand}(t)"** using equation (17), instead of using the best individual **"X*(t)"**. This approach reduces the chances of converging to a local optimum and enables the Whale Optimization Algorithm (WOA) to conduct a global search.

$$X_i(t + 1) = X_{rand}(t) - A|CX_{rand}(t) - X_i(t)| \quad (17)$$

The shrinking encircling behavior is achieved by decreasing the value of "a", and "A" is a random number within $[-a, a]$ according to equation (15). The adaptive variation of "A" allows WOA to smoothly transit between the convergence and the capacity of searching for global optimal solution: some iterations are devoted to search for the global optimal solution ($|A| \geq 1$) and the rest is dedicated to convergence ($|A| < 1$). In terms of "SP" and "M", "SP" provides a trade-off between computational costs and optimal results found by WOA, and "M" is a number large enough to guarantee the convergence of algorithm. When "p" is not smaller than "50%", whale individuals update their positions by the spiral model. The spiral equation (18) is created between positions of whales and the target whale to mimic the spiral movement of whale individuals:

$$X_i(t + 1) = X^*(t) + |X^*(t) - X_i(t)|e^{bl}\cos(2\pi l) \quad (18)$$

where "l" is a random number within $[-1, 1]$ and "b" is a constant that defines the shape of the exponential spiral model. A large value of "b" can be beneficial for the convergence of the Whale Optimization Algorithm (WOA), but it also increases the probability of becoming trapped in a local optimum within the solution space. for optimal setup of the calibration parameters in WOA, "a" should decrease from "2" to "0" in order to achieve an appropriate trade-off between convergence and the ability to search for the optimal solution. In the TASOM classifier, we optimize weight matrices in order to obtain more accuracy as the results are indicated in Table 6 and Fig.6 This table demonstrates improvements in the classification criteria when using the optimization algorithm WOA. Besides, we observe a slight increase in the program's execution time. Here, Maximum iteration number equals 100 and the constant number "a" is "0.5" and "b" is "40".

TABLE 6. Results from experiments, (1) Applying WOA in TASOM, (2) Without using WOA.

NO.	Experiment	Accuracy	Recall	MSE	General Time(s)
1	1	0.985	0.987	0.0248	48
	2	0.971	0.967	0.0299	42
2	1	0.990	0.986	0.0208	47
	2	0.972	0.976	0.0410	43
3	1	0.976	1.000	0.0309	46
	2	0.961	0.987	0.0418	41
4	1	0.972	1.000	0.0358	47
	2	0.966	1.000	0.0356	42
5	1	0.987	0.993	0.0248	48
	2	0.967	0.969	0.0312	43

FIGURE 6. The accuracy plots of using contourlet and WOA, only WOA, only contourlet in TASOM in ten experiments.

4) TUMOR CLASSIFIERS

There are various machine learning classifiers that are commonly used for classification tasks. The choice of classifier

TABLE 7. Average accuracy and run-time of all evaluated machine learning classifiers in 100 iterations.

No.	Classifiers	ACCURACY	F1-Score	Precision	Recall	*MSE*	*General Time(s)*
1	Gaussian Naïve Bayes	0.946	0.929	0.923	0.935	0.0477	31
2	MLP	0.976	0.964	0.957	0.971	0.0375	53
3	AdaBoost	0.951	0.931	0.893	0.971	0.0473	36
4	KNN	0.972	0.958	0.958	0.959	0.0319	38
5	SVM_RBF	0.971	0.961	0.948	0.973	0.0415	42
6	SVM_Linear	0.954	0.934	0.922	0.947	0.0435	39
7	SVM_Polygonal	0.963	0.944	0.931	0.957	0.0344	46
8	SVM_Quadratic	0.961	0.943	0.917	0.971	0.0401	44
9	Fuzzy C-Means	0.966	0.954	0.923	0.986	0.0361	31
10	**TASOM**	**0.985**	**0.979**	**0.961**	**1.00**	**0.0257**	**47**

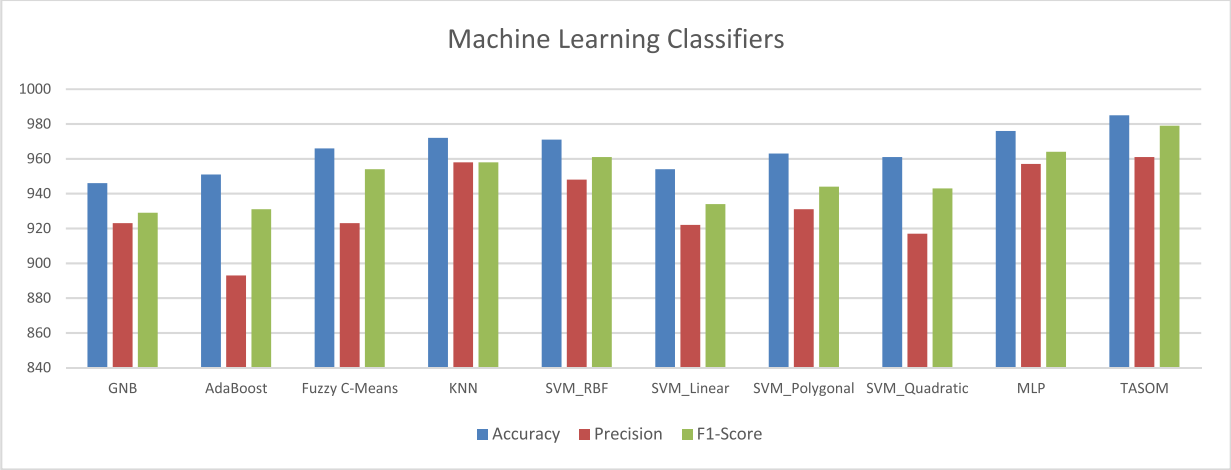


FIGURE 7. Comparison graph for the performance of classifiers, according to the results which are obtained from Table 7.

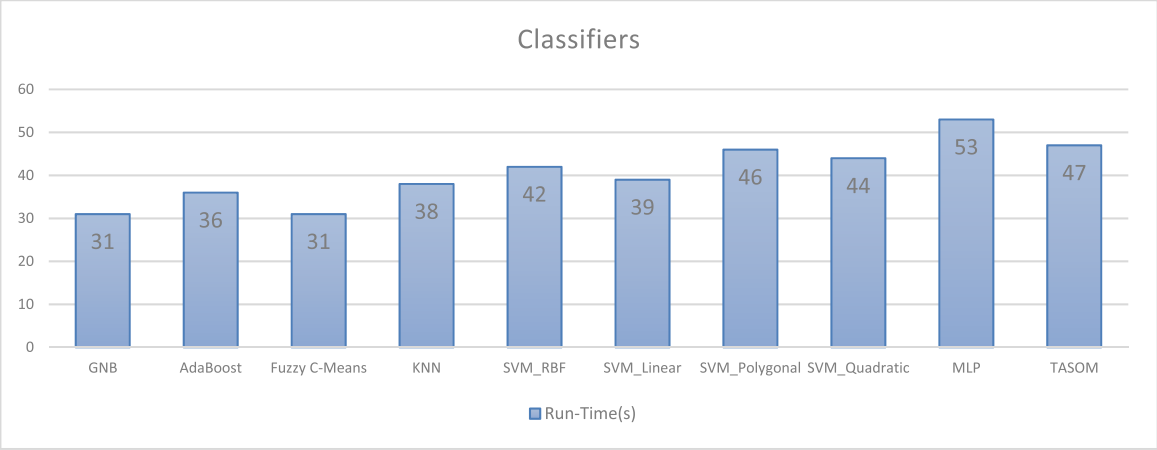
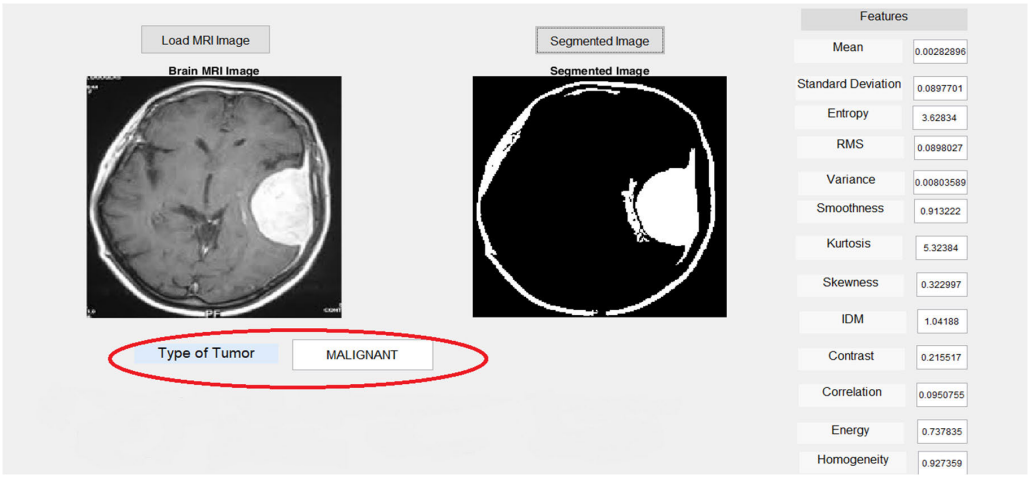
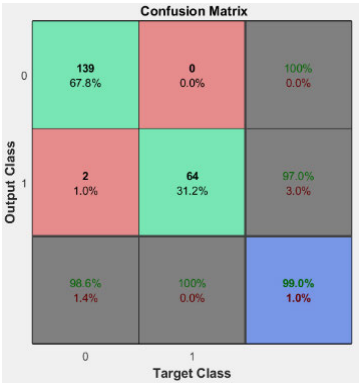


FIGURE 8. Diagram for time consumption in classifiers, according to the results which are obtained from Table 7.

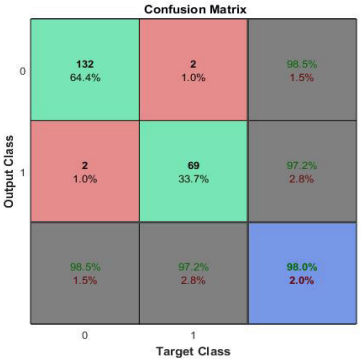
depends on the characteristics of the data, the problem domain, computational resources, interpretability requirements, and the specific performance goals. It is often beneficial to experiment with multiple classifiers and select the one that performs the best for a given task. Here are some popular ones:



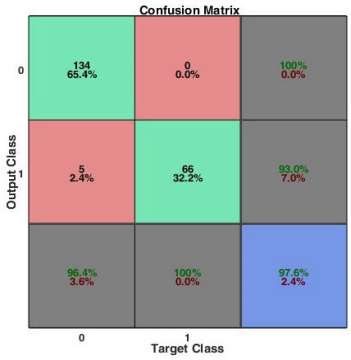
(a)



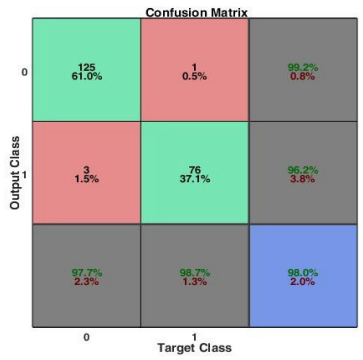
(b1)



(b2)

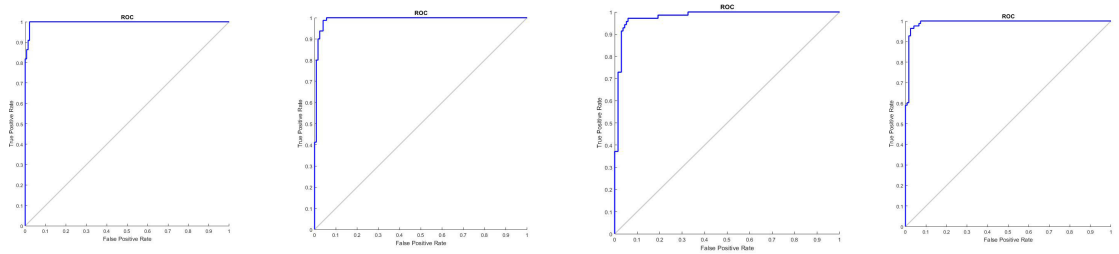


(b3)



(b4)

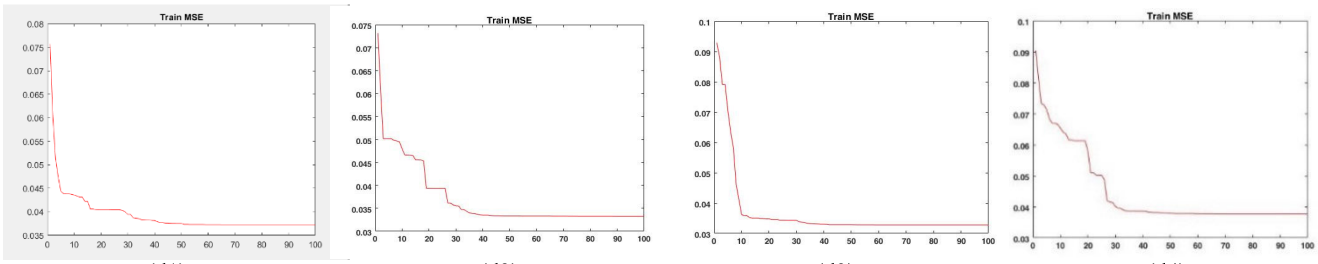
(b)



(c1) (c2) (c3) (c4)

(c)

FIGURE 9. Classification results, (a) Type of Tumor, (b) Confusion Matrix, (c) ROC Curve, (d) Train MSE.



(d1) (d2) (d3) (d4)

(d)

FIGURE 9. *(Continued.)* **Classification results, (a) Type of Tumor, (b) Confusion Matrix, (c) ROC Curve, (d) Train MSE.**

a: MULTI-LAYER PERCEPTRON

The multi-layer perceptron is a robust neural network in the field of function estimation and data classification [34], [35]. In this case, we use a three-layer network, where the layers have ''9'', ''19'', and ''1'' neuron(s), respectively. The activation function of the three layers is ''hyperbolic tangent sigmoid''. The network is trained for ''100'' epochs using the ''Levenburg and Marquardt'' training algorithm with a learning rate of 0.8.

b: GAUSSIAN NAIVE BAYES

Gaussian Naive Bayes is a probabilistic classification algorithm that leverages Bayes theorem and assumes feature independence. It estimates the mean and variance of each feature within each class using training data. When presented with a new input, it calculates the probabilities of the input belonging to each class based on the estimated mean and variance. By applying Bayes theorem, it obtains the posterior probability for each class and assigns the input to the class with the highest posterior probability [9], [36].

c: ADABOOST

AdaBoost is an ensemble learning algorithm that combines weak classifiers to create a strong classifier. It iteratively trains weak classifiers, placing more emphasis on misclassified samples in each iteration [37]. The final classifier is a weighted sum of the weak classifiers, with weights assigned, based on their accuracy. In our case, the algorithm employs ''100'' decision trees with a learning rate of ''0.1'', and uses decision tree stumps as weak learners within the ''AdaBoostM1'' framework.

d: K-NEAREST NEIGHBOR

In this classifier, the parameter ''K'' is set to three. Increasing the value of ''K'' leads to longer program run-times, but it may improve the image quality. We used the ''Euclidean'' method for calculating distances [38].

e: SUPPORT VECTOR MACHINES

The SVM (Support Vector Machine) algorithm seeks to find a hyperplane that maximizes the margin between different classes of data points. This hyperplane is chosen to optimize the distance between the closest points of the classes, known as support vectors. In binary classification, the goal is to find the hyperplane that effectively separates the two classes with the largest possible margin. In cases where a linear hyperplane is insufficient, SVM can employ a kernel function to transform the data into a higher-dimensional space, making separation achievable [12], [39]. In this research, we used SVM functions such as ''Quadratic'', ''Polynomial'', ''RBF'' and ''Linear''.

f: FUZZY C-MEANS

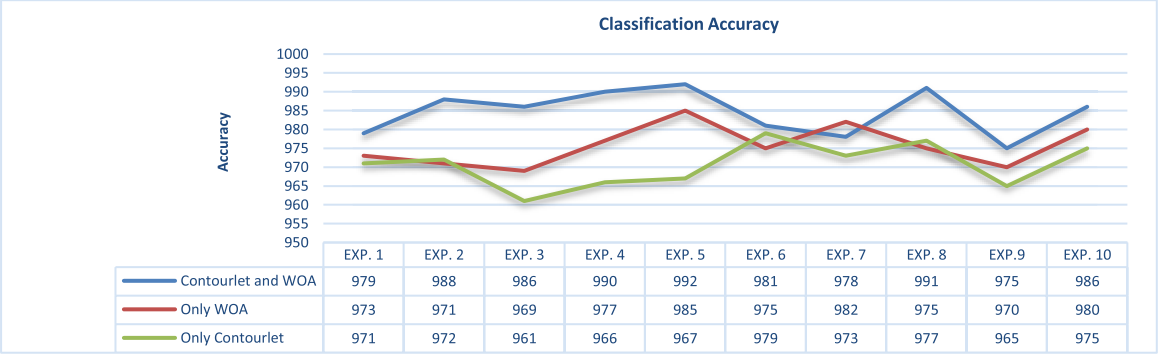
In this classification technique, each data point belongs to a class with a degree specified by a membership function. Similar to previous classifiers, we used both test and training data [40].

IV. ANALYSIS OF RESULTS

Accordance with Table 7 and Fig.7, Fig.8, it can be observed that among the 10 selected classifiers in the experimental conditions, the best results in terms of accuracy are respectively related to classifiers TASOM, MLP, KNN, SVM_RBF, FCM, SVM(Linear), AdaBoost and Gaussian Naïve Bayes. According to the criterion of program execution time for medical image classification, the fastest classifiers in order are Gaussian Naïve Bayes, FCM, AdaBoost, KNN, SVM(Linear), SVM(RBF), SVM (Quadratic), SVM(Polygonal), TASOM and MLP.

In order of results based on MSE, they are: TASOM, KNN, SVM(Polygonal),FCM,MLP,SVM(Quadratic),SVM(RBF), SVM(Linear), AdaBoost and Gaussian Naïve Bayes. Based on Fig.10 and Fig.11, the information obtained from diagrams indicate that the use of both the contourlet transform and the whale algorithm concurrently plays a significant role in improving the accuracy of results. Furthermore, the increase in execution time of the program is not considerably high and can be overlooked.

The data, obtained from the experiments in Fig.9, indicate that the proposed classifier in this research achieves remarkable results in the classification of benign and malignant brain tumors within an acceptable time. Moreover, considering the negligible computational errors in the proposed classifier and the high level of accuracy and precision in its performance,



	EXP. 1	EXP. 2	EXP. 3	EXP. 4	EXP. 5	EXP. 6	EXP. 7	EXP. 8	EXP.9	EXP. 10
Contourlet and WOA	979	988	986	990	992	981	978	991	975	986
Only WOA	973	971	969	977	985	975	982	975	970	980
Only Contourlet	971	972	961	966	967	979	973	977	965	975

FIGURE 10. Classification accuracy in three experimental conditions, evaluating the robustness of the proposed method.

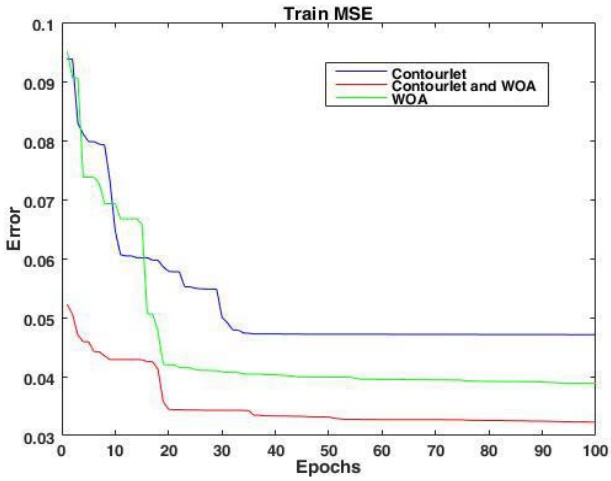


FIGURE 11. The results from the comparison of MSE in three experimental conditions.

we can easily rely on the brilliant results of this approach. At the same time, in Table.6 and Fig.6, the optimal effect of the Whale optimization heuristic algorithm on the obtained results can be clearly observed. Although the use of this tool does not have a significant impact on increasing the program execution time.

V. FUTURE WORKS

Magnetic Resonance Imaging (MRI) is a widely used diagnostic tool for brain tumors. The ability to accurately classify brain tumors using MRI is critical for proper diagnosis, treatment planning, and monitoring of treatment response [8], [41], [42]. While there have been significant advancements in the field, there is still a need for further research to improve the accuracy and efficiency of brain tumor classification using MRI [10], [43], [44].

VI. CONCLUSION

MRI brain tumor classification using a TASOM neural network has the potential to revolutionize the way physicians diagnose and treat brain tumors. The success of the experiments conducted on this system demonstrates its ability to improve the accuracy and speed of diagnosis, which is crucial for early detection and better patient outcomes. By power of this method, the system can detect patterns and features in MRI images that traditional machine learning algorithms cannot. Furthermore, the system's performance surpasses traditional methods of machine learning in terms of accuracy and runtime.

The use of this novel neural network for brain tumor classification is especially relevant given the increasing prevalence of brain tumors worldwide. With the growing demand for more efficient and accurate diagnostic tools, this system is a promising step forward in addressing the needs of patients and healthcare providers alike. It has the potential to greatly reduce the time and resources required for diagnosis and treatment, as well as enhance the accuracy of the diagnostic process. Overall, the use of this novel neural network for MRI brain tumor classification has shown great potential and could have a significant impact on the field of medical imaging in the future. Our system has achieved a classification accuracy of over 98.5% through experiments with various test samples, while maintaining an acceptable run-time.

REFERENCES

[1] M. Gurbină, M. Lascu, and D. Lascu, "Tumor detection and classification of MRI brain image using different wavelet transforms and support vector machines," in *Proc. 42nd Int. Conf. Telecommun. Signal Process. (TSP)*, Jul. 2019, pp. 505–508, doi: 10.1109/TSP.2019.8769040.

[2] J. Kang, Z. Ullah, and J. Gwak, "MRI-based brain tumor classification using ensemble of deep features and machine learning classifiers," *Sensors*, vol. 21, no. 6, p. 2222, Mar. 2021.

[3] S. H. Hazaveh, A. Bayandour, A. Khalili, A. Barkhordary, A. Farzamnia, and E. G. Moung, "Impulsive noise suppression methods based on time adaptive self-organizing map," *Energies*, vol. 16, no. 4, p. 2034, Feb. 2023, doi: 10.3390/en16042034.

[4] R. C. Gonzales and R. E. Woods, *Digital Image Processing*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2001.

[5] S. V. Vaseghi, *Advanced Digital Signal Processing and Noise Reduction*. Hoboken, NJ, USA: Wiley, 2008.

[6] B. Jahne, *Digital Image Processing: Concepts, Algorithms, and Scientific Applications*. Cham, Switzerland: Springer-Verlag, 2000.

[7] P. G. Rajan and C. Sundar, "Brain tumor detection and segmentation by intensity adjustment," *J. Med. Syst.*, vol. 43, no. 8, pp. 1–13, Aug. 2019.

[8] N. V. Shree and T. N. R. Kumar, "Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network," *Brain Informat.*, vol. 5, no. 1, pp. 23–30, Mar. 2018.

[9] Z. Ullah, M. U. Farooq, S.-H. Lee, and D. An, "A hybrid image enhancement based brain MRI images classification technique," *Med. Hypotheses*, vol. 143, Oct. 2020, Art. no. 109922.
[10] B. Ural, "A computer-based brain tumor detection approach with advanced image processing and probabilistic neural network methods," *J. Med. Biol. Eng.*, vol. 38, no. 6, pp. 867–879, Dec. 2018.
[11] F. J. Díaz-Pernas, M. Martínez-Zarzuela, M. Antón-Rodríguez, and D. González-Ortega, "A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network," *Healthcare*, vol. 9, no. 2, p. 153, Feb. 2021.
[12] T. Ateeq, M. N. Majeed, S. M. Anwar, M. Maqsood, Z.-U. Rehman, J. W. Lee, K. Muhammad, S. Wang, S. W. Baik, and I. Mehmood, "Ensemble-classifiers-assisted detection of cerebral microbleeds in brain MRI," *Comput. Electr. Eng.*, vol. 69, pp. 768–781, Jul. 2018.
[13] D. J. Hemanth, J. Anitha, A. Naaji, O. Geman, D. E. Popescu, and L. H. Son, "A modified deep convolutional neural network for abnormal brain image classification," *IEEE Access*, vol. 7, pp. 4275–4283, 2019.
[14] P. Saxena, A. Maheshwari, and S. Maheshwari, "Predictive modeling of brain tumor: A deep learning approach," 2019, *arXiv:1911.02265*.
[15] N. M. Dipu, S. A. Shohan, and K. M. A. Salam, "Deep learning based brain tumor detection and classification," in *Proc. Int. Conf. Intell. Technol. (CONIT)*, Jun. 2021, pp. 1–6, doi: 10.1109/CONIT51480.2021.9498384.
[16] A. Sarkar, M. Maniruzzaman, M. A. Alahe, and M. Ahmad, "An effective and novel approach for brain tumor classification using AlexNet CNN feature extractor and multiple eminent machine learning classifiers in MRIs," *J. Sensors*, vol. 2023, pp. 1–19, Mar. 2023, doi: 10.1155/2023/1224619.
[17] G. Gao, B. Cai, S. Xu, and T. Yan, "Watermark performance contrast between contourlet and non-subsampled contourlet transform," in *Proc. IEEE Int. Conf. Inf. Autom.*, Jun. 2012, pp. 507–511.
[18] C. Zhang, X. Fu, S. Peng, and Y. Wang, "Linear unequally spaced array synthesis for sidelobe suppression with different aperture constraints using whale optimization algorithm," in *Proc. 13th IEEE Conf. Ind. Electron. Appl. (ICIEA)*, Jun. 2018, pp. 69–73, doi: 10.1109/ICIEA.2018.8397691.
[19] P. Hu, W. Wang, C. Zhang, and K. Lu, "Detecting salient objects via color and texture compactness hypotheses," *IEEE Trans. Image Process.*, vol. 25, no. 10, pp. 4653–4664, Oct. 2016.
[20] S. S. Naqvi, W. N. Browne, and C. Hollitt, "Feature quality-based dynamic feature selection for improving salient object detection," *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 4298–4313, Sep. 2016.
[21] F. Meng, X. Li, and J. Pei, "A feature point matching based on spatial order constraints bilateral-neighbor vote," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4160–4171, Nov. 2015.
[22] M. A. Elaziz, M. Ahmadein, S. Ataya, N. Alsaleh, A. Forestiero, and A. H. Elsheikh, "A quantum-based chameleon swarm for feature selection," *Mathematics*, vol. 10, no. 19, p. 3606, Oct. 2022.
[23] A. H. Elsheikh, M. A. Elaziz, and A. Vendan, "Modeling ultrasonic welding of polymers using an optimized artificial intelligence model using a gradient-based optimizer," *Weld. World*, vol. 66, no. 1, pp. 27–44, Jan. 2022, doi: 10.1007/s40194-021-01197-x.
[24] M. A. Elaziz, A. H. Elsheikh, D. Oliva, L. Abualigah, S. Lu, and A. A. Ewees, "Advanced metaheuristic techniques for mechanical design problems: Review," *Arch. Comput. Methods Eng.*, vol. 29, no. 1, pp. 695–716, Jan. 2022, doi: 10.1007/s11831-021-09589-4.
[25] A. H. Elsheikh, H. Panchal, M. Ahmadein, A. O. Mosleh, K. K. Sadasivuni, and N. A. Alsaleh, "Productivity forecasting of solar distiller integrated with evacuated tubes and external condenser using artificial intelligence model and moth-flame optimizer," *Case Stud. Thermal Eng.*, vol. 28, Dec. 2021, Art. no. 101671.
[26] M. A. Elaziz, A. Dahou, N. A. Alsaleh, A. H. Elsheikh, A. I. Saba, and M. Ahmadein, "Boosting COVID-19 image classification using MobileNetV3 and Aquila optimizer algorithm," *Entropy*, vol. 23, no. 11, p. 1383, Oct. 2021, doi: 10.3390/e23111383.
[27] E. B. Moustafa, A. H. Hammad, and A. H. Elsheikh, "A new optimized artificial neural network model to predict thermal efficiency and water yield of tubular solar still," *Case Stud. Thermal Eng.*, vol. 30, Feb. 2022, Art. no. 101750.
[28] A. H. Elsheikh, T. A. Shehabeldeen, J. Zhou, E. Showaib, and M. A. Elaziz, "Prediction of laser cutting parameters for polymethylmethacrylate sheets using random vector functional link network integrated with equilibrium optimizer," *J. Intell. Manuf.*, vol. 32, no. 5, pp. 1377–1388, Jun. 2021.
[29] W. K. Pratt, "Image restoration models," in *Digital Image Processing*. Hoboken, NJ, USA: Wiley,2001, pp. 297–317, doi: 10.1002/0471221325.ch11.
[30] H. Shah-Hosseini and R. Safabakhsh, "TASOM: A new time adaptive self-organizing map," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 33, no. 2, pp. 271–282, Apr. 2003.
[31] H. Shah-Hosseini and R. Safabakhsh, "Automatic multilevel thresholding for image segmentation by the growing time adaptive self-organizing map," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 10, pp. 1388–1393, Oct. 2002.
[32] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2091–2106, Dec. 2005.
[33] M. Nixon and A. S. Aguado, *Feature Extraction & Image Processing*, 2nd ed. Cambridge, MA, USA: Academic, 2008.
[34] C. Szepesvari, *Image Processing: Low-Level Feature Extraction*. Edmonton, AB, Canada: Univ. Alberta, 2007.
[35] G. S. Tandel, M. Biswas, O. G. Kakde, A. Tiwari, H. S. Suri, M. Turk, J. R. Laird, C. K. Asare, A. A. Ankrah, and N. N. Khanna, "A review on a deep learning perspective in brain cancer classification," *Cancers*, vol. 11, no. 1, p. 111, 2019.
[36] K. Popuri, D. Cobzas, A. Murtha, and M. Jägersand, "3D variational brain tumor segmentation using Dirichlet priors on a clustered feature set," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 7, no. 4, pp. 493–506, Jul. 2012.
[37] H. Selvaraj, S. T. Selvi, D. Selvathi, and L. Gewali, "Brain MRI slices classification using least squares support vector machine," *Int. J. Intell. Comput. Med. Sci. Image Process.*, vol. 1, no. 1, pp. 21–33, Jan. 2007.
[38] P. John, "Brain tumor classification using wavelet and texture based neural network," *Int. J. Sci. Eng. Res.*, vol. 3, pp. 1–7, Oct. 2012.
[39] J. Cheng, W. Yang, M. Huang, W. Huang, J. Jiang, Y. Zhou, R. Yang, J. Zhao, Y. Feng, Q. Feng, and W. Chen, "Retrieval of brain tumors by adaptive spatial pooling and Fisher vector representation," *PLoS ONE*, vol. 11, no. 6, Jun. 2016, Art. no. e0157112.
[40] A. Kharrat, K. Gasmi, M. Messaoud, N. B. Ben, and M. Abid, "A hybrid approach for automatic classification of brain MRI using genetic algorithm and support vector machine," *Leonardo J. Sci.*, vol. 17, pp. 71–82, Jul. 2010.
[41] E. I. Papageorgiou, P. P. Spyridonos, D. T. Glotsos, C. D. Stylios, P. Ravazoula, G. N. Nikiforidis, and P. P. Groumpos, "Brain tumor characterization using the soft computing technique of fuzzy cognitive maps," *Appl. Soft Comput.*, vol. 8, no. 1, pp. 820–828, Jan. 2008.
[42] M. Arunachalam and S. R. Savarimuthu, "An efficient and automatic glioblastoma brain tumor detection using shift-invariant shearlet transform and neural networks," *Int. J. Imag. Syst. Technol.*, vol. 27, no. 3, pp. 216–226, Sep. 2017.
[43] M. Goyal, R. Goyal, and B. Lall, "Learning activation functions: A new paradigm of understanding neural networks," 2019, *arXiv:1906.09529*.
[44] S. Christodoulidis, M. Anthimopoulos, L. Ebner, A. Christe, and S. Mougiakakou, "Multisource transfer learning with convolutional neural networks for lung pattern analysis," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 1, pp. 76–84, Jan. 2017.