Brain tumor identification with a hybrid feature extraction method based on discrete wavelet transform and principle component analysis

Dalia Mohammad Toufiq[1], Ali Makki Sagheer[2], Hadi Veisi[3]
[1]Department of Family and Community Medicine, University of Sulaimani, Iraq
[2]Department of Computer Technical Engineering, Al-Qalam University College, Iraq
[2]College of Computer Science and Information Technology, University of Anbar, Iraq
[3]Department of Computer Engineering, University of Tehran, Iran

Article Info

ABSTRACT

The Identification of brain tumors is a critical step that relies on the expertise and abilities of the physician. In order to enable radiologists to spot brain tumors, an automated tumor arrangement is extremely important. This paper presents a technique for MR brain image segmentation and classification to identify images as normal and abnormal. The proposed technique is a hybrid feature extraction submitted to enhance the classification results and basically consists of three stages. The first stage is used a 3-level of discrete wavelet transform (DWT) to extract image characteristics. In the second stage, the principle component analysis (PCA) is applied to reduce the size of characteristics. Finally, a random forest classifier (RF) was used with a feature selection for identification. 181 MR brain images are collected (81 normal and 100 abnormal), in distinguishing normal and abnormal tissues, the experimental results obtained an accuracy of 98%, the sensitivity achieved is 99.2%, specificity achieved is 97.8%, and showed the effectiveness of the proposed technique compared with many kinds of literature. The results show that the 3L-DWT+PCA+RF still achieved the best classification results. The proposed model could apply to the brain MRI sphere classification, which will help doctors to diagnose a tumor if it is normal or abnormal in certain degrees.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Dalia Mohammad Toufiq
Department of Family and Community Medicine
University of Sulaimani
Kirkuk Road, Sulaimani, Kurdistan Region, Iraq
Email: dalia.toufiq@univsul.edu.iq

1. INTRODUCTION

Magnetic resonance imaging (MRI) is a technique that generates great quality images of the physical body's anatomical structures, especially within the brain, and it provides useful knowledge on biomedical research and clinical diagnosis [1]-[3]. MRI is described as a more appropriate and useful imaging technique for brain tumors other than methodologies. Contain knowledge in detail on tumor type, position, and size in a non-invasive manner provided by the MRI [4]. In MRI scanners, T2-w images are widely utilized to provide an initial evaluation, classify types of tumors, and differentiate tumors from non-tumor tissues [5]-[7].

As scanner resolutions were enhanced, and the thickness of slices decreased, a large number of slices were constructed, and clinicians needed more time for each patient to diagnose from their image.

Consequently, in the past 20 years, automated detection and segmentation for tumors have attracted great attention [8].

The proposed method predicated on the utilizing of T2-w images in axial viewing to identify brain anomalies. Wavelet transform is an effective method for extracting features from MR brain images because, its multi-resolution analytic property, it enables image analysis at different resolution levels [9]. The principal component analysis (PCA) was used to scale back the feature vector dimensions and to increase the discriminative power [10]. PCA is attractive because it effactually decreases the dimensions of the data, thus reducing the cost of computing of new data analysis [11].

In the previous works, features were extracted from the segmented image. Since discrete wavelet transform (DWT) can efficiently extract the information from original MR images with little loss and PCA reduce the dimensions of features to a higher degree. In this paper, in order to obtain features that ensure optimal classification results, in addition to the features that extracted from the segmented image, additional textural features extracted from the PCA component of LL sub bands 3-level wavelet decomposition. Then, a random forest classifier is suggested in this work to identify the brain image as normal and abnormal.

In literature, Bahadure et al. [12] proposed an image analysis of Berkeley wavelet transform (BWT) and support vector machine (SVM) methods for MRI-based identification and classification of brain tumors. The accuracy of 95 percent is already accomplished in this process utilizing skull stripping, which for the purpose of detection removed all non-brain tissues. MR brain image segmentation utilizing a K-means clustering algorithm with morphological filtering for tumor image detection was suggested by Joseph et al. [13]. Alfonse and Salem [14] suggested an automated system for the classification of MR images of brain tumors using the support vector machine. Utilizing fast Fourier transform for the features extraction, the accuracy of a classifier was improved and the maximum relevance technique of minimal redundancy was used for the reduction of features. 98 percent was the precision obtained from this proposed work.

Also, Yao et al. [15] proposed a technique that included 83 percent accuracy in the extraction of texture characteristics with wavelet transform and SVM. Kumar et al. [16] suggested a methodology utilizing PCA and SVM, using this technique achieved an accuracy of 94 percent. Mohsen et al. [17] classified 66 images of brain tumors into four categories: tumor-free, glioblastoma, sarcoma, and metastasis. They reached a 96.97% accuracy using a deep neural network (DNN).

In addition, Chaddad [18] suggested automated feature extraction and enhanced tumor detection using the Gaussian mixture model applied to wavelet MRI and main component analysis with an accuracy of 95 percent for both T1-weighted and T2-weighted and 92 percent for FLAIR MR. Sachdeva et al. [19] utilized artificial neural network (ANN) and PCA-ANN for the classification of multiclass MR brain tumor images, 428 MR image segmentation, and 75-90 percent accuracy.

The above-mentioned survey gives a detailed vision of the techniques invented specifically to acquire a region of interest, and characteristics of extraction techniques. When the extracted features are few resulted in low tumor identification and accuracy of detection. This research is arranged in the following sections. Section 2 provides the comprehensive procedures of the proposed model, including k-means clustering, segmentation, discrete wavelet transform, principal component analysis, and presents the concepts of random forest classifier. The experiments in section 3 use a full dataset of 181 images, showing the effects of extracting and reducing features results compared with related various techniques. Conclusions and discussions are devoted in section 4.

2.   PROPOSED METHOD

In order to extract image characteristics, the suggested approach basically used preprocessing to enhance and dedicate the region of interest (ROI) image. Then the 3 levels of DWT are applied. After that, the PCA is used to decrease the size of characteristics. Finally, a random forest classifier (RF) with a selection of identification features was used. The approach consists of five stages:

a. Preprocessing including:
   - Resizing MR images
   - Apply k-means clustering
   - Segmentation
b. Transformation and reduction (including applying 3L DWT and PCA).
c. Feature extraction.
d. Random forest training, apply new MRI brains to the trained random forest and perform the prediction.
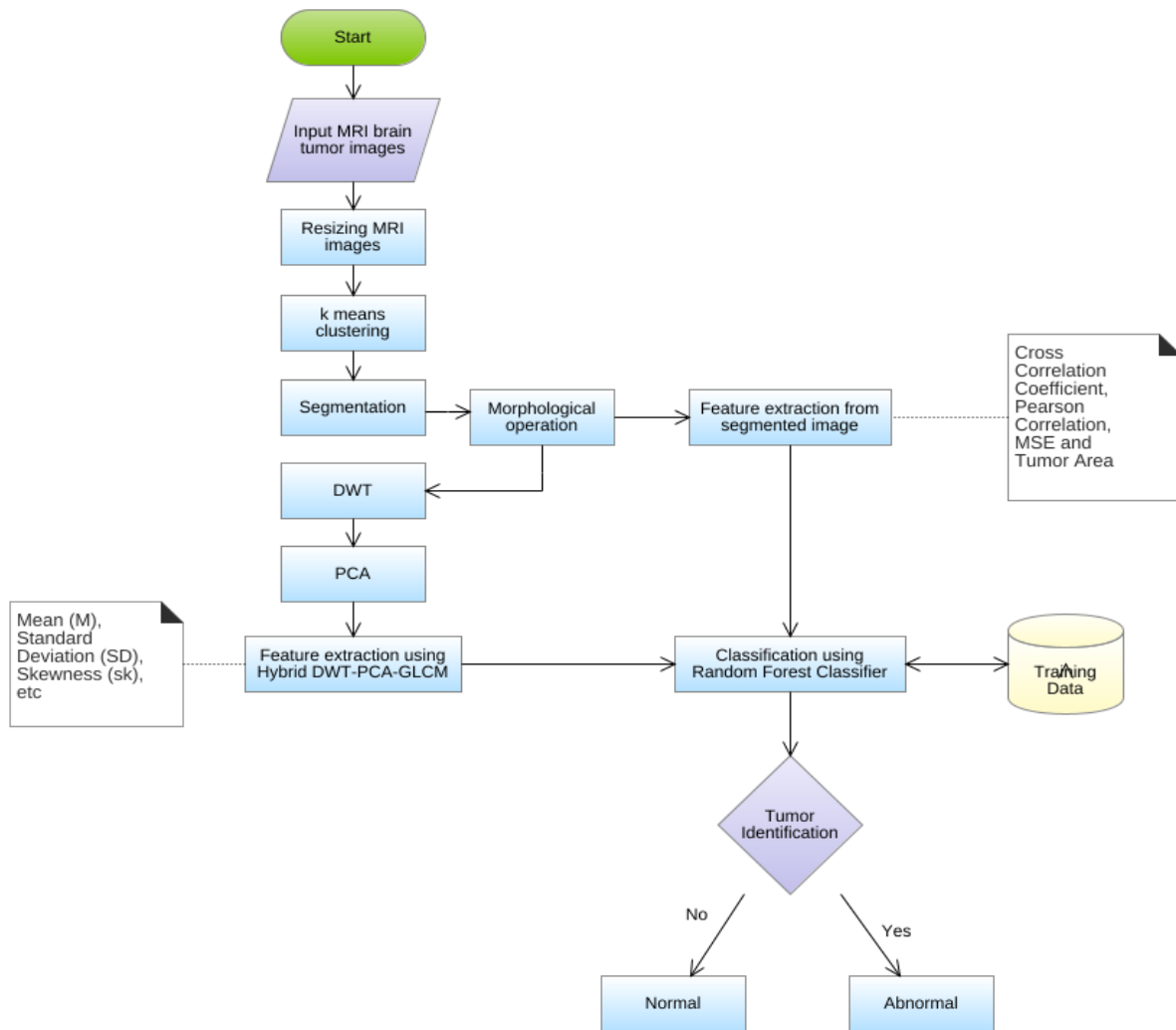   In Figure 1, the detailed processes of the proposed model are clarified.

Figure 1. The methodology of the proposed algorithm

2.1. Preprocessing
2.1.1. Resizing the dimensions of MR images
The provided MRI brain slices were collected from different scanners with different spatial resolutions. To enable the utilization of the complete set disinterestedly, the dimensions of the magnetic resonance imaging were changed using a nearest neighbor interpolation approach so that the width or height doesn't exceed 256 pixels while preserving the ratio of the image when changing its size.

2.1.2. K means clustering
Clustering is a process of grouping or partitioning a given pattern into several clusters such that similar patterns are assigned to a group which is called a cluster. Many forms of analysis use clustering to blot out the field of image segmentation. Different techniques exist and the k-means clustering algorithm is one of the most common methods. The clustering algorithm K-means is an unsupervised algorithm and the interest area from the background is a customary segment [20].

2.2. Distinct region of interest (ROI)
Segmentation is a mechanism in which the MRI is broken into distinct regions. Let the entire area of the image be stated by A. The method of segmentation can be seen as a partition of A into n sub-regions such as A1, A2, A3... An. As the segmentation must be intact, some requirements must be fulfilled; that is, every pixel should be within the region, every point should be linked in some way within the regions, regions should be disjointed [21], [22].

Dilation and erosion are the essential operations used here. Dilation attaches the pixels to the boundary region, while erosion deletes the pixels from the object boundary region. Based on the structuring components, these operations were conducted by comparing all the values of pixels in the neighborhood of the input image defined by the structuring element, dilation selects the highest value, during erosion, the rock bottom value is chosen by comparing all the pixel values in the input image region [23].

A ROI is a portion of an image on which you want to filter or do some other activity. By making a binary mask, you define a ROI, which is a binary image of the same size as the image you want to process, with pixels representing the ROI set to 1 and all other pixels set to 0. The segmentation of the affected brain MRI regions accomplished by two steps:

- Transformed the preprocessed brain MR image into a binary image with a cut-off threshold of 128 chosen in the initiative. Pixel values greater than 128 mapped as white, the others are marked as black.
- In the second phase, an erosion method of morphology used to remove sporadic white pixels. Dilation anchors the segmented region as a result, the tumor area remains without any abnormalities.

2.3. Transformation

The DWT was added separately to every dimension in the case of two-dimension images. As a result, each scale has four sub-bands (LL, LH, HH, and HL). For the next two-dimension DWT, the sub-band LL is hired. The LL sub-band is often considered the image's approximation component, while the detailed components of the image can be considered the LH, HL, and HH sub-bands. Therefore, to interpret the image detail, wavelets provide an easy hierarchical structure. Three-level decomposition of Harr wavelet was used in our proposed model.

2.4. Principal component analysis

PCA is an important method to scale down the dimension of a data set composed of an over-sized number of interrelated variables while preserving much of the variants. It is done by converting the data set into a completely new set of ordered variables aligned with their variances or significance [24]. This approach has two effects: It orthogonalizes the input vector components so that they do not correlate with each other, and uncorrelated with each other in preparation for those with the most substantial variance come first and remove those components that add the smallest amount to the variance in the data collection.

2.5. Features extraction

The analysis of texture effectively distinguishes natural and irregular tissues for human beholding and machine learning. Offers difference between normal and malignant tissues that cannot be observed by the human eye. It increases efficacy for early diagnosis, by picking effective quantitative features. In the initiative, statistical textural analysis-features (cross-correlation coefficient, pearson correlation, and tumor area) information from the segmented image intensities extracted. In the next step, textural features were obtained from the PCA components acquired from the LL sub-bands of the first three-level wavelet decomposition.

2.5.1. Feature extraction from the segmented image

In this method, four features (cross-correlation coefficient, pearson correlation, mean square error (MSE), and tumor area) were obtained from the segmented image. The textural features extracted listed:

a. Cross-correlation coefficient

It is a measure of similarity of two series as a function of the displacement of one relative to the other. The cross-correlation coefficients are more robust to changes of illumination than the MSE [25].

$$\text{Cross} - \text{Correlation Coefficient} = \frac{\sum_k (x_k - \bar{x})(y_k - \bar{y})}{\sqrt{\sum_k (x_k - \bar{x})^2}} \tag{1}$$

b. Pearson correlation coefficient

Pearson correlation evaluates if there is statistical support for a linear relationship, represented by a population correlation coefficient, between the same pairs of variables in the population. A parametric calculation is the Pearson Correlation [26].

$$\text{Pearson Correlation Coefficient} = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}} \tag{2}$$

c. Mean square error

It used by providing quantitative or similarity scores to compare two images and defined as [27]:

$$MSE = \frac{1}{m*n} \sum \sum (f(x,y) - \bar{f}(x,y))^2 \tag{3}$$

2.5.2. Feature extraction using hybrid DWT-PCA-GLCM method

One of the most commonly used image-processing implementations of the gray level co-occurrence matrix (GLCM) and texture function has been developed by Haralick *et al.* [28]. The hybrid DWT-PCA-GLCM is a DWT-PCA-based GLCM feature extraction method that combines the DWT and PCA method with GLCM. In this method, the three-wavelet decomposition levels significantly reduce the size of the input image, as shown in Figure 2.

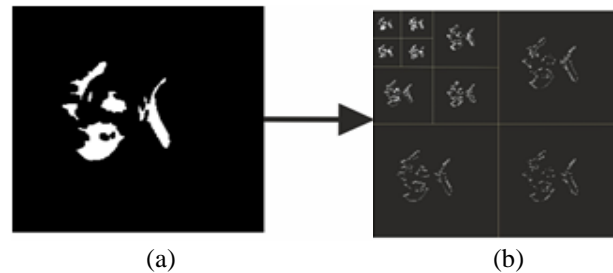(a)   (b)

Figure 2. The procedures of 3-level 2D DWT; (a) abnormal brain MRI, (b) level-3 wavelet coefficients

The wavelet coefficients image's top left corner denotes the level-3 approximation coefficients, the value of which is just 32×32=1024. The quantity of extracted features was reduced to 1024, as mentioned above. Nonetheless, it is also too big for estimation. PCA is utilized to further minimize the size of features to an optimum degree. Then the features are extracted using the GLCM algorithm from the PCA components. The statistics formulas for the features are listed:

a. Mean (M)
The image mean is determined by summing all the image pixel values divided by the total count of image pixels [27].

$$M = \frac{1}{m*n} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x,y) \tag{4}$$

b. Standard deviation (SD)
The second central moment is the standard deviation that defines the distribution of the probability of an observed population and can function as an inhomogeneity metric. A higher value implies a higher level of intensity and high contrast between an image's edges [27].

$$SD = \sqrt{\frac{1}{m*n} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (f(x,y) - M)^2} \tag{5}$$

c. Kurtosis (Kurt)
The shape of the probability distribution of a random variable is defined as Kurtosis. It denoted as Kurt(X) for the random variable X and it defined as [27]:

$$K_{urt}(X) = (\frac{1}{m*n}) \frac{\sum (f(x,y) - M)^4|}{SD^4} \tag{6}$$

d. Energy (En)
The quantifiable quantity of the degree of pixel pair repetitions is described as energy. It is defined as [28].

$$En = \sqrt{\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x,y)^2} \tag{7}$$

e. Coarseness (Cness)
Coarseness is the textural analysis of an image as an indicator of roughness [27].

$$\text{Coarseness} = \frac{1}{2^{m+n}} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x,y) \tag{8}$$

f. Homogeneity: it is defined as [29]:

$$\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} \frac{1}{1+(i-y)^2} \cdot f(x,y) \tag{9}$$

g. Variance: it is defined as [30]:

$$\frac{\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (f(x,y)-M)^2}{m*n} \tag{10}$$

h. Auto correlation predictor: it is defined as [29]:

$$\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (x*y) \cdot f(x,y) \tag{11}$$

i. Dissimilarity: it is defined as [29]:

$$\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} |x-y| \cdot f(x,y) \tag{12}$$

2.6. Random forest classifier

Classification is the method of arranging objects into different groups in an image that constitutes the final stage in the processing of images. Random forest classifiers are introduced and the outcomes are contrasted as well. Training samples were randomly selected and five cross-validations were employed to validate the robustness of the proposed system.

Random forest is a supervised learning algorithm. It's just used for classification issues, though. As we all know, a forest is made up of trees and a more robust forest means more trees. Likewise, the random forest method generates decision trees from data samples and then retrieves the estimate from every one of them, and finally selects the most efficient voting solution. It is an ensemble approach that is stronger than a single decision tree and by integrating the effect, it eliminates the over-fitting [31]. The working of the random forest algorithm is summarized in the following steps [32]:

- Start with collecting random samples from a given dataset first.
- Next, for each sample, this algorithm can create a decision tree. Then, from any decision tree, it will get the prediction output.
- Performed voting for each predicted result.
- Select the foremost voted prediction as to the final prediction

K folds are mostly partitioned purely at random, but certain folds may have a somewhat different distribution than other folds. Stratified K-fold cross-validation has also been used, where each fold has almost an equal class distribution [33], [34]. We would assume the 5-fold cross-validation in this study.

3. RESULTS AND DISCUSSION
3.1. Data set

The datasets consist of axial plane T2-weighted MR brain images of resolution 256×256 in-plane, retrieved from the Harvard Medical School website (URL: http:/med.harvard.edu/AANLIB/) and the OASIS dataset (URL: https:/www.oasis-brains.org/). Since T2 images, compared to T1 and PET modalities, are of greater contrast and better vision, we selected the T2 model. Consisting of 81 normal and 100 abnormal brain images, 181 images were chosen.

3.2. K-fold stratified cross-validation

To eliminate this overfitting in the proposed system the cross-validation is applied. The overall classification precision will not improve by cross-validation, but it will make the classifier accurate and can be extended to other separate datasets. Three types are used in cross-validation methods: K-fold cross-validation, random subsampling, and leave-one-out validation. Due to its properties, the K-fold cross-validation is implemented and uses all data for training and validation. The method is used to make the whole dataset a K-fold partition; Repeat K times for training using K-1 folds and a left fold for validation, and eventually averages the error rates of the K experiment as shown in Figure 3. Table 1, demonstrates the setting of the training images and the validation images, as 5-fold cross-validation was used.

Figure 3. 5-fold cross-validation

Table 1. The configuration (cross-validation) of training and validation images

Total No.of images	Training (145)		Validation (36)	
	Normal	Abnormal	Normal	Abnormal
181	65	80	16	20

3.3. Classification accuracy

The results of the proposed system are obtained by using actual MR brain images. The proposed system is implemented by using C# development with Visual Studio. Net framework, which runs on the Windows 10 OS, Intel Core i7 processor, and 8 GB RAM, the proposed algorithm is carried out. Since MRI scan visual diagnosis is subjective and based on the radiologist's experience, texture analysis has been thoroughly researched to enhance the diagnosis of brain MRI scans. First, during this study, k means clustering algorithm and thresholding accompanied by morphological operations have been combined, and then apply DWT and PCA with the extraction of GLCM features, assessed as a classifier tool with random forest. The study deal with the extraction of segmented area features to detect and distinguish medical brain MR images of normal and abnormal tumor cells. Performance of the classification is measured in terms of accuracy, sensitivity, specificity, and ROC curve as shown in Figure 4. The classification measures can be determined as:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{13}$$

$$\text{Specificity} = \frac{TN}{TN+FP} \tag{14}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{15}$$

As seen in Table 2, the experimental results of the proposed algorithm are contrasted with prior research. The proposed system result leads to the conclusion that it makes as possible for clinical experts to decide and diagnose.

Table 2. Comparison with previously proposed methods

Reference	Features methods	Classifier	Accuracy
Nabizadeh *et al.* [35]	First-order statistical, GLCM, GLRL, HOG, LBP	SVM	97.4%
Dvořák *et al.* [36]	Searching about the pathological area by symmetry checking	SVM	91.15%
Hasan *et al.* [29]	MGLCM	MLP	97.8%
Proposed System	Hybrid DWT-PCA-GLCM	RF	98% Accuracy 99.2% Sensitivity 97.8% Specificity

The proposed system is designed to classify and identify the brain MR image into normal and abnormal tumors. The accuracy of the system is achieved 98% for the tested dataset. Due to the statistical textural features were extracted from the PCA component of LL sub bands 3-level wavelet decomposition.

By the results that have been achieved, concluded that the proposed method outperformed and Clearly distinguishes between normal and abnormal tumors, enabling clinical experts in making accurate diagnosis decisions.
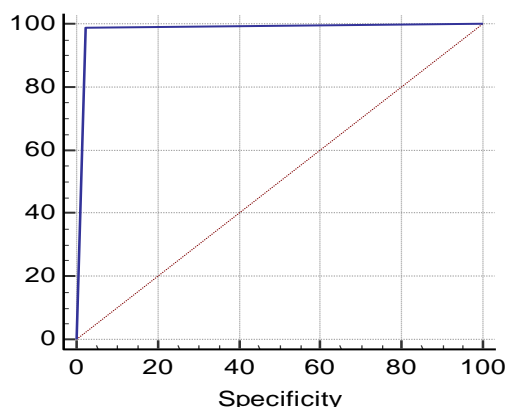
Figure 4. ROC curve of the classification results

4. CONCLUSION

In this paper, a hybrid DWT+PCA+GLCM identifier system has been developed to classify the normal and abnormal MRIs of the brain based on hybrid techniques such as discrete wavelet transforms and PCA with random forest. The foremost important contribution of this paper is a proposal of a technique that combines them with GLCM as a robust tool for identifying normal MR brain from abnormal MR brain. This algorithm helps clinicians to enhance the accuracy of the diagnosis. Because most brain tumors look hyper-intense in these images relative to normal brain tissue, it has been discovered that the statistical texture characteristics derived by GLCM are adequate to distinguish pathological patients from non-pathological patients when utilizing T2 weighted MR images. The experiments demonstrate that the proposed feature extraction tools with the RF classifier obtained 98% classification accuracy on the 181 MR images, over other popular methods in recent literature.

REFERENCES

[1] Y. Zhang, L. Wu, and S. Wang, "Magnetic Resonance Brain Image Classification By An Improved Artificial Bee Colony Algorithm," *Progress In Electromagnetics Research*, vol. 116, pp. 65-79, 2011, doi: 10.2528/PIER11031709.
[2] L. Golestanirad, A. P. Izquierdo, S. J. Graham, J. R. Mosig, and C. Pollo, "Effect Of Realistic Modeling Of Deep Brain Stimulation On The Prediction Of Volume Of Activated Tissue," *Progress In Electromagnetics Research*, vol. 126, pp. 1-16, 2012.
[3] S. A. Mohsin, "Concentration Of The Specific Absorption Rate Around Deep Brain Stimulation Electrodes During Mri," *Progress In Electromagnetics Research*, vol. 121, pp. 469-484, 2011, doi: 10.2528/PIER11022402.
[4] G. Dougherty, "Digital image processing for medical applications," *Cambridge University Press*, 2009.
[5] L. Mechtler, "Neuroimaging in Neuro-Oncology," *Neurologic Clinics*, vol. 27, no. 1, pp. 171-201, 2009, doi: 10.1016/j.ncl.2008.09.015.
[6] L. Tonarelli, "Magnetic resonance imaging of brain tumor," CEwebsource. Com, 2013.
[7] D. Belkić, K. Belkić, "Signal Processing in Magnetic Resonance Spectroscopy with Biomedical Applications," *CRC Press*, 2010, doi: 10.1201/9781439806456.
[8] B. H. Menze, *et al.*, "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993-2024, 2015, doi: 10.1109/TMI.2014.2377694.
[9] M. E. Tagluk, M. Akin, and N. Sezgin, "Classıfıcation of sleep apnea by using wavelet transform and artificial neural networks," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1600-1607, 2010, doi: 10.1016/j.eswa.2009.06.049.
[10] Y.-D. Zhang, L. Wu, and G. Wei, "A New Classifier For Polarimetric Sar Images*," Progress In Electromagnetics Research*, vol. 94, pp. 83-104, 2009 doi: 10.2528/PIER09041905.
[11] J. Camacho, J. Picó, and A. Ferrer, "The best approaches in the on-line monitoring of batch processes based on PCA: Does the modelling structure matter?," *Analytica Chimica Acta*, vol. 642, no. 1-2, pp. 59-68, 2009, doi: 10.1016/j.aca.2009.02.001
[12] N. B. Bahadure, A. K. Ray, and H. P. Thethi, "Image Analysis for MRI Based Brain Tumor Detection and Feature Extraction Using Biologically Inspired BWT and SVM," *International Journal of Biomedical Imaging*, vol. 2017, ID. 9749108, pp. 1-13, 2017, DOI: 10.1155/2017/9749108
[13] R. P. Joseph, C. S. Singh, and M. Manikandan, "Brain Tumor Mri Image Segmentation And Detection In Image Processing," *International Journal of Research in Engineering and Technology*, vol. 3, no. 1, pp. 1-5, 2014.

[14] M. Alfonse, and A. B. M. Salem, "An automatic classification of brain tumors through MRI using support vector machine," *Egyptian Computer Science Journal*, vol. 40, no. 3, pp. 1-11, 2016.

[15] J. Yao, J. Chen, and C. Chow, "Breast Tumor Analysis in Dynamic Contrast Enhanced MRI Using Texture Features and Wavelet Transform," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 1, pp. 94-100, 2009, doi: 10.1109/JSTSP.2008.2011110.

[16] P. Kumar, and B. Vijayakumar. "Brain tumor MR image segmentation and classification using PCA and RBF kernel-based support vector machine," *Middle-East Journal of Scientific Research*, vol. 23, no. 9, pp. 2106-2116, 2015, doi: 10.5829/idosi.mejsr.2015.23.09.22458.

[17] H. Mohsen, E.-S. A. El-Dahshan, E.-S. M. El-Horbaty, and A.-B. M. Salem, "Classification using deep learning neural networks for brain tumors," *Future Computing and Informatics Journal*, vol. 3, no. 1, pp. 68-71, 2018, doi: 10.1016/j.fcij.2017.12.001.

[18] A. Chaddad, "Automated Feature Extraction in Brain Tumor by Magnetic Resonance Imaging Using Gaussian Mixture Models," *International Journal of Biomedical Imaging*, vol. 2015, no. 8, pp. 1-11, 2015, doi: 10.1155/2015/868031.

[19] J. Sachdeva, V. Kumar, I. Gupta, N. Khandelwal, and C. K. Ahuja, "Segmentation, Feature Extraction, and Multiclass Brain Tumor Classification," *Journal of Digital Imaging*, vol. 26, no. 6, pp. 1141-1150, 2013.

[20] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image Segmentation Using K -means Clustering Algorithm and Subtractive Clustering Algorithm*," Procedia Computer Science*, vol. 54, pp. 764-771, 2015.

[21] R. B. Dubey, "Region growing for MRI brain tumor volume analysis," *Indian Journal of Science and Technology*, vol. 2, no. 9, pp. 26-31, 2009, doi: 10.17485/ijst%2F2009%2Fv2i9%2F29517.

[22] J. Zhao, Z. Meng, L. Wei, C. Sun, Q. Zou, and R. Su, "Supervised Brain Tumor Segmentation Based on Gradient and Context-Sensitive Features," *Frontiers in Neuroscience*, vol. 13, no. 144, 2019, doi: 10.3389/fnins.2019.00144.

[23] S. Sawakare, and D. Chaudhari, "Classification of brain tumor using discrete wavelet transform, principal component analysis, and probabilistic neural network," *International Journal For Research In Emerging Science And Technology*, vol. 1, no. 6, pp. 13-19, 2014.

[24] S. B. Gaikwad, and M. S. Joshi, "Brain Tumor Classification using Principal Component Analysis and Probabilistic Neural Network," *International Journal of Computer Applications*, vol. 120, no. 3, pp. 5-9, 2015.

[25] H. B. Mitchell, "Image Similarity Measures," *Image Fusion*, pp. 167-185, Springer, Berlin, Heidelberg, 2010.

[26] J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson Correlation Coefficient," *Noise Reduction in Speech Processing Springer Topics in Signal Processing*, pp. 1-4, 2009.

[27] 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 Informatics*, vol. 5, no. 1, pp. 23-30, 2018, doi: 10.1007/s40708-017-0075-5.

[28] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610-621, 1973, doi: 10.1109/TSMC.1973.4309314.

[29] A. M. Hasan, and F. Meziane, "Automated screening of MRI brain scanning using grey level statistics," *Computers & Electrical Engineering*, vol. 53, pp. 276-291, 2016, doi: 10.1016/j.compeleceng.2016.03.008

[30] Y. Zhang, H. Wu and L. Cheng, "Some new deformation formulas about variance and covariance," *Proceedings of International Conference on Modelling, Identification and Control*, 2012, pp. 987-992.

[31] W. Mao, and F.-Y. Wang, "Cultural Modeling for Behavior Analysis and Prediction," *Advances in Intelligence and Security Informatics*, pp. 91-102, 2012, doi: 10.1016/B978-0-12-397200-2.00008-7.

[32] S. Misra, and H. Li, "Noninvasive fracture characterization based on the classification of sonic wave travel times," *Machine Learning for Subsurface Characterization*, pp. 243-287, 2020.

[33] R. May, H. Maier, and G. Dandy, "Data splitting for artificial neural networks using SOM-based stratified sampling," *Neural Networks*, vol. 23, no. 2, pp. 283-294, 2010, doi: 10.1016/j.neunet.2009.11.009.

[34] S. Armand, E. Watelain, E. Roux, M. Mercier, and F.-X. Lepoutre, "Linking clinical measurements and kinematic gait patterns of toe-walking using fuzzy decision trees," *Gait & Posture*, vol. 25, no. 3, pp. 475-484, 2007, doi: 10.1016/j.gaitpost.2006.05.014.

[35] N. Nabizadeh, and M. Kubat, "Brain tumors detection and segmentation in MR images: Gabor wavelet vs. statistical features," *Computers & Electrical Engineering*, vol. 45, pp. 286-301, 2015, doi: 10.1016/j.compeleceng.2015.02.007.

[36] P. Dvořák, W. Kropatsch, and K. Bartušek, "Automatic Brain Tumor Detection in T2-weighted Magnetic Resonance Images," *Measurement Science Review*, vol. 13, no. 5, pp. 223-230, 2013, doi: 10.2478/msr-2013-0034.

BIOGRAPHIES OF AUTHORS
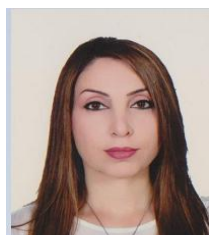
Dalia Mohammad Toufiq received the the B.Sc. in Computer and Statictics Science from University of Sulaimani, Sulaimani Iraq in 2004, the M.Sc. in Information Technology from Jawaharlal Nehro Techonolgical University (JNTUH), Hyderabad, India in 2011, working as Lecturer at University of Sulamani from 2012–Now, she is a Ph.D student in Computer Science.

Ali Makki Sagheer was born in Iraq-Basrah 1979. He got on B.Sc. of Information System in Computer Science Department at the University of Technology (2001)-Iraq, M.Sc. in Data Security from the University of Technology (2004)-Iraq and Ph.D. in Computer Science from the University of Technology (2007)-Iraq. He is interesting in Cryptology, Information Security, Cyber Security Number Theory, Multimedia Compression, Image Processing, Coding Systems and Artificial Intelligence.

Hadi Veisi received the B.Sc. in Computer Engineering-Hardware, from Shiraz University–Shiraz, Iran in 2003, the M.Sc. in Artificial Intelligence, Sharif University of Technology–Tehran, Iran in 2005, and Ph.D. degree in Computer Engineering, Specializing in Speech Processing and Natural Language Processing and Neural Networks, Sharif University of Technology–Tehran, Iran in 2011. Co-founder of 3 companies in NLP, CEO of 2 companies and the managing director of RefaTech innovation center.