Modified Tumor Diagnosis by Classification and Use of Canonical Correlation and Support Vector Machines Methods

Hamed Samadi Ghoshchi, Yaghoub Pourasad*
Faculty of Electrical Engineering, Urmia university of Technology, Urmia, Iran

y.pourasad@uut.ac.ir
Tel: +98 912 640 6595

Abstract

The main objective of this research is to investigate techniques for classifying tumor grades based on image processing. The algorithms to classify tumors are introduced, and their performance for the experimental results are investigated. In the proposed algorithm, first, the scan images of the lung are pre-processed, and then the histogram, texture, and geometric features are extracted. These characteristics are then employed in the Support Vector Machines (SVM) and Canonical Correlation Analysis (CCA) classifiers to diagnose tumors and classify benign and malignant types. These integrated approaches in investigating medical images are a vital tool to improve the diagonalization accuracy. In the current research, experimental and simulated medical images are employed. The outcomes of the developed techniques in this research are compared with the literature review to confirm the proposed approach's efficacy and reliability in diagnosing and classifying tumors. In addition to high accuracy in diagnosis, this method is also a low-cost and low-risk method. Due to its very high sensitivity and having the desired values of two criteria of precision and specificity, and the low number of features used for classification, the developed method was proposed as an efficient and appropriate method for tumor classification.

Keywords: Tumor detection, Classification, CCA, SVM, Image Processing

1- Introduction

There are two common ways to diagnose a tumor: based on computational and physical characteristics. In order to determine computational features such as standard deviation, color intensity, entropy, etc., applied algorithms are used to detect tumors [1-3]. Computational features can be used to check for the presence or absence of heterogeneous tissues in the brain. Medical images mainly include the appearance of the brain, such as cerebrospinal fluid, cranial bone tissue, gray and white particles, as well as information about tissue contrast, resolution, and location of tissue masses. Diagnosis based on these features has been gained importance due to detection in the initial stages of the tumor [4, 5]. In [3], the identification of the brain tumor area is described by predicting the tumor type and the boundary box. Magnetic resonance imaging (MRI) brain tumor images have been trained faster than baseline using Region-Based Convolutional Neural Networks (R-CNN). The faster R-CNN combines the AlexNet and region proposal network models. In [6], tissue features and possible neural network one have been used to diagnose malignant, malignant, and normal tumors. In this algorithm, wavelet conversion coefficients in different bands are used. In [7], classification of four tumors of astrocytoma, meningioma, carcinoma, and sarcoma using matrix-based features of co-occurrence of gray surfaces, network classification Neon, and the Levenber-Markard nonlinear optimization algorithm. In [8], neural network classifications of radial base function and support vector machine (SVM) have been used to detect normal and abnormal brain tumors using matrix features of co-
occurrence of extracted gray surfaces from MRI images and tomography. The reported accuracy on a database of 123 MRI images, without the application of color channels, is 73%, with a sensitivity of 77% and a specificity of 68%. In [9] have proposed an interpretive approach for converting four-dimensional data so that standard Convolution Neural Networks (CNN) two-dimensional architectures can also be utilized to do brain tumor segmentation. The findings presented on the brain tumor segmentation dataset indicate a dice score of 7.83% for the total tumor region, 6.73% for the tumor core region, and 69% for the active tumor region. In [10], segmented brain tumor segmentation data using a simple SegNet network. The average dice accuracy reported on the brain tumor segmentation database was 6.87%. The main problem in the methods proposed so far has been the low accuracy rate, and various methods have tried to increase the accuracy of the classification. In this article, we have tried to increase the accuracy of classification as much as possible based on the method based on the proposed model and using various useful features in this regard.

In recent studies, the diagnosis and classification of brain tumor diseases have been presented, such as research by image processing techniques in the diagnosis of brain tumors [11], brain tumor diagnosis using fragment MRI images by K-mean algorithm [12], the use of digital image processing in the diagnosis of brain tumors and image feature calculation techniques [13], detection of brain tumors by proximity in the nearest neighborhood and then the fragmentation of images by the waterfall algorithm [14], diagnosis Location of brain tumor using the intensity of MRI [15], diagnosis of a brain tumor in medical images using K-mean violet conversion [16]. MRI image processing and extraction of the best feature by particle swarm optimization and Genetics algorithm, then fuzzy clustering have been used to diagnose brain tumors [17]. In [18] the comparison of the random forest can also be used in addition to SVM for the Brain tumor segmentation and detection. One of the benefits of Random Forest, which exists me most, is the power of handling large data sets with higher dimensionality and accuracy. It does an excellent job at classification but not for regression issues as it did not provide an accurate continuous nature forecast. The regression case did not predict beyond the trained data variety and might overfit especially noisy data sets.

In this article, in order to improve the diagnosis of brain tumors and quick access to the location of the brain tumor mass, we have divided the brain tumor and calculated the area in the brain images using the classification method. The proposed method in this paper offers promising results when compared to the brain tumor segmentation detection system. In this study, the proposed architecture was also used for the gastric cancer dataset and achieved better performance. From the findings, it is obvious that tumor identification is predictable by utilizing Canonical Correlation Analysis (CCA), and this prediction's efficacy is appropriate.

2- Methodology

In the proposed method, we used two types of Flair and T1W brain MRI images. They should be preprocessed to normalize their pixel situation and extract extra data like the noise. Both of these images, after preprocessing, move to the fenestration state to extract the features already discussed. Based on the previous post-windowing process, we have a very important step in extracting the tissue and statistical properties that apply to each Flair and T1W brain MRI image. In the next step, which is the innovation of the proposed method, we combine the extracted features of both images using the CCA algorithm to obtain a number of new features, then reduce these new features and accept the classification over them. Based on the obtained data, we have two parts of training and trial at the classifiers step. We used trained and real images for this plan. These images are chosen by utilizing a batch K-cross validation algorithm. Lastly, the post-processing and benchmarking steps are performed.

The proposed approach in the current research (Figure 1) added new features in comparison to the reference [3], to overcome its disadvantages in accuracy, which are two CCA and a
batch K-cross validation algorithms. The batch K-cross validation algorithm produces stochastic K values, which are stochastic. The process K is reproduced to give great fidelity to the findings.

The brain tumor classification system presented in this article has an algorithm that is described below. This algorithm includes multiple steps such as: preprocessing, windowing, feature extraction, feature size reduction, classifier, and post-processing and performance calculation. Figure 1 shows a general diagram of the tumor classification system.

First, MRI images taken from the patient's brain require preprocessing of images in multiple classification algorithms. This stage (preprocessing) includes various parts such as noise cancellation, normalizing the brightness of brain MRI images, smoothing the size of brain MRI images, and removing extra information that is in the background of brain MRI images. The next step in this classification system is to brainstorm MRI images. This step consists of two parts: 1- Overlapping windows; 2. None-overlapping windows.

In the third step, which is one of the main steps of this system, from the windows formed in the previous step, we extract the features given as input to the classifier. It should be noted that for each window, a feature vector is provided.

In the fourth stage, which is called reducing the dimensions of the feature, is done due to a large number of extracted features. In the next and last step, the available attribute vectors are taught as input to the classifier. Each of these feature vectors is labeled healthy or tumorous. How to determine if it is healthy or tumorous is that in the window testing phase, if that window is in the healthy category, the center pixel of that window will be repeated as a tumor. After completing these steps, we apply the post-processing to reduce the error and calculate the efficiency of this method.

Figure 1: Schematic of image processing

One of the characteristics of first-time statisticians is the image histogram. In other words, the information related to the first-order statisticians is obtained by the image histogram. The image histogram also provides a summary of the statistical data of the image [19]. Among the characteristics of first-order statistics, entropy energy, mean, mean, elongation, asymmetry, and mean contrast (variance) can be mentioned.

1- Energy

$$E = \sum_{i=0}^{G-1} [p(i)]^2$$

(1)

2- Entropy: Indicates the degree of uncertainty in light intensity values.

$$H = -\sum_{i=0}^{G-1} p(i) \log_2[p(i)]$$

(2)

3- Average: Shows the average brightness of the whole image

$$\mu = \sum_{i=0}^{G-1} ip(i)$$

(3)

4- Skewness: Shows how flat the histogram is

$$\mu_3 = \sigma^{-4} \sum_{i=0}^{G-1} (i - \mu)^4 p(i) - 3$$

(4)

5- Asymmetry: Shows the standard deviation of the histogram around the mean

$$\mu_4 = \sigma^{-3} \sum_{i=0}^{G-1} (i - \mu)^3 p(i)$$

(5)

6- Variance: Shows changes in light intensity around the mean.
\[
\sigma^2 = \sum_{i=0}^{G-1} (i - \mu)^2 p(i) \tag{6}
\]

In the above relations, respectively: (G): number of gray surfaces of the image; P (i): The probable density of light intensity levels, expressed as follows:

\[
p(i) = \frac{h(i)}{N} \tag{7}
\]

Where, h (i) is the number of all pixels whose brightness value is I, and N represent the total pixel numbers.

2-1 Characteristics of second-order statistics

The Gray Surface Length Matrix (GLRLM) is a second-order statistic and spatial method. This matrix is such that a primary texture called the gray surface row length is considered, and the most significant number of pixels in a row that have the same gray surface. These rows have a row length and direction for a gray value [20]. To calculate this matrix, we first need to specify the number of rows of a gray area with row length; for example, in Figure 2, we see a 4-4 image that has four gray areas between the numbers 0-3. In order to be able to create the GLRLM matrix for this image, we must first consider four different directions. 0, 45, 90, and 135 degrees are the directions that should be used. As can be seen in Figure 3, the calculation of this matrix is clear. For example, in Part A of Figure 3, we first look at the gray surface row, the first row has the number zero, then we reach the length of the row, looking at the image matrix, we see that the zero pixels in the 0-degree direction are repeated four times individually. In other directions, no duplication of zero pixels is observed in the 0-degree direction. Therefore, during rows 2, 3, 4, the gray area becomes zero. This trend continues for other angles as well.

Figure 2: Hypothetical image matrix

Gray Surface Length Matrix

Figure 3: Gray Surface Length Matrix

In the figure 3 (a) is for 0 degrees, (b) for 45 degrees, (c) for 90 degrees and (d) for 135 degrees. The GLRLM matrix has the following properties:

1- Short length rows, which according to the following equation, the length value of each row is divided by the power of two (squares) of each row.

\[
RF_1(R(\theta)) = \frac{1}{T_p} \sum_{i=0}^{G-1} \sum_{l=1}^{N_a} \frac{r'(i,l|\theta)}{l^2} \tag{8}
\]

2- Long length rows that according to the following equation, the length of each row is multiplied by the power of two (squares) of each row.

\[
RF_2(R(\theta)) = \frac{1}{T_p} \sum_{i=0}^{G-1} \sum_{l=1}^{N_a} j^2 r'(i,l|\theta) \tag{9}
\]

3- Gray area distribution According to the following equation, first the number of row lengths is calculated for each amount of gray area and then this value is added and then they must be normalized to be able to calculate the gray area distribution.
\[ RF_3(R(\theta)) = \frac{1}{T_P} \sum_{i=0}^{G-1} \sum_{l=1}^{N_R} r'(i, l \mid \theta) \]  \hspace{1cm} (10)

4- The row length distribution is calculated from the following equation for the image:

\[ RF_4(R(\theta)) = \frac{1}{T_P} \sum_{i=0}^{G-1} \sum_{l=1}^{N_R} r'(i, l \mid \theta) \]  \hspace{1cm} (11)

Row Percentage according to the following equation, the ratio of the number of all rows to the number of possible rows is obtained.

\[ T_P = \sum_{i=0}^{G-1} \sum_{l=1}^{N_R} r'(i, l \mid \theta) \]  \hspace{1cm} (12)

Where, G: number of gray surfaces; P: total number of pixels; NR: Number of longest rows (number of row lengths); TP: total number of rows in images; TP is also used to normalize images.

Linear Binary Pattern (LBP) Features

The LBP performs a windowing throughout the image that can sweep the entire image. It then compares the center pixel value with the surrounding pixel and assigns binary numbers (0 and 1) to the surrounding pixels. In the next step, the LBP multiplies these numbers zero and one by the powers of two clockwise or counterclockwise, and finally adds them up. Finally, the resulting histogram represents the target tissue. For the pixel neighborhood, (R and P) are used, where P is the number of sample points on a circle with radius R. The LBP pixel value (YC and XC) is calculated from the following equations.

\[ LBP_{p,k} = \sum_{p=0}^{P-1} S(g_p - g_r) 2^p \]  \hspace{1cm} (13)

\[ S(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \]  \hspace{1cm} (14)

2-2 Features of Oriented Gradients (OG)

The operation of this actuator is such that by distributing the gradient intensities, it is able to reveal the appearance and shape of the tumor or the desired objects. This operator, one of the most important operators in object detection, is calculated on a space known as a cell and has a uniform network. Local contrast normalization can be used to increase the accuracy of this operator. In general, the summary of feature extraction from brain MRI images is summarized as follows:


2-3 System classification

In the tumor classification system, the classification stage is one of the most important stages. This stage of the system is after the stage of reducing the number of features extracted. In this step, attributes are given to the classifier input. The job of the classifier is to convert the input data, which is quantitative, into qualitative output data. This output is a real value or a vector.
Classification of brain tumors for MR images using accurate transfer learning and quantification.

Accurate classification of MRI of brain tumors plays an essential task in clinical analysis and decision-making for therapy. The critical issue in analyzing MRI is the semantic gap in the low-level visual data recorded via the MRI machine and the high-level data recorded via the human evaluator. Traditional machine learning systems for classification concentrate just on high and low-level characteristics. It utilizes just some manual elements to overcome this gap and requires proper characteristic classification and extraction techniques.

Recent developments in deep learning have shown tremendous progress, and deep CNNs have emerged as successful in task classification. Deep learning has become influential in providing features that can fully display low and high-level data and embed the characteristics analysis in learning but generally requires data sets [11]. For most medical learning scenarios, this training dataset is small. So applying deep learning and training to CNN is challenging from the start in a small dataset. Also, the hardware and computational time limitations are the other factors that convinced us to use Machine learning over deep learning.

As shown in Figures 4 (a) and 2 (b), Meningioma is generally squares of gray matter and cerebrospinal fluids. The pituitary tumor was placed nearby the sphenoid sinus, the internal carotid arteries, and the optic cysts. Glioma appears to be in an unequal pattern, as depicted in Figures 4 (c) and (d). So, the most crucial data and diagnostic characteristics of brain tumors were associated with the area of the tumor in the MRI image along with its border, tissue, intensity, and pattern [5].

SVMs are one of the supervised learning methods that are used for classification and regression. This technique is one of the relatively novel ones that has enhanced proper and acceptable efficacy in compared to older methods for classification, including perceptron neural networks. The basis of the SVM classifier is the linear classification of data, and in linear segmentation of data we try to choose the line that has the most confidence margin. Solving the optimal line equation for data is done by QP (Quadratic Programming) methods, which are known methods for solving constrained problems.

The simplest interpretation that can be given for the operation of the backup vector machine function is that in the backup vector machine algorithm, the distance of the nearest data sample to the separator line (boundary between categories) is calculated (these data samples may belong to any category). SVM uses a technique called kernel trick to convert your data and then, based on that conversion, finds the optimal boundary between possible outputs (see figure 5). In simple terms, it performs very complex conversions, then specifies how to separate your data based on the tags or outputs you have defined. One of the methods that is currently widely used for the classification problem is the SVM method. Perhaps the current
The popularity of the SVM method can be compared to the popularity of neural networks over the past decade. The reason for this is the ability to use this method to solve various problems, while techniques like the decision tree cannot be easily utilized in different type applications.

Figure 5: Data scatter diagram in support vector

The CCA algorithm was developed by [4], a statistical method that deals with the interaction between two random vectors and are used in multivariate statistical analysis. This algorithm measures the linear relationship between two multidimensional random variables. For each of those multidimensional variables, the CCA finds a base vector that has optimal conditions in terms of cross-correlation, and at the same time calculates the cross-correlation of two multidimensional variables; In other words, this algorithm seeks to find two base vectors in which the correlation matrix between the variables is diagonal, and its diameter has maximum values.

Suppose \( X = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \) and \( Y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n \) there are two sets of feature vectors with assumed dimensions that are extracted from n experimental samples. The CCA tries to find the two base vectors \( w_x \) and \( w_y \) for the two variables X and Y in such a way that the correlation between the reflection of the two vectors on those Canonical bases, \( x_i^* = w_x^T (x_i - \bar{x}) \) and \( y_i^* = w_y^T (y_i - \bar{y}) \) as \( i = 0, 1, \ldots, n \) \( \bar{x} \) which \( \bar{x} \) is the mean of X and the \( \bar{y} \) mean of Y, is mutually maximized. In other words, the relation (15) should be maximized:

\[
\rho = \frac{\mathbb{E}[xy]}{(\mathbb{E}[x^2] \mathbb{E}[y^2])^{1/2}} = \frac{\mathbb{E}[w_x^T X Y^T w_y]}{(\mathbb{E}[w_x^T X X^T w_x] \mathbb{E}[w_y^T Y Y^T w_y])^{1/2}} \quad (15)
\]

After simplifying the relation (15), the relation (16) is obtained:

\[
\rho = \frac{w_x^T C_{xy} w_y}{(w_x^T C_{xx} w_x w_y^T C_{yy} w_y)^{1/2}} \quad (16)
\]

Which \( C_{xx} \) and \( C_{yy} \) are the covariance matrices of x and y, respectively, as well as the \( C_{xy} \) mutual covariance matrix between x and y. After optimizing the relation (16), two relations (17) and (18) are obtained:

\[
Z_1 = \begin{bmatrix} X^* \\ Y^* \end{bmatrix} = \begin{bmatrix} w_x & 0 \\ 0 & w_y \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} \quad (17)
\]

\[
Z_2 = X^* + Y^* = \begin{bmatrix} w_x & w_y \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} \quad (18)
\]

Which \( Z_1 \) and \( Z_2 \) are called distinctive Canonical Correlation Discriminant Features (CCDFs), which are actually a combination of features extracted by both aggregation and addition methods. The method used in this dissertation has new characteristics in comparison to the reference [3], which are the application of two CCA algorithm as well as the K-category cross- The batch K-cross validation algorithm produces stochastic K values, which
are stochastic. As shown in figure 6, the process K is reproduced to give great fidelity to the findings. The patch-based mean-k method is also used to empty the skull (brain tissue extraction) as a preprocessing step, which has the added property of classifying the element depending on its neighborhood patterns. After the skull is removed, the classifier is applied to the brain tissue, and finally, the brain tumors are used using shape-based topological properties.

Figure 6: Create a mean skull-k patch. (a) Original image (light gray) with zero plug (dark gray) (b) patch matrix

In the windowing stage, the properties that are obtained contained 315 features, 32 were related to the directional histogram, twenty were for the row matrix of gray surface, 256 were for linear binary features, and the remained seven are for first-order ones. So, it can be derived that each window provide a property vector with dimensions of $315 \times 1$. In addition, as mentioned earlier, each brain MRI image has 4,096 windows with the size of 10×10. As a result, for per brain MRI image we have a characteristic matrix via dimensions of $4096 \times 315$.

3- Results
As mentioned earlier, the images available to us contains simulated images and real MRI ones. It should be noted that applying the pre-processing step is distinct for all simulated and real images. Simulated images firstly were required to convert the format of the images taken from the database to the standard one, such as the bmp. Next, the pixel location of the MRI images normalized. For this purpose, the size of the image should be changed to 200x200 pixels to achieve this by removing the extra pixels. In order to pre-process real images, in additional to the steps done for the simulated images, two more actions are needed. Normalizing the light intensity and implementing the histogram matching are the same steps mentioned.

3.1 Normalize the light intensity
It is one of the important steps in preprocessing. Normalization of light intensity is based on the standard deviation of MRI. The basis of this method is to calculate the standard deviation of the MRI image columns and then divide the pixels of each column by the corresponding standard deviation.

3.2 Application of histogram algorithm
In this algorithm, we act by considering an image from a group of real images as a reference ones, then calculating and matching them. The selected reference image using personal experience Researchers as well as measuring relevant criteria are achieved. Fig. 7 depicts the actual MRI image as well as the calculated histogram as a graph. Figure 8 illustrates the actual MRI image with its histogram with/without employing the algorithm.
3.2 Results of SVM classification
- The backup vector machine classifier has four main kernels: linear kernel, radial base function kernel, Gaussian kernel, and polynomial kernel.
Using each of these kernels, the SVM classifier is applied to the entire category of textural and geometric features and then to each category of features, and the best performance of each recording mode and the best kernel in which the highest number in each of criteria is obtained, will be selected as the winning category. Next, Figure 9 shows the criteria values for the backup vector machine classifier with all features. In all the diagrams in this chapter, the values of the three criteria of accuracy, sensitivity, and specificity are specified in three colors: blue, orange, and yellow, respectively.
Next, in Figures 9, the six SVM kernels are shown, first considering each feature set and then a dual combination of feature sets. The criteria for superior classification in each of the cases are given in Tables 1-3. According to the diagrams, it can be seen that the two linear and Gaussian coarse classifications give the best results. Therefore, in the case of selecting all the features, these two classifiers are selected as the top SVM classifiers. In the case of selecting each feature category, the histogram feature category with linear kernels and the texture feature category with linear kernels and Gaussian coarse have the best performance. Finally, in selecting a dual combination of features for the two modes, histogram texture and texture-geometric, with linear and coarse Gaussian kernels, the highest accuracy, sensitivity, and specificity criteria are obtained.

Table 1: Results of the criteria of the top two categories of backup vector machines with all features

Table 2: Results of the criteria of the top two categories of backup vector machines with individual characteristics

Table 3: Results of the criteria of the top two classifications of the support vector with a dual combination of features

3.3 Results of K-nearest neighbor (KNN) classification
In this section, using the KNN classifier, properties are classified and the criteria of accuracy, sensitivity and degree of specificity are calculated. In the KNN classifier, the four values of 3, 5, 7 and 9 are selected as the K parameter and the distance criterion is of the Euclidean type.
In this section, like the SVM classifier, first all the features, then each feature and finally a dual combination of features are selected for classification and the top classifier (here K top) is selected.
The classification results are shown in Figures 10. By examining these graphs, it can be seen that in the case of selecting all the features, the selection of 9 nearest neighbors, ie K = 9, gives the best criteria. Also, in selecting each feature category alone, texture properties give the best results by selecting K = 9. Finally, in the selection of two feature categories, the combination of texture and geometric feature categories for K = 7.9 results in the highest accuracy, sensitivity and specificity, which are equal to the selection of all feature properties.
for \( K = 9 \) is. Therefore, in the KNN classifier, the combination of two categories of texture and geometric properties with \( K = 7, 9 \) is selected as the superior classifier. The values of the criteria for each different feature category and \( K \) are given as a percentage in Tables 4-6.

Figure 10: Comparison of system performance in k classifier of nearest neighbor and weighted at distances \( k = 3, 5, 7, 9, 11 \)

Table 4: Results of KNN Top Classification Criteria with all features

Table 5: Results of KNN top classification criteria with individual characteristics

Table 6: Results of KNN superior classification criteria with a dual combination of features

3.4 Database and simulated parameters

The aim of this paper is to provide a method for classifying real and simulated brain MRI images in order to improve the three criteria for evaluation. Brain MRI images have similarities in the brightness of their pixels, making it difficult to distinguish healthy tissue from tumor tissue and to perform functional classification of the tumor. According to the routine of the classification system, the statistical features of these images are obtained by employing the aforementioned pre-processing and windowing. The dynamic range of the obtained values is various conditions due to the variety of the derived properties, which is a cause to the imbalance in the property matrix. According to this intend, we normalize the dynamic range to eliminate or minimize these differences. The method is that first in each property matrix column, the mean value of the column is subtracted of the related column and then divided by the column standard deviation.

3.5 Simulation results

In the initial step, the real brain MRI are evaluated, then for comparison of classifier performance; in the next step, the findings of the simulated MRI images are reported. These results are performed by applying three classifiers K-means, SVM, CCA. The focal correlation is employed to enhance the representation of the proposed approach. Table 7 presents the outcomes of the K-means, SVM, and CCA algorithms.

Table 7- Various algorithms performance for MRI for \( K=2 \)

Since the election of \( K=2 \), It was found that the procedure is duplicated, and the rates determined for each model in the determination of the offered approach are double the mean of the process. First group comprises the trained brain MRI image, and the other the trial group. As the multiple of 72 (\( K = 2, 3, 4, 6, 9 \)) were used for the selection, we now have to assign \( K \) to the number 3 in the order in which it exists. In this case, 24 images should be assigned to the training group and 12 to the experimental group. The procedure is repeated three times, and the trained and real groups are transferred once the procedure is completed. Table 8 presents the outcomes for \( K=3 \), which are presented in Table 8:

Table 8- Various algorithms performance for MRI images for \( K=3 \)

The findings for simulated brain MRI images are done and investigated for various Ks based on the before mentioned process. These outcomes are presented in Tables 9 and 10 for the simulated images.

Table 9- Various algorithms performance for MRI images for \( K=2 \)
3.6 Integrate extracted features to increase accuracy

In studies related to the classification and diagnosis of diseases with medical image processing, the results are usually expressed by three criteria: accuracy, sensitivity and degree of specificity:

TP: The number of malignant glands classified as malignant glands.
FP: The number of benign glands classified as malignant.
FN: The number of malignant glands classified as benign glands.
TN: The number of benign glands classified as benign glands.

The above relationships can be easily extracted from the confusion matrix, which is a matrix to express the performance of the classification model and displays the distribution of the categories in terms of correctness or incorrectness. Figure 6 shows the binary deformation matrix (two classes).

Figure 11 presents a correlation diagram of sensitivity patterns by applying a focal correspondence interpretation in simulated brain MRI images according to the various k values in batch k cross-validation procedure, utilizing the K-means classifier. Based on the amount of simulated brain MRI, if k = 2, 20 images are chosen as trained images and remained as trial ones, and the amount of iterations is 2. As k = 3, then 26 images are picked as trained images and the others as trial ones, and the amount of repeat will be 3. If k = 6, 30 they are chosen as trained images and the others as trial ones, and the amount of repetitions will be 6 and so on for the k=9.

As presented in Fig. 11, using the K-means algorithm, implementing the CCA to simulated brain MRI improved the sensitivity criteria up to 0.6% to 1.56%, indicating that tumor pixels' apprehension ratio has risen. Figure 10 designates comparison diagrams of sensitivity using the focal correlation procedure for classifying brain tumors in real images according to the various k values in batch cross-validation utilizing the K-means. The amount of actual brain MRI images handled is 40 in this condition. In this experiment, k = 2, k = 3, k = 6 and k = 9 have taken. Tumor classification outcomes reveal that the sensitivity pattern increases utilizing the CCA employed at the characteristic extraction step, specially for real brain ones. Employing the CCA algorithm developed the sensitivity pattern up to 22%. The three parameters of specificity, sensitivity, and accuracy criteria were evaluated for the K-MEANS classifier in the stated and the reference approach [20]. The total amount of images applied in the [21] was 26, whilst in this study, 70 images were utilized, so the database employed here is more extensive than the reference so that the results can be reliable [21].

In [21], only the accuracy factor is considered, and two other factors have not been listed. In the present research, just the most beneficial outcomes from the reference [21] to the SVM classifier were discussed, and the consequences that were lower than the standard accuracy are withdrawn.

In [21] to enhance safety, the k-batch cross-validation procedure is just utilized for k = 10, i.e., most of the data is employed as trained images, and only about 15% employed as experimental ones. However this research, we selected the various k = 2, k = 3, k = 6, and k = 9, which utilize 50%, 34%, 25%, 15%, and 10% of the data as experimental ones, respectively, which made the intended approach in the current work more authentic than.

Figure 12 is proposed for the sensitivity comparison of simulated brain MRI images classifiers in the current developed method and [21].
were not presented in [21] for classifying real brain MRI, just the presented approaches charts in the current research utilize K-means, CCA, and SVM. The real brain MRI is presented in Fig. 10 for the precision factor.

Figure 12- Accuracy factor for three classifiers (simulated)

As presented in fig. 12, the sensitivity factor in simulated brain MRI using the CCA procedure raised up to 3% than [20] regarding to the amount of various batches (various ks) in the k batch cross-validation algorithm. It is clear from the fig. 12, the application of K-MEANS functions better than other ones. This research concentrates more on the effectiveness of classifier in the intended CCA formation.

Figure 13- Accuracy factor for three classifiers (experiment)

As depicted in fig. 13, the accuracy factor in real brain MRI images using the CCA procedure improved up to 2% compared to [21]. Furthermore, by estimating the precision of real images and linking the three intended purposes for classification via [21], it can be concluded that applying the CCA algorithm enhances the system's performance. Additionally, by analyzing the three applied approaches in the current research for classification, it was found that K-means do better than the SVM and CCA and upgrade system efficiency.

4 Conclusion

Diagnosis of the disease by a physician by viewing MRI images and determining the tumor's exact location to improve the treatment process will be very significant. Computerized techniques provide more accurate and comprehensive results, so in this study, with the aim of providing image processing to support physicians in making decisions in this regard, the general results of this study were performed by preprocessing, and finally, the desired features are extracted and the tumor area was obtained. In this study, a new and efficient method for classifying tumors from images was proposed. Then, using the mentioned characteristics, tumors were classified into two categories: benign and malignant. The proposed classification algorithm was able to classify tumors with 90%, 100%, and 89% accuracy, sensitivity, and specificity, respectively. Using this method in medical centers can be of great help to doctors and radiologists to increase the accuracy of diagnosis and saving time.

By comparing the results of the proposed method with the previously presented methods, the proposed method obtained the highest sensitivity, and in many studies, the highest classification accuracy. Considering that the criterion of susceptibility means the ability of the classifier to correctly diagnose the disease [22, 23] and the criterion of the degree of specificity means the ability of the classifier to diagnose the absence of the disease correctly, so it can be concluded that the criterion of susceptibility in research related to endocrine diagnosis Medical images are of great importance. This is because correctly diagnosing the presence of a disease or cancer is much more important and vital than diagnosing its absence in a candidate. Therefore, the proposed method is an efficient and appropriate method for tumor classification due to its very high sensitivity and the desired values of two criteria of accuracy and specificity, and the low number of features used for classification. Following recommendations provided for future studies:

1- To increase accuracy, other classifiers such as artificial neural network or KNN can be used. 2- In this paper, to increase the sensitivity criterion, a combination of features extracted from two types of brain MRI images T1-w and T2-flair based on CCA algorithm was used. To increase the accuracy, it can be used all four types of images and integrate their features.
3- One of the features of brain MRI images is the wavelet transform feature. In addition to extracting the statistical features used in this article, by using wavelet transform can increase the accuracy. 4- There are different methods to reduce the dimensions of the feature matrix, which can be used by the independent component analysis. 5 - One of the methods to change the three variables and evaluation criteria is to change the size of overlapping windows; the size of overlapping windows can be changed from 10x10 to 20x20, which reduces the computational complexity as much as possible.

References:


**Yaghoub Pourasad** is an assistant professor at the Urmia University of Technology. He his BS and MS degrees in Electrical Engineering from the K. N. Toosi University of received in 2002 and 2004, respectively, and his PhD degree in Electrical Engineering Technology
Hamed Samadi Ghoushchi is a researcher at the Urmia University of Technology. He holds an MSc in Electrical Engineering in 2018. His current research interests include Image Processing, Deep learning and Bioelectrical Engineering.


Figures

Figure 1: Schematic of image processing

Figure 2: Hypothetical image matrix

<table>
<thead>
<tr>
<th>Row Length/Gray Surface</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Row Length/Gray Surface</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 3: Gray Surface Length Matrix

<table>
<thead>
<tr>
<th>Row Length/Gray Surface</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Figure 4. Brain tumors of on weighing T1. The area in the rectangle includes the tumor. (A) The meningioma is nearby the skull; (c) glioma containing edema and necrosis; And (d) glioma surrounded via inflammation.

Figure 5: Data scatter diagram in support vector.

Figure 6: Create a mean skull-k patch. (a) Original image (light gray) with zero plug (dark gray) (b) patch matrix.
Figure 7 - images and its references histogram

Figure 8 - image, during using the algorithms

a) all features

b) histogram characteristics
c) textural characteristics

d) geometric properties

e) histogram and texture characteristics

f) histogram and geometric features

g) Geometric and textural properties

Figure 9: Comparison of the results of six support vector classifiers
a) all features

b) histogram features

c) tissue

d) geometric features

e) histogram and tissue features

f) histogram and tissue features
g) Geometric and textural features

Figure 10: Comparison of system performance in k classifier of nearest neighbor and weighted at distances $k = 3, 5, 7, 9, 11$

Figure 11- Sensitivity for K-means classifier by using canonical correlation
Figure 12 - Accuracy factor for three classifiers (simulated)

Figure 13 - Accuracy factor for three classifiers (experiment)
Table 1: Results of the criteria of the top two categories of backup vector machines with all features

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Coarse Gaussian</th>
<th>Linear</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>90.8</td>
<td>90.8</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Specificity</td>
<td>89</td>
<td>89</td>
</tr>
</tbody>
</table>

Table 2: Results of the criteria of the top two categories of backup vector machines with individual characteristics

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Histogram</th>
<th>Tissue</th>
<th>Geometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top Classifier</td>
<td>Linear</td>
<td>Linear, Coarse Gaussian</td>
<td>Linear</td>
</tr>
<tr>
<td>Accuracy</td>
<td>90.8</td>
<td>90.8</td>
<td>80</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Specificity</td>
<td>89</td>
<td>89</td>
<td>78</td>
</tr>
</tbody>
</table>

Table 3: Results of the criteria of the top two classifications of the support vector with a dual combination of features

<table>
<thead>
<tr>
<th>Criteria</th>
<th>+ Histogram Tissue</th>
<th>+ Histogram Geometry</th>
<th>+ Geometry Tissue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top Classifier</td>
<td>Linear, Coarse Gaussian</td>
<td>Linear</td>
<td>Linear, Coarse Gaussian</td>
</tr>
<tr>
<td>Accuracy</td>
<td>90.8</td>
<td>86.2</td>
<td>90.8</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Specificity</td>
<td>89</td>
<td>84</td>
<td>89</td>
</tr>
</tbody>
</table>

Table 4: Results of KNN Top Classification Criteria with all features

<table>
<thead>
<tr>
<th>Criteria</th>
<th>K=9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>90.8</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>100</td>
</tr>
<tr>
<td>Specificity</td>
<td>89</td>
</tr>
</tbody>
</table>

Table 5: Results of KNN top classification criteria with individual characteristics

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Histogram</th>
<th>Tissue</th>
<th>Geometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-number</td>
<td>K=9</td>
<td>K=9</td>
<td>K=9</td>
</tr>
<tr>
<td>Accuracy</td>
<td>86.2</td>
<td>90.8</td>
<td>76.9</td>
</tr>
<tr>
<td>sensitivity</td>
<td>85</td>
<td>100</td>
<td>71</td>
</tr>
<tr>
<td>------------</td>
<td>----</td>
<td>-----</td>
<td>----</td>
</tr>
<tr>
<td>specificity</td>
<td>87</td>
<td>89</td>
<td>78</td>
</tr>
</tbody>
</table>

Table 6: Results of KNN superior classification criteria with a dual combination of features

<table>
<thead>
<tr>
<th>Criteria</th>
<th>+ Histogram Tissue</th>
<th>+ Histogram Geometry</th>
<th>+ Geometry Tissue</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-number</td>
<td>K=7.9</td>
<td>K=5</td>
<td>K=7.9</td>
</tr>
<tr>
<td>Accuracy</td>
<td>89.2</td>
<td>86.2</td>
<td>90.8</td>
</tr>
<tr>
<td>sensitivity</td>
<td>92</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>specificity</td>
<td>88</td>
<td>84</td>
<td>89</td>
</tr>
</tbody>
</table>

Table 7- Various algorithms performance for MRI for K=2

<table>
<thead>
<tr>
<th>Classification method</th>
<th>Ref</th>
<th>CCA</th>
<th>SVM</th>
<th>K-means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage of sensitivity</td>
<td>----</td>
<td>92.5</td>
<td>91.5</td>
<td>91.5</td>
</tr>
<tr>
<td>Percentage of specificity</td>
<td>----</td>
<td>95.1</td>
<td>94.2</td>
<td>95.1</td>
</tr>
<tr>
<td>Percentage of accuracy</td>
<td>92.1</td>
<td>95.0</td>
<td>93.7</td>
<td>953</td>
</tr>
</tbody>
</table>

Note: Canonical Correlation Analysis (CCA); Support Vector Machine (SVM)

Table 8- Various algorithms performance for MRI images for K=3

<table>
<thead>
<tr>
<th>Classification method</th>
<th>Ref</th>
<th>CCA</th>
<th>SVM</th>
<th>K- means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage of sensitivity</td>
<td>----</td>
<td>94.1</td>
<td>92.1</td>
<td>92.0</td>
</tr>
<tr>
<td>Percentage of specificity</td>
<td>----</td>
<td>94.3</td>
<td>94.2</td>
<td>93.5</td>
</tr>
<tr>
<td>Percentage of accuracy</td>
<td>93.0</td>
<td>94.2</td>
<td>94.1</td>
<td>93.4</td>
</tr>
</tbody>
</table>

Note: Canonical Correlation Analysis (CCA); Support Vector Machine (SVM)

Table 9- Various algorithms performance for MRI images for K=2

<table>
<thead>
<tr>
<th>Classification method</th>
<th>Ref</th>
<th>CCA</th>
<th>SVM</th>
<th>K- means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage of sensitivity</td>
<td>96.2</td>
<td>97.2</td>
<td>97.8</td>
<td>97.1</td>
</tr>
<tr>
<td>Percentage of specificity</td>
<td>95.7</td>
<td>97.6</td>
<td>97.5</td>
<td>96.3</td>
</tr>
<tr>
<td>Percentage of accuracy</td>
<td>96.7</td>
<td>97.5</td>
<td>97.6</td>
<td>96.4</td>
</tr>
</tbody>
</table>

Note: Canonical Correlation Analysis (CCA); Support Vector Machine (SVM)
Table 10- Various algorithms performance for MRI images for $K=3$

<table>
<thead>
<tr>
<th>Classification method</th>
<th>Ref</th>
<th>CCA</th>
<th>SVM</th>
<th>K-means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of sensitivity</td>
<td>96.2</td>
<td>96.9</td>
<td>97.8</td>
<td>98.1</td>
</tr>
<tr>
<td>Percentage of specificity</td>
<td>95.7</td>
<td>97.7</td>
<td>97.4</td>
<td>96.1</td>
</tr>
<tr>
<td>Percentage of accuracy</td>
<td>96.7</td>
<td>97.5</td>
<td>97.5</td>
<td>96.4</td>
</tr>
</tbody>
</table>

Note: Canonical Correlation Analysis (CCA); Support Vector Machine (SVM)