

# Tumor Detection using Brain MRI and Low-dimension Co-occurrence Feature Approach

Marta Mirkov<sup>1</sup>, Ana Gavrovska<sup>1</sup>

**Abstract:** Research in medical imaging focuses on methods useful in computer-aided diagnosis systems. In modern times, these systems often have automatic detection of regions of interest, and imaging technologies offer numerous advantages, like the possibility of developing reliable assisting algorithms. Magnetic Resonance Imaging (MRI) provides compelling features for brain tumor detection due to good soft tissue contrast and has important clinical value. To help clinicians in making diagnoses, current algorithms for processing and medical image classification may depend on intricate deep learning designs that demand large hardware resources and lengthy execution times. This is with no doubt helpful in understanding disease mechanisms and in labeling difficult instances for brain tumor identification. On the other hand, statistical low-dimension feature sets including co-occurrence-based ones could be useful in dealing with tumor detection avoiding possible complexity. In this paper, statistical approaches for feature extraction and reduction are analyzed for MRI brain tumor classification, and the evaluation of the results is presented on one of the publicly available brain tumor detection database commonly used for machine learning tasks. Bayes and  $k$ NN classifiers are applied for the analysis, as well as four distance metrics, and two methods for feature reduction. The results seem promising in developing a simple and less hardware-demanding procedure.

**Keywords:** Brain tumor detection, Region of Interest, Magnetic Resonance Imaging, Statistical moments, Feature extraction and reduction, Machine learning.

## 1 Introduction

In recent years, research in medical imaging has focused on the implementation of machine and deep learning algorithms for assisting in making proper diagnosing decisions related to various diseases. Some of them employ ionizing radiation intending to get adequate organ representation. Not surprisingly, systems that do not use this form of radiation are widely applied in medical decision-making, such as ultrasound scanning and magnetic resonance

---

<sup>1</sup>University of Belgrade, School of Electrical Engineering, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia; E-mails: [marta.mirkov@gmail.com](mailto:marta.mirkov@gmail.com); [anaga777@etf.bg.ac.rs](mailto:anaga777@etf.bg.ac.rs).

imaging (MRI) [1, 2]. In brain tumor detection, ultrasound imaging is at disadvantage compared to MRI due to the attenuation of these waves through the skull bones. On the other hand, MRI provides brain images with good soft-tissue contrast and offers more possibilities throughout measurements. In fact MRI enables an efficient characterization for identification of various disorders of the central nervous system including demyelinating diseases, cerebrovascular disease, dementia, epilepsy, infectious disease, Alzheimer's disease, etc. This is mainly due to the good contrast between gray and white matter [2]. Imaging is performed in milliseconds, and it enables a rich set of possibilities for developing a highly efficient algorithm that can be helpful in both functional and anatomical challenges.

A brain tumor is described as an abnormal growth of unwanted malignant cells that disrupts the operation of functioning cells in brain tissue. The lives of patients can be saved by early identification and prompt diagnosis of malignancies. Thus, mathematical and programming algorithms can be quite useful in order to efficiently reveal possible brain anomalies. The study of such approaches has a long history, and one of the major challenges is a realization of tools and algorithms that are efficient as assistance in diagnosing without dealing with unnecessary resources. To develop a straightforward system for brain tumor diagnosis, statistical feature extraction may reveal whether in some image cases low-complexity models can be used. Here, the application of Bayes and  $k$ NN classifiers is considered, as well as the application of four metrics, but with the examination of feature reduction for obtaining a low-dimension, i.e. reduced feature approach. Namely, PCA (Principal Component Analysis) and LDA (Linear Discriminant Analysis) methods are tested on statistically based techniques that can provide efficient and reliable results [3, 4].

The organization of the rest of this paper is as follows. In Section II, traditional methods for MRI-based brain tumor detection are considered, as well as some of the recent related work. The main steps in the experimental analysis are described in Section III, where further details for feature extraction, reduction, classifiers design, and performance evaluation are explained. Experimental results in this paper show further possibilities for obtaining low-dimension feature set. This is presented in Section IV. Finally, the conclusion including possibilities for future work can be found in Section V.

## **2 Feature Engineering and GLCM Features**

Tumors can be quickly and accurately identified, which can potentially save a patient's life. Brain abnormalities are often identified with great success using mathematical and technological approaches. As a result, related research focuses on machine learning and statistical feature extraction techniques for brain tumor identification. Machine learning in medical imaging is found beneficial, whereas

Tumor Detection using Brain MRI and Low-dimension Co-occurrence Feature Approach for segmentation and classification tasks statistical features are particularly valuable [5]. The texture is one of the most significant image and region of interest features for the appearance description, especially in the identification of brain tumors [6, 7]. Fundamental research questions include the differentiation of visual qualities, where statistically defined properties and spatial distribution of gray levels in the MRI image are crucial.

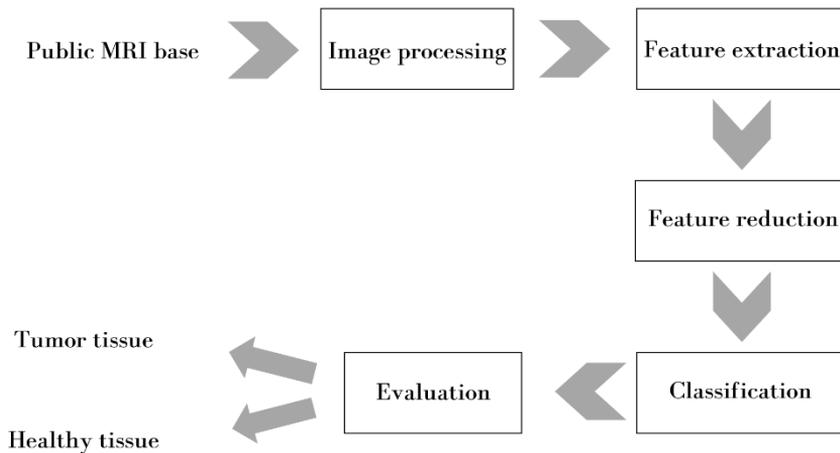
Co-occurrence matrices are one of the most widely used mathematical representation of visual texture. The spatial relationship between pixels is taken into consideration within the Gray Level Co-occurrence Matrix (GLCM) to analyze the texture using statistical methods. Energy, correlation, contrast, and homogeneity are four traditional properties that can be retrieved, and are used in [5 – 7]. These features are specifically extracted and form inputs for models like neuro-fuzzy ones, as well as for various machine and deep learning solutions [8 – 12]. Support vector machines are used as a classifier in [6], while deep convolutional neural networks are built in [11] and [12], using one of the publically available datasets for brain tumor identification [13]. Despite common features, other descriptors can be considered to deal with automatic brain tumor detection challenges. More parameters may produce improved results, but it is of importance to analyze common methods for feature number reduction. Handcrafted characteristics and feature engineering are still useful for getting reliable results in diagnostics.

### **3 Brain Tumor Detection and Feature Analysis**

The cancer tissue is supposed to stand out from the neighboring normal tissue in MRI images due to the contrast that MRI images provide. However, the choice of features for classification is still a matter of debate. Along with the common GLCM features, in this paper totally nine features are extracted for brain tumor classification. This is a relatively small number compared to the literature for this challenge [6, 8, 14]. On the other hand, a smaller feature vector dimension is crucial for algorithm execution, particularly when the evaluated dataset indicates that a higher level of complexity is not necessary. Therefore, two algorithms for feature reduction are tested. The block diagram of the proposed analysis is shown in Fig. 1.

The experimental analysis carried out in this paper includes five main steps:

1. image (pre-)processing,
2. tumor region of interest segmentation,
3. feature extraction,
4. classification, and
5. evaluation of the dataset.



**Fig. 1** – Main steps for brain tumor detection evaluation.

### 3.1 Image (pre-)processing and tumor region of interest segmentation

Segmentation of regions of interest corresponding to brain tumors involves multiple steps. An input MRI image is preprocessed using a high-pass filter and intensities are adjusted to capture relevant pixel candidates. The binarization is then used to identify similar regions. In the final step, solidity for tumor segmentation is found.

A high-frequency filter exposes sudden changes in an image [15]. The employed filter is a fifth-order Gaussian High-Pass (HP) filter with a cutoff frequency of 55 Hz. The grayscale range of the input image is reduced, and the contrast is improved by saturating the lowest two percent and the top one percent of all pixel intensity values, which reveals possible tumor locations.

The image histogram displays a bimodal distribution with a deep and sharp valley between two pointed peaks because the image contrast has been modified. This allows the application of Otsu's approach for automatic thresholding [15]. After the thresholding, this binary image contains a variety of white patches, some of which may be tumor tissue. If two neighboring pixels are connected through eight connectivity, they are considered a part of a single region of interest.

Pre-experimental analysis showed that solidity may be an interesting property [1]. Thus, solidity is computed, as well for each region that has been labeled. The region's total solidity is quantified as the convex hull area divided by the image area. The convex hull area grows, and the computed solidity falls as the object form moves away from a closed circle. Tumor areas are more likely to be present in images with high solidity [7, 16]. It is assumed that by contrasting the calculated solidity of the labeled image with a higher value of solidity (closer

Tumor Detection using Brain MRI and Low-dimension Co-occurrence Feature Approach to 1), a tumor can be properly identified if it exists. If the solidity is greater than 0.6 in this instance, the tumor is detected. The region of interest may represent a tumor area where a relatively large number of connected white pixels exist within the binary image. Still, small white regions can be detected despite the fact there is no tumor tissue in the MRI image. A solution is to exploit the properties of the segmented region(s) of interest to make a differentiation between healthy and tumor tissue. This difficulty may be resolved by computing features that give sufficient region area and region shape information.

### 3.2 Feature extraction

The region of interest segmentation is followed by feature extraction. Here, exclusively statistical descriptors are analyzed for binary classification. The main reason for the integration of statistical features, like the GLCM ones, reflects in the fast and easy-to-use way of quantizing region characteristics. Except for GLCM-based features, and statistical moments, a few handcrafted features are tested to further stimulate the practicability of the feature vector [17].

Gray Level Co-occurrence Matrix (GLCM) is used for texture analysis [4]. We consider two pixels at a time, called the reference and the neighbor pixel and characterize a specific spatial relationship between them. Each entry of the GLCM holds the count of the number of times that pair of intensities appears in the image with the defined spatial relationship. The frequency at which a pixel with intensity value  $i$  and a pixel with value  $j$  are horizontally adjacent determines the GLCM. Offset is not employed to specify the horizontal spatial connections between pixels. The GLCM size depends on how many gray levels exist. In Fig. 2 example of calculating the GLCM values is shown. For input image with intensity levels from zero to three, GLCM will be  $4 \times 4$ . For feature extraction, the dimension of GLCM is  $8 \times 8$  pixels since the number of intensity values is decreased by scaling to eight levels.

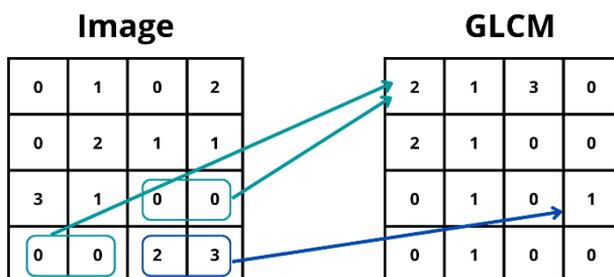


Fig. 2 – Example of calculating the GLCM values.

Energy, correlation coefficient, contrast descriptor, and homogeneity descriptor are the usual statistical handcrafted features that are taken from the calculated GLCM. In **Table 1** four features are briefly summarized. Energy

estimates the sum of squared elements from GLCM and represents the first feature. The second feature is a correlation, which is a well-known approach for describing the dependency between pixels. The contrast value stands here as the third feature, and it estimates the local variations. The fourth feature represents homogeneity, which here describes the closeness of the distributed pixels. For each detected region, the abovementioned features are calculated.

The only feature that applies comparison, meaning that it is reference-based, is the correlation feature. Joint probabilities are applied, and linear dependency between neighbouring pixels is found. Note that the means and standard deviations that correspond to  $p_x$  and  $p_y$  are calculated, where  $p_x(i)$  represents the  $i$ -th entry in the marginal-probability matrix obtained by summing the rows of  $GLCM(i, j)$  and  $p_y(i)$  is the  $i$ -th entry in the marginal-probability matrix obtained by summing the columns of  $GLCM(i, j)$ .

**Table 1**  
The usual feature description\*.

| No. | Feature description and expression no. | Expressions (1) – (4)   |
|-----|--|---|
| 1   | Energy (1)                             | $E = \sum_{i,j=0}^{N-1} p^2(i, j)$  |
| 2   | Correlation (2) **                     | $Corr = \sum_{i,j=0}^{N-1} \frac{(i - \mu_x)(j - \mu_y)p(i, j)}{\sigma_x \sigma_y}$ |
| 3   | Contrast (3)                           | $C = \sum_{i,j=0}^{N-1}  i - j ^2 p(i, j)$  |
| 4   | Homogeneity (4)                        | $H = \sum_{i,j=0}^{N-1} \frac{p(i, j)}{1 +  i - j }$                                |

\*  $N$  represents the number of pixels;  $i$  and  $j$  the location of pixel;

$p(i, j)$  the pixel intensity at the location  $(i, j)$ .

\*\* Values  $\mu_x, \mu_y$ , are the means and  $\sigma_x$  and  $\sigma_y$  are the standard deviations of  $p_x$  and  $p_y$ , respectively.

Besides the GLCM features, statistical moments also give information about intensity distribution. Hence, they are commonly used as features in medical image classification [17, 18]. Let random variable  $I$  represent the gray levels of an image region. The first-order histogram  $P(I)$  is defined as:

$$P(I) = \frac{\text{number of pixels with gray level } I}{\text{total number of pixels in the region}}. \tag{5}$$

Based on the definition of  $P(I)$ , the mean  $m_1$  and central moments  $\mu_k$  of  $I$  are given by [16]:

$$m_1 = E[I^1] = \sum_{I=0}^{N_g-1} I^1 P(I), \quad (6)$$

$$\mu_k = E[(I - E[I])^k] = \sum_{I=0}^{N_g-1} (I - m_1)^k P(I), \quad (7)$$

where  $N_g$  is the number of gray levels, and  $k \in \{2, 3, 4\}$ .

The most frequently applied central moments that are inexpensive to computation are: variance, skewness, and kurtosis calculated by  $\mu_2$ ,  $\mu_3$ , and  $\mu_4$  respectively using (7) [17]. The variance corresponds to histogram width and measures the deviation of gray levels from the mean. Skewness is a measure of the histogram asymmetry degree around the mean. With the fourth statistical moment in mind, the kurtosis measures the histogram sharpness. One way to deal with tumor tissue is through higher skewness since it is reasonable to assume that tumor tissue is rich in high-intensity values compared to healthy tissue. It is also a reliable feature for image classification [15]. In this paper, mean, variance, and skewness are found for pre-processed images meaning images with adjusted contrast.

Two additional features are extracted for the experimental analysis, and both of them are calculated directly from the binary image. The area of the region(s) of interest can point out the higher number of white pixels expected in the number of tumor tissue images. Since the solidity parameter has been found useful in the final step of a region of interest segmentation, it is added as a feature for the classification. The calculation is easy to perform as the area is divided by its convex hull areas.

Using exclusively GLCM features or just first-order statistical moments for a region of interest may not be sufficient as input in the classification and for obtaining high accuracy results. Nine features are tested to exploit further feature reduction based on two common methods.

### 3.3 Feature reduction

Feature extraction may produce a relatively large number of features meaning that challenges in detection are presented with plentiful descriptions. Therefore, it is commonly expected that the classifier is able to make an improved decision while receiving more information leading to classification accuracy increases with a higher number of features. However, a large number of features can also lead to the creation of the dimensionality curse misinterpretations and higher complexity of the algorithm [18, 19]. Hence, there are dimension reduction methods that are developed so that the most promising features, which carry the most information, can be selected among others, mostly redundant, for the classification task.

Typically applied statistical methods of dimensional reduction are: Principle Component Analysis (PCA) method, vector quantization approach, topographic measurements based method, the dimension reduction methods based on the scattering matrix or entropy, etc. [18, 19].

The basic idea behind the PCA algorithm is that some  $n$ -th dimensional vector  $\mathbf{X}$  can be mapped into a new random vector  $\mathbf{Z}$  by the linear transformation as shown in equation (8):

$$\mathbf{Z} = \mathbf{A}^T \mathbf{X} . \quad (8)$$

The matrix  $\mathbf{A}$  determines the dimension of the new vector  $\mathbf{Z}$ , and it should be chosen so that the loss of information caused by the dimension reduction is minimal. The criterion function aims optimization in order to minimize the mean square errors of approximation:

$$\overline{\varepsilon_{opt}^2} = \sum_{i=m+1}^n \lambda_i , \quad (9)$$

where  $m$  is the desired dimension,  $n$  is the current dimension, and  $\lambda_i$  is the  $i$ -th eigenvalue. This criterion is equal to the sum of the eigenvalues along those coordinates that were omitted during the dimension reductions. As it is in the interest of the expression (9) to get the minimum value, it becomes clear that during the reduction the coordinates of the vectors whose eigenvalues are of little value should be ignored.

Linear Discriminant Analysis (LDA) method, unlike the previous method, considers the separability of classes, if it exists [18]. In the case of two classes, the following matrices are formed:

$$\mathbf{S}_w = \Sigma_1 \mathbf{P}_1 + \Sigma_2 \mathbf{P}_2 , \quad (10)$$

$$\mathbf{S}_b = \mathbf{P}_1 (\mathbf{M}_1 - \mathbf{M}_0) (\mathbf{M}_1 - \mathbf{M}_0)^T + \mathbf{P}_2 (\mathbf{M}_2 - \mathbf{M}_0) (\mathbf{M}_2 - \mathbf{M}_0)^T , \quad (11)$$

$$\mathbf{S}_m = \mathbf{S}_w + \mathbf{S}_b , \quad (12)$$

where  $P_i$  represents the probability of class  $i$  occurrence,  $\mathbf{M}_i$  is the mathematical expectation of class  $i$ ,  $\mathbf{M}_0$  is the combined mathematical expectation for all classes and  $\Sigma_i$  denotes the covariance matrix of class  $i$ . Matrix  $\mathbf{S}_w$  represents the within-class scatter matrix,  $\mathbf{S}_b$  the inter-class scatter matrix, and  $\mathbf{S}_m$  the mixed one matrix. The challenge is the same as with the PCA method, where it is necessary to determine the matrix  $\mathbf{A}$  from equation (8), thus minimizing the selected criterion. The most frequently applied criterion is  $J = \ln |\mathbf{S}_b^{-1} \mathbf{S}_w|$ . At this point the transformation matrix  $\mathbf{A}$ , which minimizes criterion  $J$ , can be written in the following form:

$$\mathbf{A} = [\Psi_1 \quad \Psi_2 \quad \dots \quad \Psi_m] , \quad (13)$$

Tumor Detection using Brain MRI and Low-dimension Co-occurrence Feature Approach where  $\Psi_1, \Psi_2, \dots, \Psi_m$  are the eigenvectors of the matrix  $S_b^{-1} S_w$ , which correspond to the largest  $m$  eigenvalues. By choosing  $m$ , a new reduced dimension of the vector is selected. Each of these two methods (PCA and LDA) has its own advantage. PCA takes the vectors with the highest variances, while LDA maintains the separability of the classes.

### 3.4 Classifiers design

For Bayes classifier design, the posterior probabilities  $q_i(X)$ , which represent the conditional likelihood that sample  $X$  belongs to class  $i$  if its precise realization is known, need to be defined. The Bayes theorem can be used to determine these probabilities if prior probabilities of class occurrence and posterior density probability functions of measured vectors  $q_i(X)$  are known [17 – 19]. Feature vectors are calculated for both images of tumor tissue and images of healthy tissue, representing the first ( $\omega_1$ ) and the second class ( $\omega_2$ ), respectively. The following fundamental decision rule for sample  $X$  can be utilized based on conditional probabilities:

$$q_1(x) > q_2(x) \Rightarrow X \in \omega_1, \quad (14)$$

$$q_2(x) > q_1(x) \Rightarrow X \in \omega_2. \quad (15)$$

Although the probability density functions of the classes are unknown, it can be assumed that if there is a higher sample number, the probability density functions of the classes can be assumed to be Gaussian using the central limit theorem [20]:

$$f(x) = \frac{1}{|\Sigma|^{1/2} (2\pi)^{n/2}} e^{-\frac{1}{2}(x-M)^T \Sigma^{-1}(x-M)}, \quad (16)$$

where  $n$  is the dimension of vector  $X$ ,  $M$  is its mathematical expectation, and  $\Sigma$  is the feature vector's covariance matrix. These values are taken from the training set for both classes.

The non-parametric classification is useful in cases where classification information based on hypothesis testing is lacking.  $K$ -nearest neighbours, often known as  $k$ NN, is one of the most useful techniques. The method assigns appropriate group to the observation point based on how its neighbours are labeled. The number of nearby samples included in the categorization is indicated by the parameter  $k$  in the  $k$ NN algorithm [22, 23]. A new circle sample that needs to be categorized can be seen in Fig. 3. When  $k$  takes, for example value 3, the whole circle in the illustration reflects the situation where one square and two stars are found as its neighbours. The circle sample is included in the same class as the stars since there are more stars than there are squares. However, if four is taken for  $k$ , the circle will be classified into the squares class because there are more of them in the surroundings. In conclusion,  $k$  is an essential parameter [23].

Also, the success of the classification depends on methods used for defining what the nearest neighbours are. Some of the methods that are going to be considered in this paper are: Euclidean, Chebyshev, Mahalanobis distance, and cosine similarity [24].



**Fig. 3** – Graphical representation of *k*NN method, where circle sample needs to be assigned to rectangle or star class.

### 3.5 Evaluation of results

The whole dataset is consisted from 253 images, where 98 images represent healthy brain and 155 images with tumor tissue [13]. Difference of number of samples in each class is important to considerate in aim to choose right evaluation metric. These difference shows that used database is unbalanced and requires metrics that are used for unbalanced dataset, which will be represented in next paragraph.

Both tested classifiers, Bayes and *k*NN, are evaluated on the test set, which is made from 30% of the whole dataset selected randomly, while the rest was used for the training set. The whole algorithm was reapplied five times to ensure that randomly chosen sets are the best ones for classification, and the results represent averaged values.

The performance is evaluated using metrics extracted from a confusion matrix. Two columns and two rows constitute the confusion matrix in this binary categorization. The instances in a real class can be represented in each row of the matrix, while the examples in a predicted class as can be shown in each column. Images containing tumor tissue are marked as “positive”, while those containing healthy tissues are marked as “negative.” For performance evaluation, true positive rate (TPR), true negative rate (TNR), and balanced accuracy (BACC) are calculated as in (17) – (19), respectively. Sensitivity is found as:

$$\text{Sensitivity} = \frac{TP}{P}, \quad (17)$$

Tumor Detection using Brain MRI and Low-dimension Co-occurrence Feature Approach where  $TP$  represents true positive, the number of samples that has been positive and are detected as positive, and  $P$  represents the whole set of positive samples. Similarly, specificity or  $TNR$  is calculated as:

$$Specificity = \frac{TN}{N}, \quad (18)$$

where  $TN$  represents true negative meaning the number of samples that has been negative and are detected as negative, and  $N$  represents the whole set of negative samples. Best parameter for evaluation is balanced accuracy because the dataset isn't balanced. Finally, the balanced accuracy is found as:

$$BACC = \frac{TPR + TNR}{2}. \quad (19)$$

Another useful way to evaluate the method performance is by ROC (Receiver Operating Characteristic) curve. The parameters associated with the statistical classifiers can be varied in order to change the TP and FP rates [18]. Each set of parameter values can result in a different (TP, FP) pair, i.e. working point. It is possible to trade a lower (higher) FP rate for a higher (lower) TP detection rate by choice of the corresponding parameter values. Area under the curve (AUC) is a metric for evaluation of the classifier performance. When a perfect classifier has a TP rate one and FP rate zero, this results in having AUC of value one. A random guess results in AUC of value 0.5.

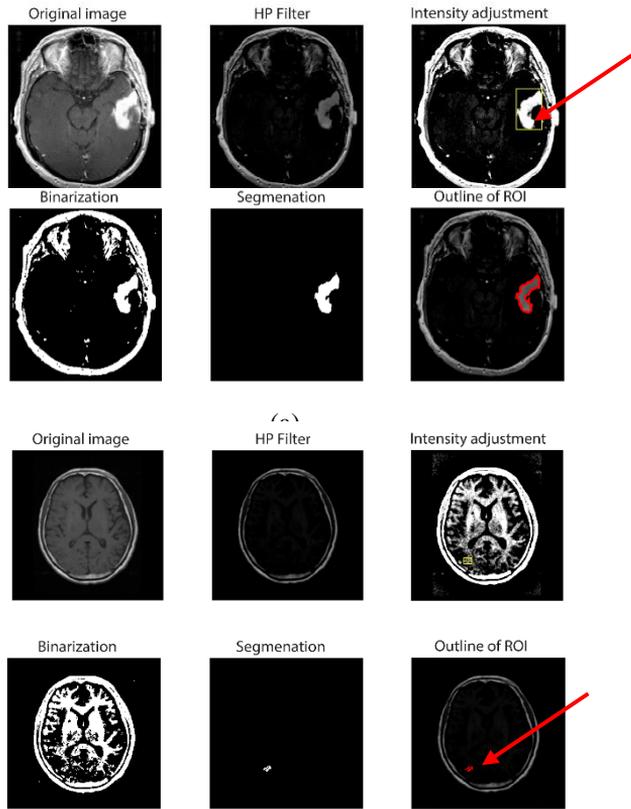
## 4 Experimental Results

Results of the proposed tumor segmentation steps are shown in Fig. 4. Cancer tissue is detected, and in a normal tissue image only a few white pixels are found as possible candidates for tumor.

The proposed tumor segmentation steps and feature extraction are used for the classification based on Bayes and  $kNN$  classifiers. Moreover, the influence of feature vector dimension on results is analyzed. Feature reduction is performed on feature vectors. It is obtained that less than four features is not enough for the classification task. The effects of distance type and different values of  $k$  (1 to 40) on accuracy are also analyzed for  $kNN$  methods.

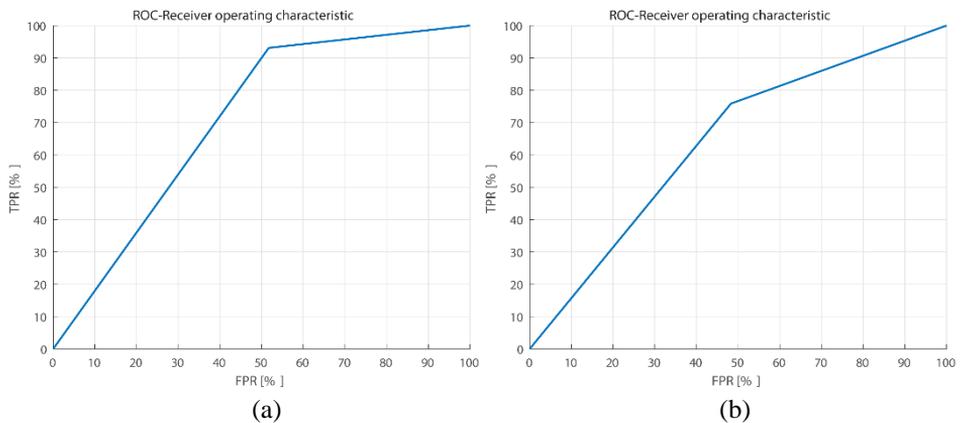
Here, dimension reduction serves for checking whether smaller feature dimensions can be used without affecting the results. Two methods, LDA and PCA, are tested. Both methods are considered for dimensionality reduction to four and six features, with the goal of accuracy analysis. Overall results can be seen in **Table 2**.

In Figs. 5 and 6, ROC for both classifiers for four feature approach is shown. For both classifiers, higher TPR has achieved for LDA feature reduction method, which also can be seen from **Table 2**.



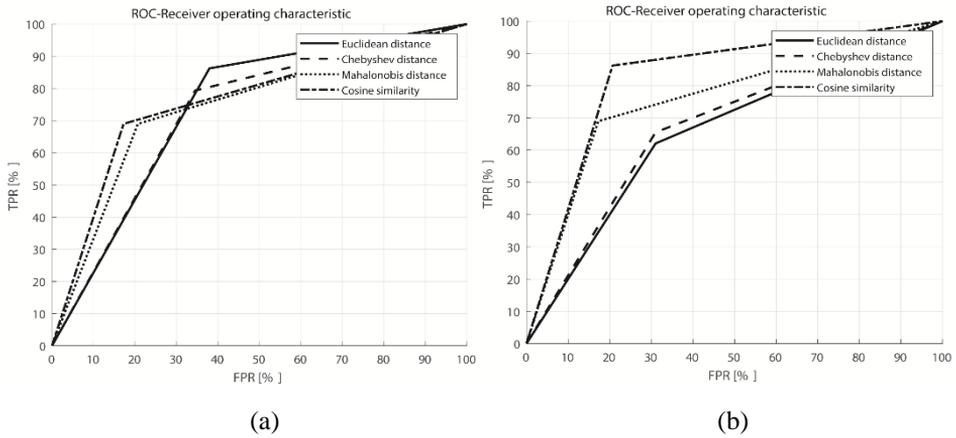
(b)

**Fig. 4** – Results of segmentation for (a) cancer tissue and (b) normal tissue.



**Fig. 5** – ROC curve for the Bayes classifier using four features obtained by (a) LDA method and (b) PCA method.

## Tumor Detection using Brain MRI and Low-dimension Co-occurrence Feature Approach



**Fig. 6** – ROC curves for the  $k$ NN classification using four features obtained by (a) LDA method and (b) PCA method.

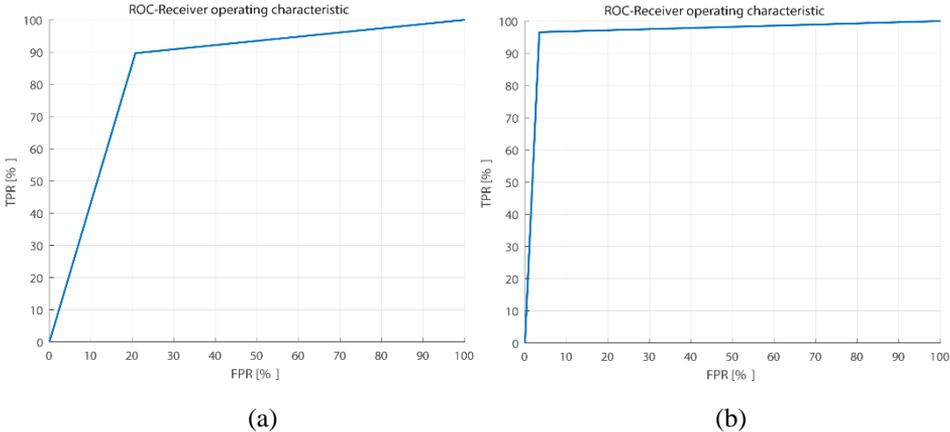
**Table 2**  
Overall test results.

| Test no. | Reduction method (vector dimension) | Classifier | BACC [%] | TPR [%] | TNR [%] |
|----------|-------------------------------------|------------|----------|---------|---------|
| 1        | LDA (4)                             | Bayes      | 81       | 84      | 78      |
| 2        | LDA (4)                             | $k$ NN     | 80       | 82      | 78      |
| 3        | LDA (6)                             | Bayes      | 98.3     | 96.6    | 100     |
| 4        | LDA (6)                             | $k$ NN     | 98       | 100     | 96      |
| 6        | PCA (4)                             | Bayes      | 72       | 74      | 70      |
| 7        | PCA (4)                             | $k$ NN     | 82.3     | 93.1    | 72.4    |
| 8        | PCA (6)                             | Bayes      | 96.5     | 93.5    | 100     |
| 9        | PCA (6)                             | $k$ NN     | 96.5     | 96.5    | 96.5    |

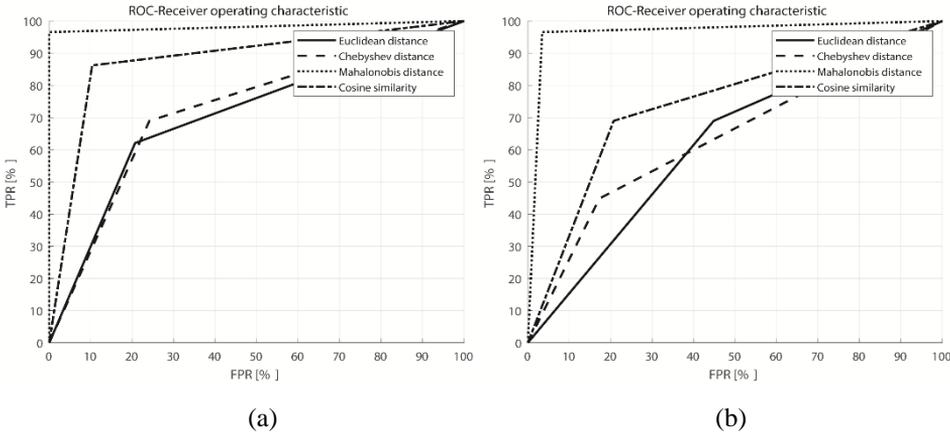
In Fig. 7 and Fig. 8, ROC for both classifiers for six feature approach is shown. For both classifiers, higher TPR has achieved for LDA feature reduction method, which also can be seen from **Table 2**. Using six features increased accuracy and decreased false positive rate.

LDA method gives better results due to using more information for reduction than PCA method. The  $k$  value for this challenge is analyzed for feature dimensions of four and six, obtained by LDA method. For feature vector consisted of six values the best results are obtained when  $k$  equals 11, which

matches with literature proposition for  $k$  (choosing  $k \sim \sqrt{N}$  where  $N$  represents the number of samples) and for four features the  $k$  has value 2. Mahalanobis distance, as well as LDA method use covariance matrix of feature vectors so it is expected for those algorithms to give reliable results. Mahalanobis distance gave the best result in both cases, which can be seen in Fig. 9.

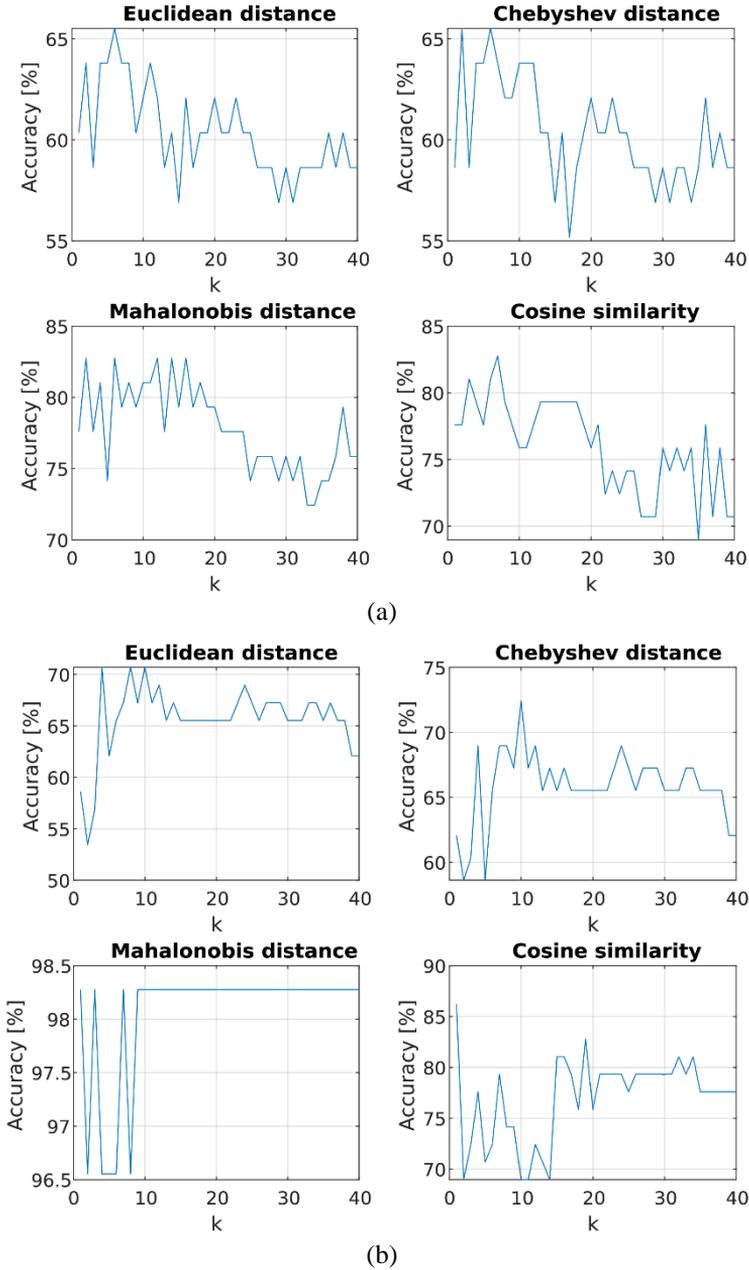


**Fig. 7** – ROC curves for the Bayes classification using six features obtained by (a) LDA method and (b) PCA method.



**Fig. 8** – ROC curves for the  $k$ NN classification using six features obtained by (a) LDA method and (b) PCA method with noticeable advantage of Mahalanobis distance.

Tumor Detection using Brain MRI and Low-dimension Co-occurrence Feature Approach



**Fig. 9** – Accuracy versus parameter  $k$  for four types of distances for (a) six and (b) four features (LDA method).

## 5 Conclusion

The analyzed classification methods give satisfying results compared to the methods based on neural networks or fuzzy logic [9 – 12]. With the aim of obtaining such results, several techniques for image processing, feature selection and feature reduction are applied, as well as two classifiers. The results primarily depend on the training and the test set, so the results shown are averaged over several choices. In this paper the  $k$ NN method with Mahalanobis distance metric is selected as a method with high accuracy.

Methods based on neural networks or fuzzy logic can give improved results, however, they require a lot of hardware power and are often time-consuming in the cases where it is not needed. In this sense, these classical methods have an advantage. In practice, these methods can be used as parts of an automated system to help medical diagnostics.

Future work will be oriented through using additional types of features with other datasets to distinguish cases where more complex approaches are inevitable without wasting resources unnecessarily. Testing the whole algorithm with larger datasets with different tumor types is expected as the next step. This will require deeper analysis of different tumor type images. Detailed feature analysis for this challenge is crucial for developing algorithms similar to the proposed.

## 6 References

- [1] M. Mirkov, A. Gavrovska: Application of Bayes and  $k$ NN Classifiers in Tumor Detection from Brain MRI Images, Proceedings of the IX International Conference on Electrical, Electronic and Computing Engineering (IcETRAN), Novi Pazar, Serbia, June 2022, pp. 268 – 272.
- [2] S. Webb: The Physics of Medical Imaging, Taylor and Francis Group, New York, London, 1988.
- [3] R. A. Sadek: An Improved MRI Segmentation for Atrophy Assessment, International Journal of Computer Science Issues, Vol. 9, No. 3, May 2012, pp. 569 – 574.
- [4] P. K. Bhagat, P. Choudhary, Kh. M. Singh: A Comparative Study for Brain Tumor Detection in MRI Images Using Texture Features, Ch. 13, Sensors for Health Monitoring: Advances in Ubiquitous Sensing Applications for Healthcare, Vol. 5, pp. 259 – 287, 1<sup>st</sup> Edition, Elsevier, London, 2019.
- [5] J. Jagtap, N. Patil, C. Kala, K. Pandey, A. Agarwa, A. Pradhan: Statistical Characterization of Tissue Images for Detection and Classification of Cervical Precancers, arXiv:1112.4298 [physics.med-ph], December 2011, pp. 1 – 25.
- [6] S. A. Medjahed: A Comparative Study of Feature Extraction Methods in Images Classification, International Journal of Image, Graphics and Signal Processing, Vol. 7, No. 3, February 2015, pp. 16 – 23.
- [7] K. K. Kumar, M. D. Thiagarajan, S. Maheswaran: An Efficient Method for Brain Tumor Detection Using Texture Features and SVM Classifier in MR Images, Asian Pacific Journal of Cancer Prevention, Vol. 19, No. 10, October 2018, pp. 2789 – 2794.

## Tumor Detection using Brain MRI and Low-dimension Co-occurrence Feature Approach

- [8] D. Mery, D. Filbert: Classification of Potential Defects in Automated Inspection of Aluminum Castings Using Statistical Pattern Recognition, Proceedings of the 8<sup>th</sup> European Conference on Non-Destructive Testing (ECNDT), Barcelona, Spain, June 2002, pp.1–13.
- [9] T. M. Hsieh, Y.- M. Liu, C.- C. Liao, F. Xiao, I.- J. Chiang, J.- M. Wong: Automatic Segmentation of Meningioma from Non-Contrasted Brain MRI Integrating Fuzzy Clustering and Region Growing, BMC Medical Informatics and Decision Making, Vol. 11, 2011, p. 54.
- [10] K. Sharma, A. Kaur, S. Gujral: Brain Tumor Detection based on Machine Learning Algorithms, International Journal of Computer Applications, Vol. 103, No. 1, October 2014, pp. 7–11.
- [11] R. Ranjbarzadeh, A. B. Kasgari, S. J. Ghouschi, S. Anari, M. Naseri, M. Bendechache: Brain Tumor Segmentation based on Deep Learning and an Attention Mechanism Using MRI Multi-Modalities Brain Images, Scientific Reports, Vol. 11, No. 1, May 2021, p. 10930.
- [12] A. Çinar, M. Yildirim: Detection of Tumors on Brain MRI Images Using the Hybrid Convolutional Neural Network Architecture, Medical Hypotheses, Vol. 139, June 2020, p. 109684.
- [13] N. Chakrabarty: Brain MRI Images for Brain Tumor Detection Dataset, Available at: <https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection>
- [14] N. Aggarwal, R. K. Agrawal: First and Second Order Statistics Features for Classification of Magnetic Resonance Brain Images, Journal of Signal and Information Processing, Vol. 3, No. 2, May 2012, pp. 146–153.
- [15] M. Popović: Digital Image Processing, Akademska misao, Belgrade, 2006. (In Serbian).
- [16] Sanallah, M. A. Javid, S. A. Buzdar: A Novel Computer Aided Diagnostic System for Quantification of Metabolites in Brain Cancer, Biomedical Signal Processing and Control, Vol. 66, April 2021, p. 102401.
- [17] H. J. Baek, H. S. Kim, N. Kim, Y. J. Choi, Y. J. Kim: Percent Change of Perfusion Skewness and Kurtosis: A Potential Imaging Biomarker for Early Treatment Response in Patients with Newly Diagnosed Glioblastomas, Radiology, Vol. 264, No. 3, September 2012, pp. 834–843.
- [18] A. Meyer-Baese, V. J. Schmid: Pattern Recognition and Signal Analysis in Medical Imaging, 2<sup>nd</sup> Edition, Elsevier, Amsterdam, Boston, 2014.
- [19] K. Fukunaga: Introduction to Statistical Pattern Recognition, 2<sup>nd</sup> Edition, Academic Press, San Diego, San Francisco, New York, Boston, 1990.
- [20] B. Efron: Bayes' Theorem in the 21<sup>st</sup> Century, Science, Vol. 340, June 2013, pp. 1177–1178.
- [21] S. G. Kwak, J. H. Kim: Central Limit Theorem: The Cornerstone of Modern Statistics, Korean Journal of Anesthesiology, Vol. 70, No. 2, April 2017, pp. 144–156.
- [22] G. Guo, H. Wang, D. Bell, Y. Bi, K. Greer: *k*NN Model-Based Approach in Classification, Proceedings of the OTM Confederated International Conferences – On the Move to Meaningful Internet Systems, Catania, Italy, November 2003, pp. 986–996.
- [23] P. Nair, I. Kashyap: Classification of Medical Image Data Using K Nearest Neighbor and Finding the Optimal K Value, International Journal of Scientific and Technology Research, Vol. 9, No. 4, April 2020, pp. 221–226.
- [24] R. Ehsani, F. Drabløs: Robust Distance Measures for *k*NN Classification of Cancer Data, Cancer Informatics, Vol. 19, 2020, pp. 1–9.