Pattern Recognition Letters 000 (2017) 1–10

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

A distinctive approach in brain tumor detection and classification using MRI

Javeria Amin[a], Muhammad Sharif[a], Mussarat Yasmin[a,*], Steven Lawrence Fernandes[b]

[a] COMSATS Institute of Information Technology, Wah 47040, Pakistan
[b] Department of Electronics and Communication Engineering, Sahyadri College of Engineering & Management, Mangalore, Karnataka, India

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Cells
Tumors
Segmentation
Lesion
Tissues

ABSTRACT

A very exigent task for radiologists is early brain tumor detection. Brain tumor raises very fast, its average size doubles in just twenty-five days. If not treated properly, the survival rate of the patient is normally not more than half a year. It can rapidly lead to death. For this reason, an automatic system is required for brain tumor detection at an early stage. In this paper, an automated method is proposed to easily differentiate between cancerous and non-cancerous Magnetic Resonance Imaging (MRI) of the brain. Different techniques have been applied for the segmentation of candidate lesion. Then a features set is chosen for every applicant lesion using shape, texture, and intensity. At that point, Support Vector Machine (SVM) classifier is applied with different cross validations on the features set to compare the precision of proposed framework. The proposed method is validated on three benchmark datasets such as Harvard, RIDER and Local. The method achieved average 97.1% accuracy, 0.98 area under curve, 91.9% sensitivity and 98.0% specificity. It can be used to identify the tumor more accurately in less processing time as compared to existing methods.

1. Introduction

Brain tumor begins due to abnormal development of cells that proliferate in an uncontrolled way. Tumors originate from brain cells around the membranes of the brain (meninges), glands or nerves. Tumors can specifically wreck brain cells. They can harm cells via providing more pressure inside the skull [1]. Brain tumors are shown in Fig. 1.

Malignant tumors are most dangerous forms of tumor causing fourteen thousand deaths every year. Despite broad efforts in research over numerous decades, the middle Overall Survival (OS) remains still at only fifteen months for the malignant glioma, Glioblastoma Multiforme (GBM) [2]. Due to severity level, the tumor has been divided into different grades. In grade 1 least dangerous tumor are normally related to prolonged survival. They grow gradually and practically have a usual appearance when seen through a microscope. Treatment via surgery might be successful for this type of tumor grade. Pilocytic astrocytoma, ganglioglioma, and gangliocytoma are cases of grade 1 brain tumor. Grade 2 tumor grows slowly and looks anomalous under a microscopic instrument. Few spread into neighboring tissues and repeat, sometimes as grade high tumor [3]. A grade 3 tumor is malignant but generally, there is not a major contrast between grade 2 and grade 3 tumors. This tumor has a tendency to frequently repeat as grade 4. Grade 4 is the maximum malignant tumor. It duplicates quickly having a strange appearance when seen in the microscopic instrument and effectively grows into the neighboring tissues of the brain resulting in the appearance of new vessels. These tumor cells have zones of dead cells in their middles. GBM is an example of grade 4 tumor [4].

Image processing methods are mostly utilized for tumor detection. One of them is segmentation of images. The target of picture segmentation is to segment a picture into homogeneous districts hence finding shapes of the region [5–6]. Magnetic Resonance Imaging (MRI) or computed Tomography (CT) scan are used to examine brain life structure. MRI scan is more effective when contrasted with CT scan. It does not use radiation [7]. Tumor comprises of various biologic tissues; only single kind of MRI cannot give whole information related to anomalous tissues. Joining distinctive complementary information can upgrade the segmented region of tumors. Features of MR images utilized for segmentation comprise of weighted three pictures (T2, T1, and Proton Density (PD)) for every slice axial. The segmentation strategies have been entirely effective especially in the improvement stages of infected tissues [8–10].

Main contributions of this research include the following:

- An automatic method is proposed to detect and classify brain tumor at image and lesion levels. Proposed methodology con-

* Corresponding author.
E-mail address: mussarat@ciitwah.edu.pk (M. Yasmin).

https://doi.org/10.1016/j.patrec.2017.10.036
0167-8655/

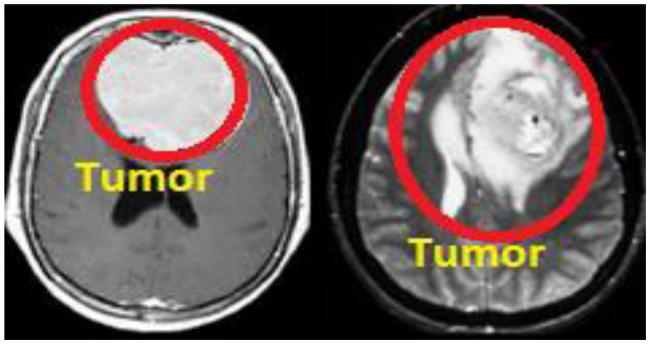

Fig. 1. Brain tumor.

sists of several steps, for example, preprocessing, extraction of features and classification.

- In preprocessing step, different methods are applied for the segmentation of candidate lesions. Then texture, shape and intensity based features set are applied for each candidate lesion. The texture features are good descriptors of the tumor along with shape and intensity features because the combination of texture, shape, and intensity of a lesion has better discrimination information.
- Finally, a number of tests are performed on the selected features set to choose a more appropriate classification strategy. For this purpose, geometrical family Support Vector Machine (SVM) is applied with different cross validations. Three variants of SVM are tested i.e., Linear, Gaussian and Cubic kernel functions. The method is tested on Local, Harvard and RIDER datasets based on the performance parameters such as Area Under the Curve (AUC) and Accuracy (ACC).

The article organization is as follows. Section II describes related work for tumor detection and section III mentions complete proposed method steps. Section IV discusses results after the performance evaluation of suggested method.

2. Related work

Medicinal image segmentation is considered as an imperative examination theme. Many analysts have proposed different systems and image segmentation methods for brain tumor detection; some are described in this section.

Morphological operators are used for tumor segmentation [11]. Non-negative Matrix Factorization (NMF) is utilized for brain tumor detection [12]. The neuro-fuzzy method is applied for the detection of various tissues such as White Matter (WM), Gray Matter (GM) and Cerebrospinal Fluid (CSF) [13]. Piece Wise-Triangular-Prism-Surface-Area (PTPSA) method is used for tumor detection and automated Bayesian regularization with feed forward multilayer neural network is utilized for the discrimination between tumors and non-tumor regions [14]. Fuzzy C-Means (FCM) clustering method is used for tumor detection [15]. K-means clustering method is applied for tumor diagnosis [16]. The suggested method comprises of three stages such as K-means clustering, guided grid based local standard deviation of coarse and fine grain localization. Extraction of brain tumor region is divided into two fragments. One fragment comprises of normal cells of the brain containing GM, CSF, and WM. The second fragment comprises of tumor cells. Segmentation method is used to isolate the region of interest. Image fusion method gives good results in the fusion of many images [17]. Spatial Fuzzy C-means (SFCM) clustering technique is utilized to examine the images of Positron Emission Tomography (PET) datasets [18]. Fuzzy clustering with FCM method is applied for tumor detection using MRI [19]. Thresholding, fast Fourier transforms and advanced morphological method are used for the detection of the tumor. The technique is tested on T2 weighted MRI modality [20]. Histogram-based Gravitational Optimization Algorithm (HGOA) is utilized to detect brain lesion. The method achieved an accuracy of 88.1% and 91.5% [23]. FCM segmentation method gives better results for the segmentation of lesion region [24]. K-means clustering and FCM methods are used for the segmentation of tumor tissues. According to the performance evaluation, it is observed that K-means performed superior as compared to FCM method [25]. Gabor wavelet features are tested on several classifiers for discrimination between the healthy and tumor tissues [26]. Gray Level Co-Occurrence Matrix (GLCM) with KSVM has been used for brain tumor detection [27].

In this article, an automatic technique is proposed for the segmentation and classification of brain pathology at both image and lesion levels.

At image level, the Gaussian filter is used for image smoothing and the optimal value of threshold and different morphological operations are utilized for discrimination between the healthy and tumor images.

At lesion level, an unsupervised K-means clustering approach is applied for classification between the grade 1, 2 also called benign (low-grade tumor) and grade 2, 3 malignant (high-grade tumor). Then shape, texture and intensity features are used for classification. In the classification stage, three kernels of the geometrical family are tested to compare the accuracy of proposed method.

3. Proposed methodology

An automated system is presented for the detection of brain tumor at the lesion and image levels. Suggested system is tested on MRI. Tumor detection in MRI is more efficient because of its low radiation, high contrast, and spatial resolution. MR images give information related to location and size of a brain tumor but these images are not able to categorize the tumor grade. Clinicians, therefore, progress to an insidious system where spinal tap and biopsy are time consuming and painful. This incapacity encourages proposing an accurate system to enhance the detection capacity of MR images. Furthermore, the initial stage does not require going for a biopsy as the purpose of proposed system is to classify the brain tumors using MR images. The suggested technique comprises of three major steps that are preprocessing, extraction of features and analysis of classification. In preprocessing, different methods are applied to segment the region of interest. Then intensity, shape and texture features are extracted for each candidate lesion. SVM with different kernels is applied for classification. The automatic proposed system helps in premier detection and efficient provision of rapid outcomes in clinical results. Fig. 2 shows major steps of presented method.

3.1. Preprocessing

For brain tumors detection, data is acquired from publicly and local datasets that consist of anomalous MR images. The artifacts in MR images are to be eliminated, therefore, preprocessing is applied. The processing time is increased when image background is used because it has not necessary information. So skull, background, scalps and eyes are eliminated and not included in the region of interest. Brain and skull can be removed by the Brain Surface Extractor (BSE) method. It is used to identify the edges and perform morphological operations such as erosion. It also performs image masking and cleanup surface. Fig. 3 shows the skull removal.

Then the input image (f(x, y)is converted into gray scale. 5×5 Gaussian filter is applied to noise reduction and smoothing shown in Fig. 4. Fig. 5 presents Gaussian filter results. In Eq. (1), G(x,y) represents Gaussian filter and σ shows the standard deviation of

Please cite this article as: J. Amin et al., A distinctive approach in brain tumor detection and classification using MRI, Pattern Recognition Letters (2017), https://doi.org/10.1016/j.patrec.2017.10.036

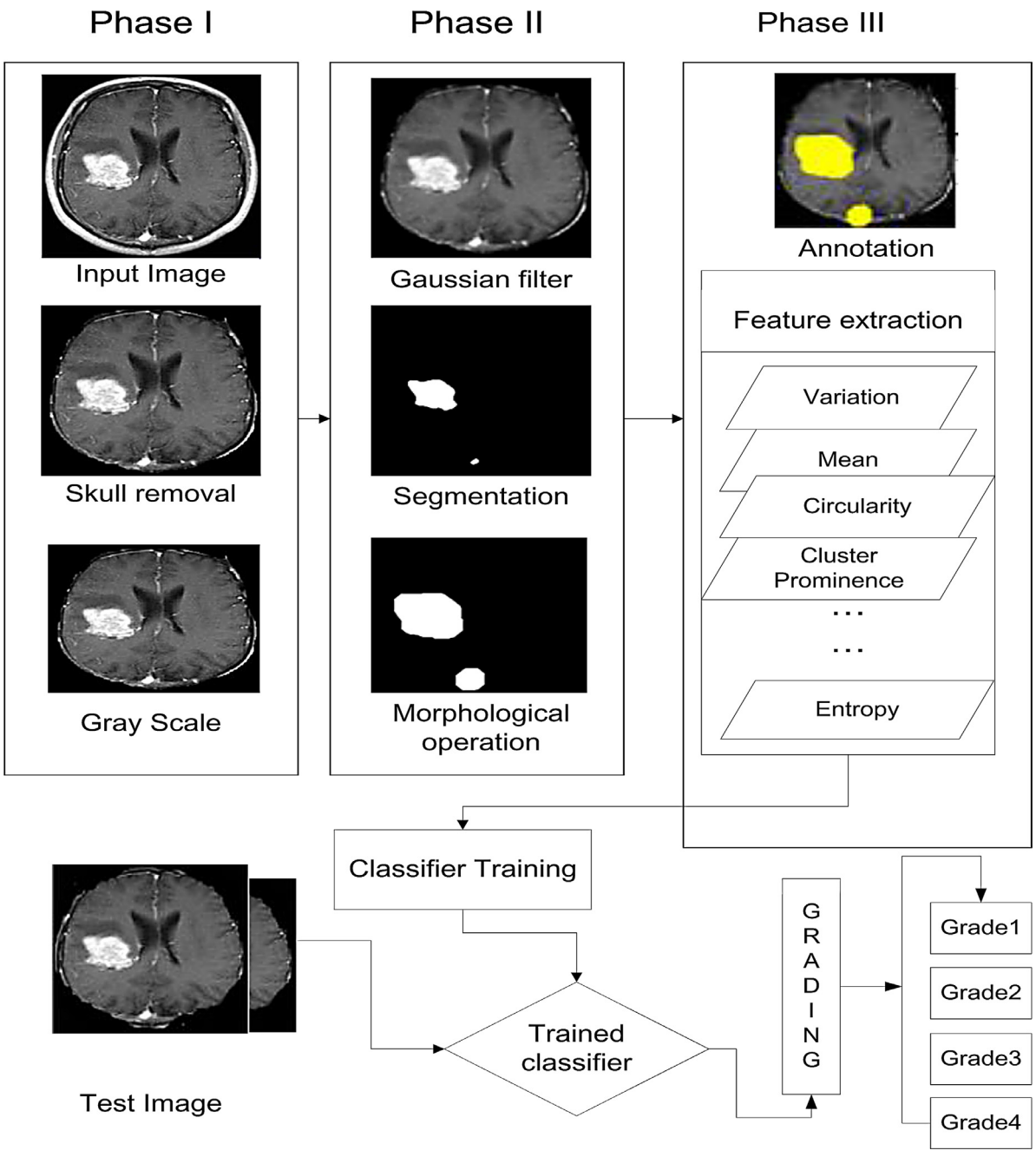

Fig. 2. Major steps of proposed method.

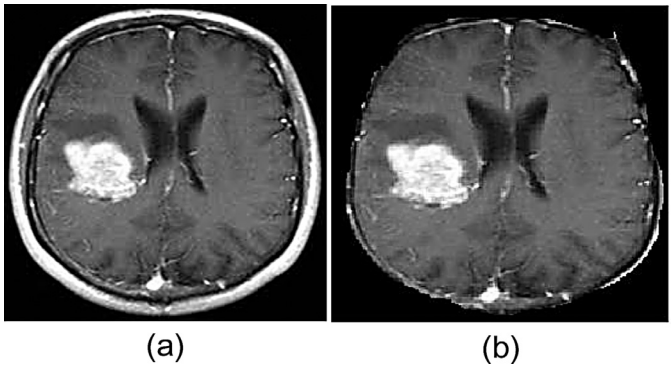

Fig. 3. (a) Input image (b) skull removal.

the Gaussian distribution.

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

In Eq. (2), candidates lesion can be segmented by a threshold selected value T1 = 0.8, because it has been observed during performance evaluation that optimal values are less or greater than 0.8. In this way, False Positive Rate (FPR) is increased and Positive Predicted Value (PPV) and ACC are decreased. Procedure for the selection of optimal threshold value is shown in Table 1.

Fig. 6 presents the tumor results. In Eq. (3), morphological reconstruction by dilation is performed by a disc shaped structuring element with radius 9 because the segmented area of tumor comes along with information loss and dilation recovers the missing pixels. $\oplus$ represents the dilation operation as given in Eq. (3). Algorithm 1 given ahead is for the detection of tumor at the image

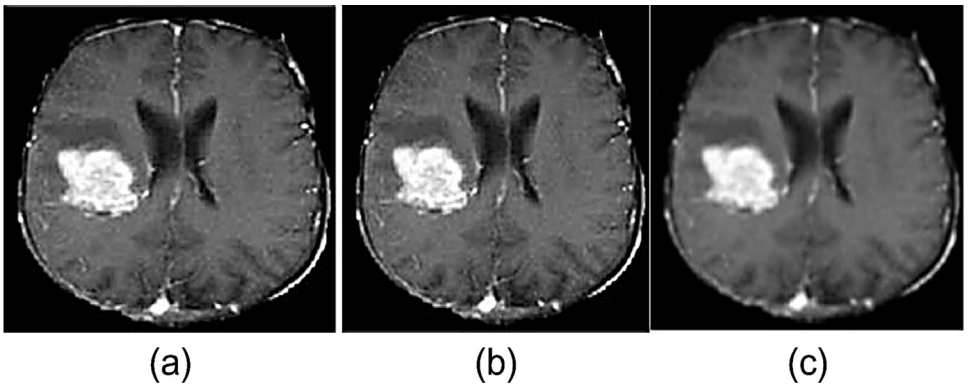

Fig. 4. (a) Skull removal (b) gray scale (c) smoothing.

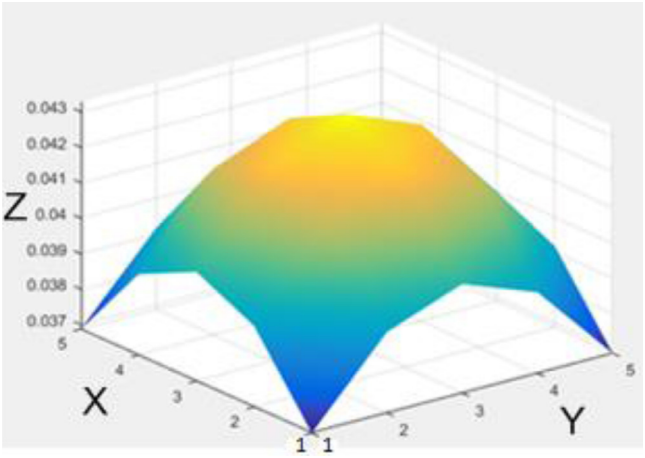

Fig. 5. 5×5 Gaussian filter to remove noise and smoothing the image.

Algorithm 1 Brain tumor detection at image level.

Input: $I_{Input(x,\,y)}$
Output: $I_{Output(x,\,y)}$
Step1: $I_{Input} \leftarrow$ Skull removal on I_{Input}
Step2: $I_{Gray} \leftarrow \frac{I^{R}_{Input} + I^{G}_{Input} + I^{B}_{Input}}{3}$
Step3: $I_{Enhanced(x,\,y)} \leftarrow$ Gaussian filter on I_{Gray}
Step4: $I_{Segmentation} \leftarrow$ Thresholding on $I_{Enhanced(x,\,y)}$
Step5: $I_{Dilation} \leftarrow I_{Segmentation} \oplus s$
return $I_{Output(x,\,y)}$

Table 1
Results of optimal threshold values for tumor detection.

T1	PPV	FPR	False positive (FP)
0.5	0.3451	0.1844	11,013
0.6	0.6437	0.0538	3212
0.7	0.8073	0.0227	1358
0.8	0.9510	0.0040	240
0.9	0.7070	0.0089	550

level.

$$g(x,y) = \begin{cases} 1, & if\ f(x,y) \le T \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

$$g(x,y) = f(x,y) \oplus s \tag{3}$$

For the segmentation of brain tumor at lesion level, after skull removal, the input image is converted to Lab color space with Eqs. (4)–(8) because it gives accurate color balance correction via output curves modifying in a and b components and lightness contrast is adjusted using L component.

$$X, Y, Z \in (\text{CIE data}) \tag{4}$$

$$X_n, Y_n, Z_n \in (\text{CIE values}) \tag{5}$$

Which can be transformed to $L^*a^*b^*$ as follows.

$$L^* = 116\, f\left(\frac{Y}{Y_n}\right) - 16 \tag{6}$$

$$a^* = 500\left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right] \tag{7}$$

$$b^* = 200\left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right] \tag{8}$$

'f' denotes the function which is used for defining L, a and b changes. Then unsupervised learning algorithm is applied such as

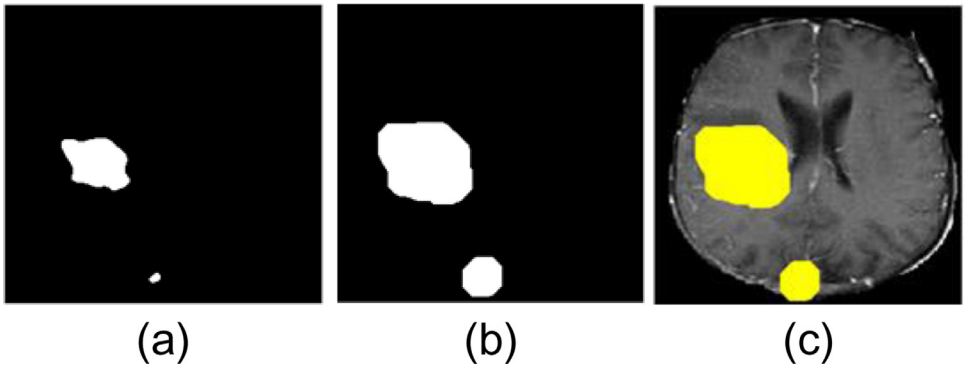

Fig. 6. Results of brain tumor (a) segmentation (b) morphological dilation (c) marking/annotation.

Please cite this article as: J. Amin et al., A distinctive approach in brain tumor detection and classification using MRI, Pattern Recognition Letters (2017), https://doi.org/10.1016/j.patrec.2017.10.036

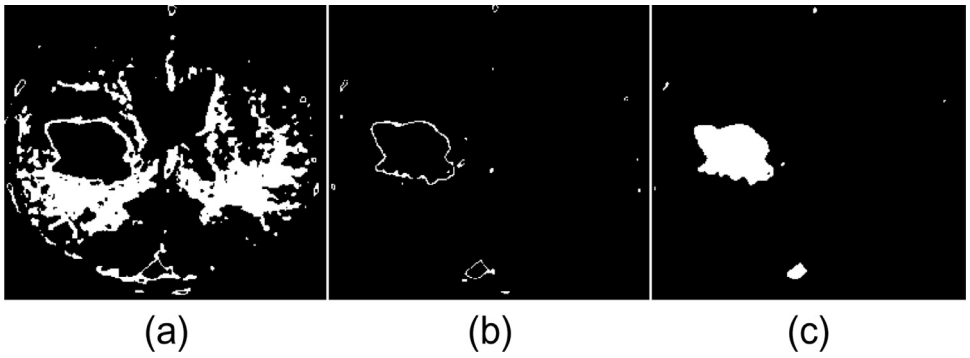

Fig. 7. Results of K-means clustering (a) cluster1 (b) cluster 2 (c) cluster 3.

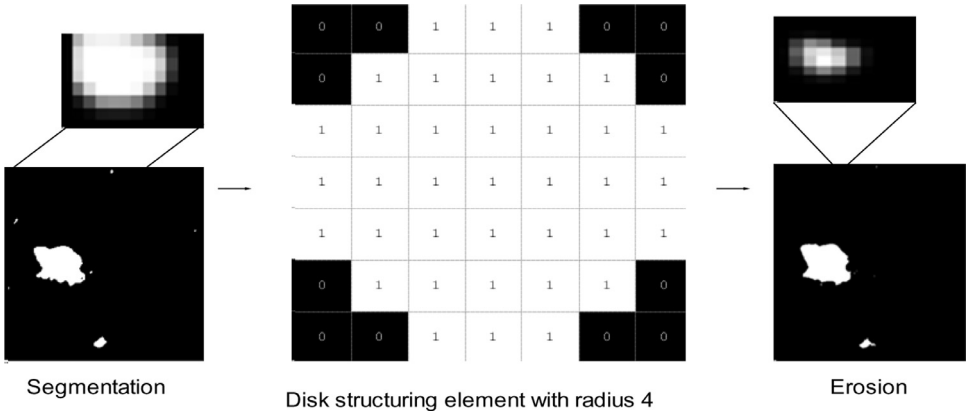

Fig. 8. Process of morphological erosion.

k-means to solve the problem in clusters. In the proposed method, three clusters are used to classify the data because desired results have been achieved in cluster 3. Increasing the size of the cluster will not improve the detection rate rather execution time will be increased. K centers are defined for every cluster. K centers are placed in wiliness way because different results appear in different locations. So, these centers are located distant away from each other. In the last step, a loop is created in which step by step variation occurs in K centers until no further variation is done. The region of interest is selected in the required cluster. $x_i - v_j$ represents the Euclidean distance between x_i and v_j, c_i denotes data points number in the ith cluster and c denotes total cluster centers. The process of K-means clusters is shown in Fig. 7. In cluster 1, the ratio of non-tumor pixels is increased. In cluster 2, the boundary of tumor pixels is obtained and required pixels are achieved in cluster 3.

$$g(x, y) = \sum_{i=1}^{c} \sum_{j=1}^{c_i} (\|x_i - v_j\|)^2 \tag{9}$$

To obtain the required candidate lesions, morphological erosion is applied on g(x,y) and binary image using Eq. (10) to remove extra pixels that are shown in Fig. 8.

$$g(x, y) \ominus B = \{z \in E | B_z \subseteq g(x, y)\} \tag{10}$$

Disk shaped structuring element (B) with radius 4 is used. E denotes integer grid and B_z is B via vector Z translation. Fig. 9 shows effect of erosion on segmented image.

Morphological operations are essential for accurate detection of tumor pixels. If not used, it increases the ratio of non-tumor pixels indicating tumor pixels as a result of which low grade tumor is shown as a high grade tumor. So it is mandatory to apply morphological operations such as erosion and dilation for the correct detection of the tumor. Although erosion removes extra pixels and reduces the FPR, it does not eliminate actual tumor pixels. Even if some tumor pixels are missed, a dilation operation is applied to recover them. Algorithm 2 for the detection of brain tumor at lesion level is mentioned below.

Algorithm 2 Brain tumor detection at lesion level.

Input: $I_{Input(x, y)}$
Output: $I_{Output(x, y)}$
Step1: $I_{Input} \leftarrow$ Skull removal on I_{Input}
Step2: $I_{Lab} \leftarrow L^* + a^* + b^*$
Step3: $I_{Segmentation} \leftarrow K - means$ on I_{Lab}
Step4: $I_{erosion} \leftarrow I_{Segmentation} \ominus s$
Step5: $I_{Dilation} \leftarrow I_{erosion} \ominus s$
return $I_{Output(x, y)}$

3.2. Feature vector

As an automatic system is proposed for tumor detection at image and lesion levels, so for this purpose mixture of features such as texture, shape, and intensity are selected as features set that provide better results for classification. In this article, hybrid features set is used which is consisting of texture, shape and intensity features as mentioned in Table 2.

GLCM of an image ($\emptyset^{\varphi}$) is matrix p(k, l, λx, λy) where k and l are gray level pixels in the image. The matrix element p(k,l) calculates how frequently a pixel with gray level value k comes about horizontally or vertically diagonal to contiguous pixels with the valuel. The component (p(k, l)|d, θ) of a matrix is relative frequency having two pixels at a particular angle (θ) with a definite displacement distance d. Pixel spatial relationship is defined by using offset (λx, λy) which can be determined via direction of (θ)

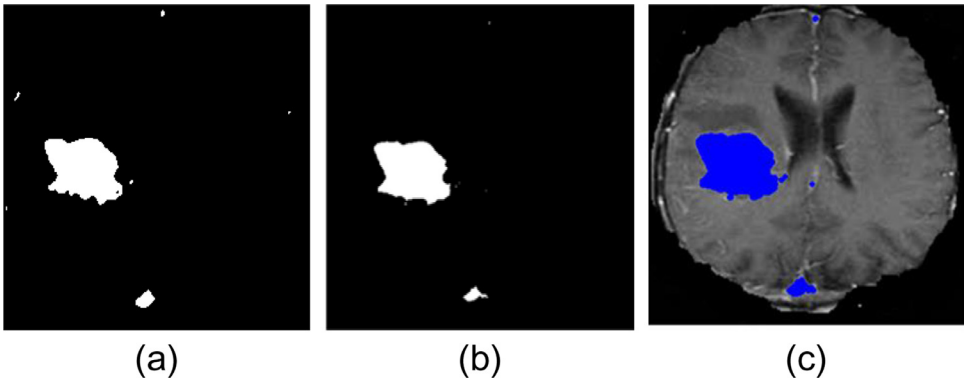

Fig. 9. Results of erosion on segmented image.

Table 2
Features extraction.

No.	Equations	Description
$\emptyset^{f1}$	Energy$(\emptyset^{E}) = \sum_k \sum_l p(k,l)^2$	Property of object that is transferred to another object.
$\emptyset^{f2}$	Variance$(\sum \sigma^2) = \sum_{k=0}^{\emptyset^{r}-1} \sum_{l=0}^{\emptyset^{r}-1} (k-\emptyset^{u})^2 p(k,l)$	The average squared difference from the mean.
$\emptyset^{f3}$	Entropy$(\emptyset^{\sum H}) = -\sum_{k=2}^{2\emptyset^{f}-2} p_{x+y}(k)\log(p_{k+1}(k))$	Used to measure disorder of asystem.
$\emptyset^{f4}$	Dissimilarity $(\emptyset^{D}) = \sum_k \sum_l P(k,l)\lvert k-l\rvert$	Measures the variation in the gray level image.
$\emptyset^{f5}$	Homogeneity$(\emptyset^{\alpha 1}) = \frac{\sum_k^{\emptyset^{r}-1} \sum_l^{\emptyset^{r}-1} p(k,l)}{1+\lvert k-l\rvert}$	Measures non-zero entries in GLCM.
$\emptyset^{f6}$	Cluster shade$(\emptyset^{s}) = \sum_{k=0}^{\emptyset^{r}-1} \sum_{l=0}^{\emptyset^{r}-1} \{k+l-\emptyset^{\mu x}-\emptyset^{\mu y}\}^3 p(k,l)$	The number of operations is reduced.
$\emptyset^{f7}$	Cluster Prominence$(\emptyset^{p}) = \sum_{k=0}^{\emptyset^{r}-1} \sum_{l=0}^{\emptyset^{r}-1} \{k+l-\emptyset^{\mu x}-\emptyset^{\mu y}\}^4 p(k,l)$	Measures the asymmetry or skewness.
$\emptyset^{f8}$	Area$_R(\emptyset^{A}) = \sum_{i=x}^{x+width} \sum_{j=height}^{y+height} S_R(i,j)$	Describes shape of objects in nodule candidate region $S_R(i,j)$.
$\emptyset^{f9}$	Perimeter$(\emptyset^{P}) = 2l+2w$, where l denotes length and w is width	Counts the total boundary pixels.
$\emptyset^{f10}$	Circularity$(\emptyset^{C}) = \frac{(4*\emptyset^{A}*\pi)}{(\emptyset^{p2})}$	How object is close to a true circle.

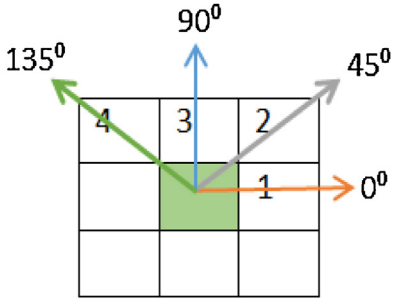

Fig. 10. GLCM: all pixels have adjacent relationship in a given direction.

and distance d. The consistency of p(k,l) is dependent on gray level ($\emptyset^{r}$) and image size. Pixel adjacency is observed based on four angles $\theta = 0^0$, 45°, 90° and 135°. Required offset value d in GLCM defines adjacency pixels via certain distance. Value of offset is set as d = 1.

There are four directions (angles) of pixels against each feature shown in Figs. 10 and 11 to achieve the maximum count of pixels. Also calculated three measures: mean, range and variance against each feature to make them direction independent and then fused each measure with geometric features to get three features set. The GLCM features are calculated as follows:

Let $\emptyset^{f}$ be any feature to calculate its mean ($\emptyset^{\mu f}$), range ($\emptyset^{\text{Range f}}$) and variance ($\sigma^2$). There are four angles as 0°, 45°, 90°, 135° and $\emptyset^{f(1)}$, $\emptyset^{f(2)}$, $\emptyset^{f(3)}$, $\emptyset^{f(4)}$ are the four measures of ($\emptyset^{f}$) against each angle. Hence $\emptyset^{\mu f}$, $\emptyset^{\text{Range f}}$ and σ^2 of $\emptyset^{f}$ are calculated as follows from Eqs. (11)–(13):

$$\emptyset^{\mu f} = \frac{\emptyset^{f1} + \emptyset^{f2} + \emptyset^{f3} + \emptyset^{f4}}{4} \tag{11}$$

$$\emptyset^{\text{Range f}} = \max(\emptyset^{f}) - \min(\emptyset^{f}) \tag{12}$$

$$\sigma^2 = \frac{\left[\emptyset^{f(1)} - \emptyset^{\mu f}\emptyset^{f}\right]^2 + \left[\emptyset^{f(2)} - \emptyset^{\mu f}\emptyset^{f}\right]^2 + \left[\emptyset^{f(3)} - \emptyset^{\mu f}\emptyset^{f}\right]^2 + \left[\emptyset^{f(4)} - \emptyset^{\mu f}\emptyset^{f}\right]^2}{4} \tag{13}$$

3.3. Classification

A number of tests are performed on the selected features set to choose a more appropriate classification strategy. For this purpose, geometrical family SVM is used. Three variants of SVM are tested i.e. Linear, Gaussian and Cubic kernel functions.

3.3.1. Support vector machine (SVM)

In features vector formulation, SVM is used for tumor classification. SVM is a leading classifier in statistical learning domain. SVM uses kernel trick for handling the high dimensional data. Regularization parameter is used in SVM to avoid overfitting, local and global minima. The main idea behind SVM is to drive a hyperplane by maximizing margins between the classes while using support vector. Given a set of examples $\{(x_1, y_1), (x_2, y_2)\ldots (x_r, y_r)\}$, SVM is used to find the solution of optimization problem as given in Eqs. (14)–(15).

$$\text{Minimize}: \quad \frac{< w.w >}{2} + C\sum_{i=1}^{r} \xi_i \tag{14}$$

$$\text{Subject to}: \quad y_i(< w.x_i > + b) \geq 1 - \xi_i, \; i = 1, 2, \ldots, r \tag{15}$$

$$\xi \geq 0, \quad i = 1, 2, \ldots, r$$

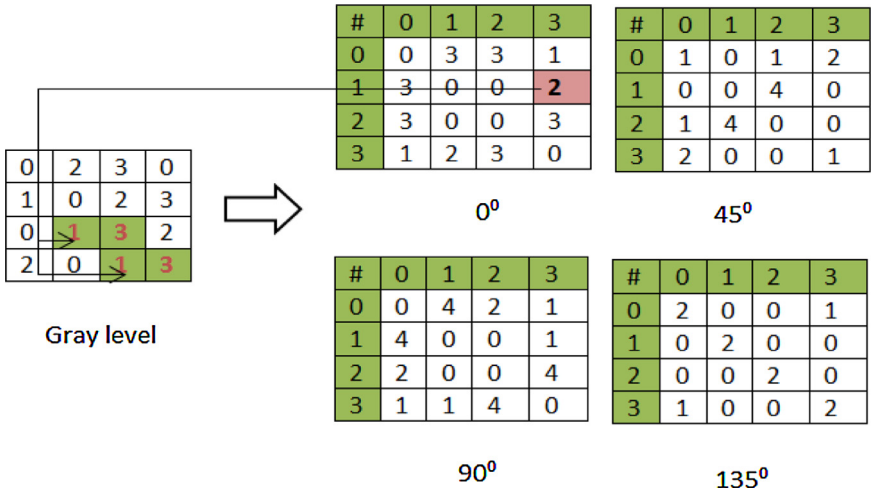

Fig. 11. Pixels adjacency on different angles (a) 0° (b) 45° (c) 90° (d) 135°.

Where $x_i \epsilon R^r$ is an r dimensional space and $y_i = \pm 1$, where 1 and −1 represent cancerous and non-cancerous cells respectively. C represent the convex optimization. Besides linear classification, SVM also performs nonlinear classification with high precision by using kernel tricks. In this regard, SVM transforms features space from a lower dimension to a high dimensional space. Three different kernels including Gaussian Radial Based Function (RBF), Linear and Cubic are used in the proposed methodology to improve accuracy and reduce the FPR. These kernels are mentioned using Eqs. (16)–(18).

$$\text{Linear Kernel } K(X_i, X_j) = X_i^T . X_j \tag{16}$$

$$\text{Radial} - \text{Basis Function } K(X_i, X_j) = \exp\left(-\frac{X_i - X_j^2}{2\sigma^2}\right) \tag{17}$$

$$\text{CubickernelK}(X_i, X_j) = (X_i^T . X_j + 1)^2 \tag{18}$$

Better speculation of classifiers is achieved using cross-validation strategy. In cross validation method, 5, 10, 15, 20, 25 and 30 fold is used for classification. 5 fold means that whole data is alienated into five sets of which one set is for the intention of testing and four are considered for training. Similarly, in the next phase, 10 fold means to divide the whole data into ten sets such that one set is for testing purpose and remaining nine are used for training. This procedure continues for the 15, 20, 25 and 30 fold in different iterations.

4. Datasets for experiments

Three datasets are used for the performance evaluation of proposed system.

85 images with ground truth annotation are obtained from Nashtar Hospital Multan in which 46 images show the sign of tumor and 39 images are healthy.

Harvard dataset consists of 100 images in which 65 images contain tumor and 35 images are healthy [21].

Cancer imaging archive (TCIA) organized RIDER brain image data that consists of 126 patients in which each case per patient comprises of multiple studies. Data is collected from Henry Ford Hospital (RETRO) and TJU Institute. By following this data, information about the patients suffering from different stages of brain diseases is obtained i.e., Astrocytoma, GBM, and Oligodendrogliomas along with race (white, black). Pathologists assign grades 1, 2, 3, 4 due to the survival rate of disease in which 41 images belong to grade 2, 36 are from grade 1, 26 images belong to grade 3 and 23 are from grade 4 tumor [22].

Table 3
Proposed method results on tumor and non-tumor on Nashtar hospital dataset using Linear kernel.

Cross validation	ACC	AUC	Sensitivity	Specificity	FNR	FPR
5 fold	98.8%	0.98	96.9%	100%	0.03	0.00
10 fold	92.2%	1.00	80.9%	100%	0.19	0.00
15 fold	96.0%	0.97	94.1%	100%	0.05	0.00
20 fold	96.6%	0.98	95.0%	100%	0.05	0.00
25 fold	96.0%	0.99	94.1%	100%	0.05	0.00
30 fold	95.5%	0.96	90.4%	100%	0.09	0.00

Table 4
Proposed method results on tumor and non-tumor on Nashtar hospital dataset using Cubic kernel.

Cross validation	ACC	AUC	Sensitivity	Specificity	FNR	FPR
5 fold	93.5%	0.94	90.4%	95.6%	0.09	0.04
10 fold	79.9%	0.91	90.4%	72.5%	0.09	0.27
15 fold	93.3%	0.94	94.1%	91.4%	0.05	0.08
20 fold	94.6%	0.94	94.1%	95.7%	0.05	0.04
25 fold	94.0%	0.94	94.1%	93.6%	0.05	0.06
30 fold	77.3%	0.97	95.2%	64.8%	0.04	0.35

Table 5
Proposed method results on tumor and non-tumor on Nashtar hospital dataset using Gaussian kernel.

Cross validation	ACC	AUC	Sensitivity	Specificity	FNR	FPR
5 fold	96.1%	0.97	95.2%	96.7%	0.04	0.03
10 fold	97.6%	0.99	96.9%	98.0%	0.03	0.02
15 fold	85.9%	0.99	96.9%	78.8%	0.03	0.21
20 fold	89.4%	1.00	83.0%	100%	0.17	0.00
25 fold	87.1%	1.00	79.2%	100%	0.20	0.00
30 fold	98.1%	0.98	95.2%	100%	0.04	0.00

4.1. Results and discussion

For performance evaluation of presented method, tumor detection on image level has been done on Harvard and local datasets. Moreover, detailed classification has been performed on RIDER dataset. Tumor detection is made on lesion level and its results are compared with ground truth annotation that is publicly available. ACC and AUC etc. of suggested method taken from benchmark datasets are mentioned in Tables 3–11. Eqs. (19)–(25) are used to

Table 6
Proposed method results on tumor and non-tumor on Harvard dataset using Linear kernel.

Cross validation	ACC	AUC	Sensitivity	Specificity	FNR	FPR
5 fold	98.1%	1.00	92.5%	100%	0.08	0.00
10 fold	99.4%	1.00	98.4%	100%	0.02	0.00
15 fold	98.7%	0.99	96.8%	100%	0.01	0.00
20 fold	96.1%	0.94	90.4%	100%	0.10	0.00
25 fold	95.5%	0.93	88.8%	100%	0.05	0.00
30 fold	94.8%	1.00	87.3%	100%	0.06	0.00

Table 7
Proposed method results on tumor and non-tumor on Harvard dataset using Cubic kernel.

Cross validation	ACC	AUC	Sensitivity	Specificity	FNR	FPR
5 fold	93.5%	0.96	92.0%	94.5%	0.08	0.05
10 fold	87.7%	0.95	92.0%	84.6%	0.08	0.15
15 fold	81.2%	0.90	87.3%	76.9%	0.12	0.23
20 fold	75.3%	0.89	87.3%	67.0%	0.12	0.33
25 fold	83.1%	0.91	0.88%	79.1%	0.12	0.20
30 fold	78.6%	0.90	88.8%	71.4%	0.11	0.28

Table 8
Proposed method results on tumor and non-tumor on Harvard dataset using Gaussian kernel.

Cross validation	ACC	AUC	Sensitivity	Specificity	FNR	FPR
5 fold	95.5%	0.97	95.2%	95.6%	0.04	0.04
10 fold	96.8%	0.97	95.2%	97.8%	0.04	0.02
15 fold	97.4%	0.99	96.8%	97.8%	0.03	0.02
20 fold	98.7%	1.00	96.8%	100%	0.03	0.00
25 fold	86.4%	0.98	69.8%	97.8%	0.30	0.02
30 fold	92.2%	1.00	80.9%	100%	0.19	0.00

Table 9
Proposed method results on grading the input image at lesion level using Linear kernel on RIDER dataset.

Cross validation	ACC	AUC	Sensitivity	Specificity	FNR	FPR
5 fold	81.3%	0.86	92.5%	62.5%	0.08	0.38
10 fold	75.0%	0.73	77.5%	70.8%	0.23	0.30
15 fold	85.9%	0.87	92.5%	75.0%	0.08	0.25
20 fold	79.1%	0.86	92.5%	58.3%	0.08	0.42
25 fold	90.0%	0.88	92.3%	85.7%	0.08	0.15
30 fold	78.1%	0.76	92.5%	66.6%	0.08	0.34

Table 10
Proposed method results on grading the input image at lesion level using Gaussian kernel on RIDER dataset.

Cross validation	ACC	AUC	Sensitivity	Specificity	FNR	FPR
5 fold	93.3%	1.00	92.1%	100%	0.08	0.00
10 fold	88.3%	1.00	88.2%	55.5%	0.12	0.45
15 fold	84.0%	0.87	77.7%	87.5%	0.23	0.13
20 fold	80.0%	0.92	66.6%	87.5%	0.34	0.13
25 fold	89.3%	0.93	83.3%	100%	0.17	0.00
30 fold	96.8%	0.97	95.2%	97.8%	0.05	0.03

Table 11
Proposed method results on grading the input image at lesion level using Cubic kernel on RIDER dataset.

Cross validation	ACC	AUC	Sensitivity	Specificity	FNR	FPR
5 fold	95.0%	1.00	92.1%	100%	0.08	0.00
10 fold	96.7%	1.00	95.9%	100%	0.05	0.00
15 fold	86.2%	0.77	94.4%	72.7%	0.06	0.28
20 fold	80.0%	0.92	66.6%	0.87%	0.34	0.13
25 fold	100%	1.00	100%	100%	0.00	0.00
30 fold	85.0%	0.88	85.7%	84.6%	0.15	0.16

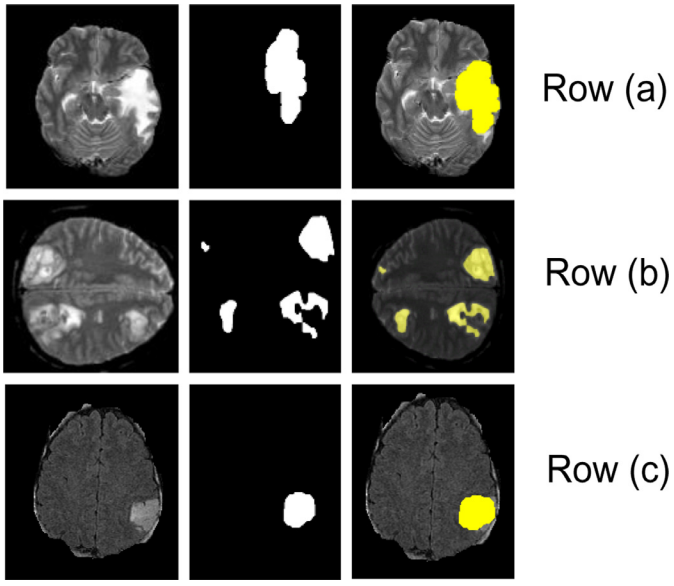

Fig. 12. Proposed method results of tumor detection. Row (a) shows local dataset results, Row (b) shows Harvard dataset results and Row (c) shows RIDER dataset results.

obtain ACC, AUC, sensitivity, specificity, PPV, FPR and False Negative Rate (FNR) respectively. Total lesions pixels indicate True Positive (TP). True Negative (TN) denotes the non-tumor pixels. False Positive (FP) represents the wrong detection of healthy pixels. Lesion pixels that are not detected by the proposed algorithm are indicated as False Negative (FN). TPR denotes true positive rate.

$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{19}$$

$$\mathrm{AUC} = \int_{\infty}^{-\infty} \mathrm{TPR}(\mathrm{T})\mathrm{FPR}'(\mathrm{T})\mathrm{dT} \tag{20}$$

$$\mathrm{Sensitivity} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{21}$$

$$\mathrm{Specificity} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \tag{22}$$

$$\mathrm{PPV} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{23}$$

$$\mathrm{FPR} = 1 - \mathrm{Specificity} \tag{24}$$

$$\mathrm{FNR} = 1 - \mathrm{Sensitivity} \tag{25}$$

The suggested method successfully detects the tumor in MRI shown in Fig. 12.

The overall performance of proposed system is improved as compared to previous methods because accurate classification depends on the segmentation of lesions and extraction of descriptors. The proposed method accurately segments the tumor at image and lesion levels. The extracted feature set consists of texture, shape, and intensity features which help in good classification. With the classifier, the tumor is detected at the image level and the tested image is graded as a tumor or healthy image as mentioned in Tables 3–8. Similarly, when the tumor is detected via classifier at lesion level, the tested image is labeled as grade 1, 2 also called benign (low grade tumor) and grade 3, 4 called malignant (highgrade tumor) as mentioned in the Tables 9–11.

Three variant kernels of SVM are applied on benchmark datasets for the comparison of presented methodology results. At

Table 12
Overall performance of proposed method using variant kernels of SVM classifier with different cross-validation approach.

Datasets	Cross validation	Linear		Cubic		Gaussian	
		ACC	AUC	ACC	AUC	ACC	AUC
Local	5 fold	98.8%	0.98	93.5%	0.94	96.1%	0.97
	10 fold	92.2%	1.00	79.9%	0.91	97.6%	0.99
	15 fold	96.0%	0.97	93.3%	0.94	85.9%	0.99
	20 fold	96.6%	0.98	94.6%	0.94	89.4%	1.00
	25 fold	96.0%	0.99	94.0%	0.94	87.1%	1.00
	30 fold	95.5%	0.96	77.3%	0.97	98.1%	0.98
RIDER	5 fold	81.3%	0.86	95.0%	1.00	93.3%	1.00
	10 fold	75.0%	0.73	96.7%	1.00	88.3%	1.00
	15 fold	85.9%	0.87	86.2%	0.77	84.0%	0.87
	20 fold	79.1%	0.86	80.0%	0.92	82.8%	1.00
	25 fold	90.0%	0.88	100%	1.00	89.3%	0.93
	30 fold	78.1%	0.76	85.0%	0.88	96.8%	0.97
Harvard	5 fold	98.1%	1.00	93.5%	0.96	95.5%	0.97
	10 fold	99.4%	1.00	87.7%	0.95	96.8%	0.97
	15 fold	98.7%	0.99	81.2%	0.90	97.4%	0.99
	20 fold	96.1%	0.94	75.3%	0.89	98.7%	1.00
	25 fold	95.5%	0.93	83.1%	0.91	86.4%	0.98
	30 fold	94.8%	1.00	78.6%	0.90	92.2%	1.00

Table 13
Proposed method comparison with the existing methods.

Existing methods	Year	Results
Sudharani and Sarma [20]	2016	89.20% ACC, 88.9% sensitivity, 90% specificity
Nabizadeh et al. [23]	2014	91.50% ACC
Subashini et al. [24]	2016	91.00% ACC
Abdel-Maksoud et al. [25]	2015	95.06% ACC
Nabizadeh and Kubat [26]	2015	79.3 ± 0.3 ACC, 77.3 ± 1.4 sensitivity, 81.3 ± 1.2 specificity
Proposed Method	Average 97.1% ACC, 0.98 AUC, 91.9% sensitivity and 98.0% specificity	

Table 14
Running time comparison of the individual proposed method steps with the existing methods.

Steps of techniques	Running time of the existing methods	Running time of the proposed method
Tumor slice detection	12 min [23]	13.32 sec MATLAB (2016b) image processing and vision toolbox on Intel Corei7 3.4 GHz
	13.92 sec [25] on MATLAB (R2011a) core i5/2.4 GHZ	
Feature extraction	17 min on Gabor wavelet feature extraction, 18 min on statistical feature extraction [23] on Intel Xeon CPU X5472 machine at 3 GHz	58.52 sec on shape, texture and intensity features extraction on MATLAB 2016b image processing and vision toolbox on Intel Corei7 3.4 GHz
SVM Linear kernel	0.52 sec on [24]	0.30 sec
SVM Cubic kernel	–	0.22 sec
SVM Gaussian kernel	–	0.20 sec

the linear kernel 5 fold cross validation, maximum 98.8% ACC, 0.98 AUC, 96.9% sensitivity and 100% specificity are achieved on the local dataset. Similarly, at the linear kernel 10 fold cross-validation, 99.4% ACC, 1.00 AUC, 98.4% sensitivity and 100% specificity are obtained at Harvard dataset. Whereas in the case of RIDER dataset obtained is 100% ACC, 1.00 AUC, sensitivity 100% and 100% specificity at the cubic kernel on 25 fold cross validation. It is important to note that AUC (based on ROC) and accuracy are not the same concept. Overall accuracy is based on one specific cut point while ROC tries all cut points to plot the sensitivity and specificity. So as accuracy is compared on one cut point, it varies for different such cut points. Results of overall comparison of all benchmark datasets is shown in Table 12. The proposed method comparison with the existing methods is given in Table 13. Execution time comparisons of proposed method with the existing techniques are mentioned in Table 14.

5. Conclusion

In this article, an automated system is presented for the segmentation and classification of brain cancer utilizing Magnetic Resonance Imaging (MRI). Testing of proposed method is performed at image and lesion levels for detailed assessment. Performance measures such as area under the curve (AUC) and accuracy (ACC) are used for the evaluation of suggested method on publicly available and one local dataset. The suggested methodology can be used to increase tumor detection procedure at an early stage before the complications phase. Outcomes of experiments reveal the outperformance of presented methodology as compared to previous methods. This work will be helpful in examining the brain tumors precisely and accurately.

References

[1] Haj-Hosseini Neda, Peter Milos, Camilla Hildesj, Martin Hallbeck, Johan Richter, Karin Wrdell, Fluorescence spectroscopy and optical coherence tomography for brain tumor detection, in: SPIE Photonics Europe, Biophotonics: Photonic Solutions for Better Health Care, Brussels Belgium, SPIE-International Society for Optical Engineering, 2016, pp. 9887–9896.
[2] Y. Liu, A. Carpenter, H. Yuan, Z. Zhou, M. Zalutsky, G. Vaidyanathan, H. Yan, T Vo-Dinh, Goldnanostar as theranostic probe for brain tumor sensitive PET-optical imaging and image-guided specific photothermal therapy, Cancer Res. 76 (14) (2016) 4213.
[3] D.N. Louis, A. Perry, G. Reifenberger, A. von Deimling, D. FigarellaBranger, W.K. Cavenee, H. Ohgaki, O.D. Wiestler, P. Kleihues, D.W. Ellison, The World Health Organization classification of tumors of the central nervous system: a summary, Acta Neuropathol. 131 (6) (2016) 803–820.
[4] S.J. Choi, J.S. Kim, J.H. Kim, S.J. Oh, J.G. Lee, C.J. Kim, Y.S. Ra, J.S Yeo, J.S Ryu, D.H. Moon, [18F] 3-deoxy-3-fluorothymidine PET for the diagnosis and grading of brain tumors, Eur. J. Nuclear Med. Mol. Imag. 32 (6) (2005) 653–659.

Please cite this article as: J. Amin et al., A distinctive approach in brain tumor detection and classification using MRI, Pattern Recognition Letters (2017), https://doi.org/10.1016/j.patrec.2017.10.036

[5] V. Janani, P. Meena, Image segmentation for tumor detection using fuzzy inference system, Int. J. Comput. Sci. Mobile Comput. (IJCSMC) 2 (5) (2013) 244–248.
[6] B. Dong, A. Chien, Z. Shen, Frame based segmentation for medical images, Commun. Math. Sci. 32 (4) (2010) 1724–1739.
[7] J. Patel, K. Doshi, A study of segmentation methods for detection of tumor in brain MRI, Adv. Electron Electr. Eng. 4 (3) (2014) 279–284.
[8] M. Rohit, S. Kabade, M.S. Gaikwad, Segmentation of brain tumour and its area calculation in brain MRI images using K-mean clustering and Fuzzy C-mean algorithm, Int. J. Comput. Sci. Eng. Technol. (IJCSET) 4 (5) (2013) 524–531.
[9] H.A. Aslam, T. Ramashri, M.I.A. Ahsan, A new approach to image segmentation for brain tumor detection using pillar K-means algorithm, Int. J. Adv. Res. Comput. Commun. Eng. 2 (3) (2013) 1429–1436.
[10] V. Sehgal, Z. Delproposto, D. Haddar, E.M. Haacke, A.E. Sloan, L.J. Zamorano, G. Barger, J. Hu, Y. Xu, K.P. Prabhakaran, I.R. Elangovan, Susceptibility-weighted imaging to visualize blood products and improve tumor contrast in the study of brain masses, J. Magnet. Reson. Imaging 24 (1) (2006) 41–51.
[11] A. Mustaqeem, A. Javed, T. Fatima, An efficient brain tumor detection algorithm using watershed and thresholding based segmentation, Int. J. Image Graph. Signal Process. 4 (10) (2012) 34–39.
[12] S.J. Prajapati, K.R. Jadhav, Brain tumor detection by various image segmentation techniques with introduction to non-negative matrix factorization, Brain 4 (3) (2015) 600–603.
[13] Dipak Kumar Kole, Amiya Halder, Automatic brain tumor detection and isolation of tumor cells from MRI Images, Int. J. Comput. Appl. 39 (16) (2012) 26–30.
[14] K.M. Iftekharuddin, J. Zheng, M.A Islam, R.J. Ogg, Fractal-based brain tumor detection in multimodal MRI, Appl. Math. Comput. 207 (1) (2009) 23–41.
[15] T. Logeswari, M. Karnan, An improved implementation of brain tumor detection using segmentation based on hierarchical self organizing map, Int. J. Comput. Theory Eng. 2 (4) (2010) 1793–8201.
[16] S.K. Bandhyopadhyay, T.U. Paul, Automatic segmentation of brain tumour from multiple images of brain MRI, Int. J. Appl. Innovat. Eng. Manage. (IJAIEM) 2 (1) (2013) 240–280.
[17] Tuhin Utsab Paul, Samir Kumar Bandhyopadhyay, Segmentation of brain tumor from brain MRI images reintroducing K -Means with advanced dual localization method, Int. J. Eng. Res. Appl.(IJERA) 2 (3) (2012) 226–231.
[18] A. Meena, R. Raja, Spatial fuzzy c means pet image segmentation of neurodegenerative disorder, Comput. Vis. Pattern Recognit. 4 (1) (2013) 50–55.
[19] P. Vasuda, S. Satheesh, Improved Fuzzy C-Means algorithm for MR brain image segmentation, Int. J. Comput. Sci. Eng. 2 (5) (2010) 1713–1715.
[20] K. Sudharani, T.C. Sarma, K.S. Prasad, Advanced morphological technique for automatic brain tumor detection and evaluation of statistical parameters, Procedia Technol. 24 (2016) 1374–1387.
[21] D. Summers, Harvard whole brain atlas: www. med. harvard.edu/AANLIB/home.html, J. Neurol. Neurosurg. Psychiatry 74 (3) (2003) 288–288.
[22] Armato, Sam, Reinhard Beichel, Luc Bidaut, Larry Clarke, Bar- bara Croft, Chuck Fenimore, Marios Gavrielides et al., RIDER(Reference Database to Evaluate Response) Committee Combined Report, 9/25/2008 Sponsored by NIH, NCI, CIP, ITDB Causes of and Methods for Estimating/Ameliorating Variance in the Evaluation of Tumor Change in Response-to Therapy, https://wiki.cancerimagingarchive.net/display/Public/Collections.
[23] N. Nabizadeh, N John, C. Wright, Histogram-based gravitational optimization algorithm on single MR modality for automatic brain lesion detection and segmentation, Expert Syst.Appl. 41 (17) (2014) 7820–7836.
[24] MM Subashini, SK Sahoo, V Sunil, S Easwaran, A non-invasive methodology for the grade identification of astrocytoma using image processing and artificial intelligence techniques, Expert Syst.Appl. 43 (2016) 186–196.
[25] Abdel-Maksoud Eman, Elmogy Mohammed, Al-Awadi Rashid, Brain tumor segmentation based on a hybrid clustering technique, Egypt. Inf. J. 16 (1) (2015) 71–81.
[26] N. Nabizadeh, M. Kubat, Brain tumors detection and segmentation in MR images: Gabor wavelet vs. statistical features, Comput. Electr. Eng. 45 (2015) 286–301.
[27] Megha Kadam, Avinash Dhole, Brain tumor detection using GLCM with the help of KSVM, Int. J. Eng. Tech. Res. (IJETR) 7 (2) (2017) 2454–4698.

Please cite this article as: J. Amin et al., A distinctive approach in brain tumor detection and classification using MRI, Pattern Recognition Letters (2017), https://doi.org/10.1016/j.patrec.2017.10.036