Research article

Automated multi-class brain tumor types detection by extracting RICA based features and employing machine learning techniques

Sadia Anjum[1], Lal Hussain[2,3,5,*], Mushtaq Ali[1], Adeel Ahmed Abbasi[2,4] and Tim Q. Duong[5]

[1] Department of IT, Hazara University, Mansehra 21120, KPK, Pakistan

[2] Department of Computer Science & IT, University of Azad Jammu and Kashmir, King Abdullah Campus, Muzaffarabad 13100, Pakistan

[3] Department of Computer Science & IT, University of Azad Jammu and Kashmir, Neelum Campus, Athmuqam 13230, Pakistan

[4] School of Computer Science and Engineering, Central South University, 932 Lushan S Rd, Yuelu District, Changsha, Hunan, China

[5] Department of Radiology, Albert Einstein College of Medicine and Montefiore Medical Center, 111 East 210th Street, Bronx, NY 10467, USA

* **Correspondence**: Email: lall_hussain2008@live.com.

Abstract: Among the other cancer types, the brain tumor is one the leading cause of cancer across globe. If the tumor is properly identified at an earlier stage, then the chances of the survival can be increased. To categorize the brain tumor there are several factors including texture, type and location of brain tumor. We proposed a novel reconstruction independent component analysis (RICA) feature extraction method to detect multi-class brain tumor types (pituitary, meningioma, and glioma). We then employed the robust machine learning techniques as support vector machine (SVM) with quadratic and linear kernels and linear discriminant analysis (LDA). For training and testing of the data validation, a 10-fold cross validation was employed. For the multi-class classification, the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy and AUC were, respectively, 97.78%, 100%, 100%, 99.07, 99.34% and 0.9892 to detect pituitary using SVM Cubic followed by meningioma with accuracy (96.96%0, AUC (0.9348) and glioma with accuracy (95.88%), AUC (0.9635). The findings indicates that RICA feature based proposed methodology has more potential to detect the multiclass brain tumor types for improving diagnostic efficiency and can further improve the prediction accuracy to achieve the clinical outcomes.

Keywords: machine learning; feature extraction; image analysis; meningioma; glioma; pituitary

1.   Introduction

A brain tumor is comprised of abnormal cells in central spinal canal or brain or intracranial hard neoplasms, which are either benign or malignant [1]. In 2019 in the United States [2], about 86, 010 new cases of non-malignant and malignant brain tumor are estimated to be analyzed. There were 79, 718 deaths recognized to malignant brain between 2012 and 2016 with annual average mortality rate of 4.42. The mortality rate in adults and childern due to brain tumor has increased.

The brain tumor subtypes classification is challenging based on several factors. The experts still facing challenging to improve the detection accuracy by developing the latest technology. Several approaches are required to identify the brain tumor. The brain tumor is one of the fatal forms of cancer among other cancer types having aggressive nature, heterogenous characteristics, and low survival rate. Due to several factors such as type, location and texture properties, the brain tumor is categorized into different types (e.g. Meningioma, CNS Lymphoma, Glioma, Acoustic Neuroma, Pituitary etc.) [3]. The clinically rate of incident of meningioma, pituitary, and glioma among other brain tumor types is 15%, 15% and 45% respectively [4]. The patient survival can be predicted and diagnosed based on the tumor type through which they can decide the relevant treatment choice ranging from chemotherapy to radiotherapy. Thus in order to properly planning and monitoring the brain tumor, tumor grading is highly desired [5].

The glioma is the major tumor type which has further three types such as 1) Ependymomas, 2) Astrocytomas, and 3) Oligodendrogliomas. From the glial cells, it originates and surround nerve cells. According to the genetic features, it can be further identified which help the prediction of future treatment and behavior. The meningioma is another type of tumor, which originates in brain. It occur in women and grow slowly without any symptom [6]. The pituitary tumor type grows in the pituitary gland. The pituitary glands are benign and don't spread in the whole body [6].

The researchers recently employed many artificial intelligences-based machine learning methods to predict the tumor. The feature extraction is the most crucial part in the machine learning techniques for computing the most relevant features, which is still a challenging task for researchers. In order to the select and compute the most relevant feature is a tedious task, which require the prior knowledge about the domain of the problem. The morphological features to detect the brain tumor types can led easily to misclassification as different tumor types have similar resemblance. The extracted features are then fed as input to the different brain tumor type [7]. Recently, the researchers have computed different feature extraction methods including Elliptic Fourier descriptors (EFDs), texture, scale invariant Fourier transform (SIFT) and morphological features. Rathore et al. [8] used ensemble methods to detect the colon biopsy by computing hybrid features. Rathore et al. [9] also computed geometric features for prediction of colon cancer. Hussain et al. [10] extracted EFDs, SIFT, texture, entropy and morphological features to detect the prostate cancer. Moreover, Asim et al. [11] computed the hybrid features to detect the Alzheimer disease (AD). The graphical method is expensive, and computer aided diagnosis (CAD) methods could not properly capture the background knowledge regarding the morphological features as these methods are based merely on the texture properties. To properly detect the brain tumor with its location, the radiologists analyzed the image features which are dependent on their personal skills and expertise. The hand-crafted features are still a tedious and challenging task as selecting and computing more relevant features is still challenging.

In the past, researchers employed various machine learning (ML) algorithms by computing various features extracting approaches in the medical fields. The Gray level co-occurrence matrix (GLCM) and Berkeley wavelet transform (BWT) features were extracted by [12] to detect brain tumor. Moreover, Reboucas et al. [13] computed GLCM features to analyze the human tissue densities. Dhruv et al. [14] studied the GLCM and Haralick texture features for the analysis of 3D medical images. Hussain et al. [10] applied support vector machines (SVM) with its kernels to detect prostate cancer by extracting combination of feature extracting strategies. Zheng et al. [15] integrated the SVM and graph cuts for medical image segmentation. Taie and Ghonaim [16] applied Chicken Swarm Optimization (CSO) based algorithms alongwith SVM for brain tumor's disease diagnosis. Abd-Ellah et al. [17] used kernel SVM to classify the brain tumor MRIs. Alquran et al. [18] applied SVM to detect the melanoma skin cancer. Wang et al. [19] proposed stationary wavelet entropy (SWE) to extract brain image features. They obtained improved classification performance results by replacing wavelet entropy (WE), discrete wavelet transform (DWT) and wavelet energy (WN) with the proposed SWE. The SWE averaged the variants of DWT. Zhang et al. [20] computed the Hu moment invariant (HMI) features from a specific MR brain image and then fed these HMI features to generalized eigenvalue proximal SVM (GEPSVM) and twin support vector machine (TSVM). The proposed methods outperformed in detection of brain tumor.
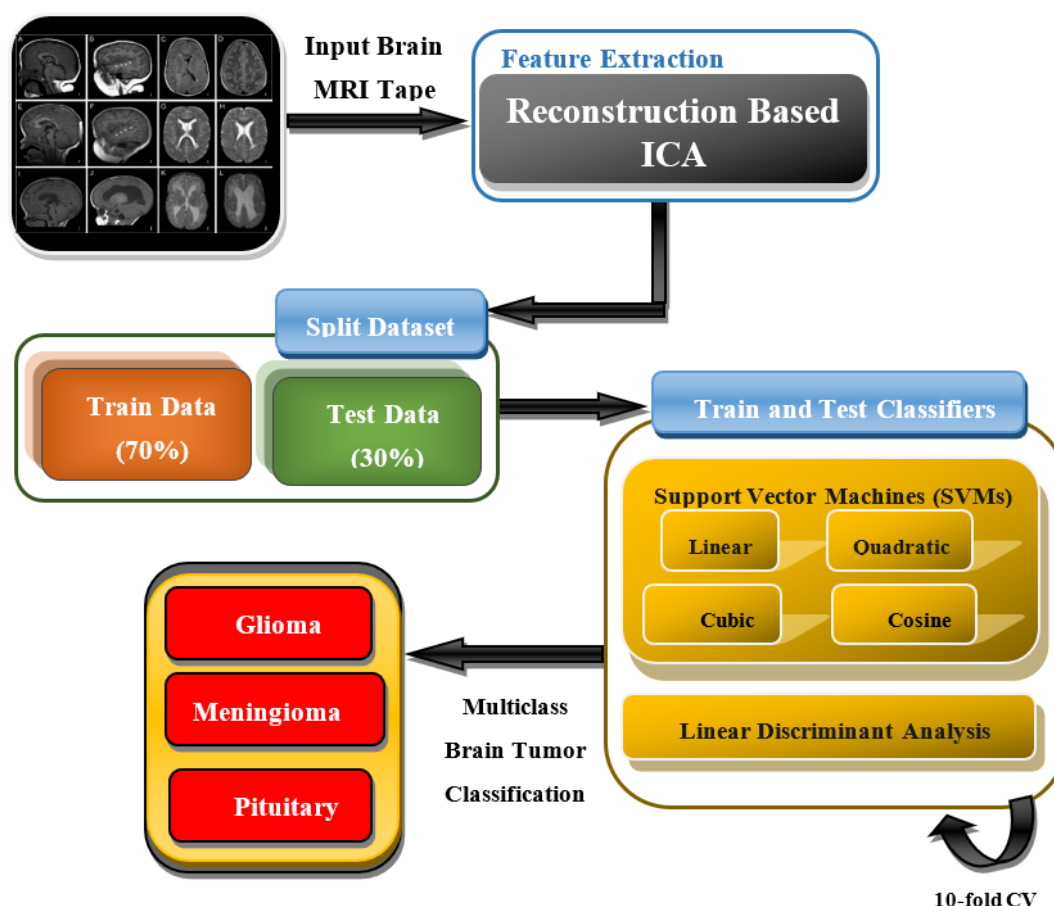

Figure 1. Schematic Diagram to detect Multi-class Brain tumor types by computing RICA features and employing Machine learning techniques.

In this study, we extracted traditional features such as entropy, morphological, texture, EFDs, SIFT and proposed new feature extraction approach based on the RICA features to classify a multi-class brain tumor types and applied ML techniques.

The Figure 1 shows the schematic diagram to detect the Multi-class brain tumor types (i.e. Meningioma, Glioma and Pituitary) by extracting RICA based features from Brain MRIs and applied ML techniques such as SVM with its kernels and LDA with 10-fold cross validation. After extracting the features, the MRI data was split into 70% for training and 30% for testing.

2. Material and methods

2.1. Dataset

The brain tumor CE-MRI dataset used in this study were taken from the publicly available database provided by the School of Biomedical Engineering, Southern Medical University, Guangzhou, China (https://figshare.com/articles/dataset/brain_tumor_dataset/1512427). The data details are used in the previous studies of detailed in [21] brain tumor adaptive sparse pooling, [22] brain tumor via region augmentation proposed by Cheng et al. [21,22] which contains 3064 T1-weighted contrast-enhanced MRI images acquired from Nanfang Hospital and General Hospital, Tianjin Medical University, China from 2005. There are three types of brain tumor from 233 patients including glioma (1426 slices), meningioma (708 slices) and pituitary (930 slices). All images were acquired from 233 patients in three planes: axial (994 images), sagittal (1025 images) and coronal (1045) image plane. The data is labelled as meningioma with 1, glioma with 2 and pituitary tumor with 3. In MR images, the experienced radiologists have designated the suspicious regions of interest (ROIs). The dataset was originally provided in matlab. mat format where each file stores a struct with a label which specify the type of tumor for a particular patient ID, brain image, image data in 512×512 unit 16 formats, vector storing the coordinates of the discrete points on tumor border, and a binary mask image with 1 indicating the tumor region. The images have an in-plane resolution of 512×512 with pixel size 0.49×0.49 mm^2. The thickness of slice is 6mm and gap of the slice is 1mm. Each patient contains approximately 1–6 images where most of patients have 1-3 images and very few patients have 4–6 images. The detail of CE-MRI data partitioning is detailed in section 2.4 and Table 1 below:

Table 1. Details of the CE-MRI dataset.

Tumor type	Number of Patients	Number of MR images	MRI view	Number of MR images
Meningioma	82	708	Axial	209
			Coronal	268
			Sagittal	231
Glioma	89	1426	Axial	494
			Coronal	437
			Sagittal	495
Pituitary	62	930	Axial	291
			Coronal	319
			Sagittal	320
Total	**233**	**3064**		**3064**

In this study, we divided data into train and test based on patient-ID, where 70% of patients data was used for training and 30% for testing purpose for tumor type based on single slice assigned to each tumor type as performed in the previous studies Cheng et al. Abiwinanda et. al.[23], Cheng et al. [22], Sajjad et al. [24], Zia et al. [25], Badža and Barjaktarović [26], Gumaei et al. [27], Swati et al. [4], and Huang et al. [28]. In order to overcome the problem of overfitting, 10-fold cross-validation was also performed.

2.2. *Feature extraction*

For improving the detection performance, the extraction of most relevant features is one of the most important steps. We extracting hybrid features as employed in our recent studies such as to detect prostate cancer by extracting combination of features [10], congestive heart failure with multimodal features [29], arrhythmia detection with hybrid features [30] proposed by Hussain et al. [10,29,30]. In this study, we computed traditional features based on morphological features and texture features, alongwith robust RICA features from multi-class brain tumor (pituitary, glioma and meningioma) and applied ML methods including SVM with its kernels and LDA. The RICA features based on their sparsity and robust to noise is more robust, and sigmoid nonlinearity imaging data. The brain tumor types are categorized into several factors such as type, location, texture of tumor. Thus, the traditional features may not provide detection performance better. On contrast, the RICA features seemed to be more appropriate to compute the multivariate information hidden in the brain tumor types. The traditional features extracted were of following categories:

2.2.1. Texture features

Texture feature have effectively utilized in solving classification related issue [31] especially to classify colon biopsies by employing microscopic image analysis for feature identification [32], Fractal analysis [33] proposed by Esgiar et al. [32,33]. Texture features are obtained from Gray-level co-occurrence matrix (GLCM). GLCM covers the spatial relationship of the Gray-level in an image. Any entry (i, j) th in co- occurrence matrix explain the occurrence of Gray-level i and j, their relative orientation Θ and their distance d. Commonly Θ correspond in four direction (0^0, 45^0, 90^0, $135^{0)}$). There are around 15 feature which obtained using GLCM which we studied as Angular second moment, Entropy, Correlation, Local Homogeneity, Shade, Variance, Average, Sum, Prominence, Difference Entropy, Sum Entropy, Difference variance Contrast, Sum variance and Information measure of correlation. The texture features extracted from brain tumor types are reflected in Table 2 below.

These features can be computed from GLCM matrix G, where x and y represent indices of rows and columns of matrix G. p_{xy} is the xy^{th} term of matrix G divided by the sum of elements. The term μ_x and μ_y are the mean, σ_x and σ_y are the standard deviation of *x*th row and *y*th column of matrix G.

Table 2. Texture Features.

Features	Formulas	Description
Contrast (t)	$$\sum_{x=1}^{K}\sum_{y=1}^{K}(x-y)^2 p_{xy}$$	It is used to measure the contract between current pixel and its neighbor.
Correlation (ρ)	$$\sum_{x=1}^{K}\sum_{y=1}^{K}\frac{(x-\mu_x)(y-\mu_y)p_{xy}}{\sigma_x\sigma_y}$$	It is used to measure the degree of correlation between current pixel and its neighbor.
Dissimilarity (Dis)	$$\sum_{x=1}^{K}\sum_{y=1}^{K}|x-y|p_{xy}$$	It is used to measure the difference in images.
Entropy	$$\sum_{x=1}^{K}\sum_{y=1}^{K}p_{xy}(-lnp_{xy})$$	It is used to get the encoded information from an image.
Energy (n)	$$\sum_{x=1}^{K}\sum_{y=1}^{K}p_{xy^2}$$	It is used to measure the uniformity of an image.
Homogeneity (h)	$$\sum_{x=1}^{K}\sum_{y=1}^{K}\frac{p_{xy}}{1+|x-y|}$$	It is used to calculate the spatial closeness of elements in G to the diagonal of the matrix.
Randomness (r)	$$-\sum_{x=1}^{K}\sum_{y=1}^{K}p_{xy}\,log_2 p_{xy}$$	It is used to measure the randomness of the elements of the GLCM.
Mean (μ)	$$\mu_x=\sum_{x=1}^{K}\sum_{y=1}^{K}x(p_{xy})$$ $$\mu_y=\sum_{x=1}^{K}\sum_{y=1}^{K}y(p_{xy})$$	This formula is used to calculate the sum of all values and P is the probability mass function.
Variance (σ²)	$$\sigma_x^2=\sum_{x=1}^{K}\sum_{y=1}^{K}(p_{xy})(x-\mu_x)^2$$ $$\sigma_y^2=\sum_{x=1}^{K}\sum_{y=1}^{K}(p_{xy})(y-\mu_y)^2$$	This equation is used to measure how far a set of numbers is spread out from their mean.
Standard Deviation (σ)	$$\sigma_x=\sqrt{\sigma_x^2}\ \&\ \sigma_y=\sqrt{\sigma_y^2}$$	It is used to quantify the amount of dispersion of different values of a data set.

2.2.2. Morphological features

Morphology of skins plays vital part in deciding either the tissues are malignant or oppositely normal. Morphological features give an approach to change over the image morphology values. These features are obtained from images through changing the morphology of image within set of quantitative values utilizes in classification and they have extensively been utilized as a part of classification [34] segmentation [35]. Morphological feature module (FEM) taking input in form of the binary batch also finds associated factors in the clusters. Researchers in the past extracted few morphological features

such as Perimeter (p), Eccentricity (y), Area (a), Convex Area (x), Euler Number (l), Orientation (e), Compactness (0), Length of major (m1), and Minor Axes (m2) etc. In this study, we computed the following morphological features as reflected in Table 3:

Table 3. Morphological Features.

Features	Formulas	Description
Area (A)	Total number of pixels in a region	Total count of pixels that a specific region of image contains
Perimeter (P)	Pixels at the boundary of an image	Total count of pixels at the boundary of the image
Solidity	$\dfrac{Area}{ConvexArea}$	To calculate the density of an object, ratio between area and full convex object.
Roundness	$\dfrac{4 \times \Pi \times Area}{(ConvexPerimeter)^2}$	This equation is used to illustrate the difference between line and circle from other region of image.
Convex Area	Total no of pixels in a convex image	It is used to count total no of pixels in convex image.
Convexity	$\dfrac{ConvexPerimeter}{Perimeter}$	This equation is used to calculate the perimeter ratio between object itself and convex full of object.
Compactness	$\dfrac{4 \times \Pi \times Area}{(Perimeter)^2}$	It is used to find the degree of deviation from a circle. This shows the ratio between the object areas with circle area.
Maximum Radius (Max_R)	MAX(DISTANCE(C(x,y), BOUNDARY(x,y)))	This formula is used to calculate the maximum distance from boundary of the image to the center of the image, x and y are two points on the image.
Minimum Radius (MIN_R)	MIN(DISTANCE(C(x,y), BOUNDARY(x,y)))	This formula is used to calculate the minimum distance from boundary of the image to the center of the image.
Euler Number (EUL_NO)	No of objects in region – No of holes in these objects	This formula provides the difference between effected and unaffected area of an image.
Standard Deviation	$\sqrt{\dfrac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$	It is used to calculate the contrast of an image.
Entropy	$\sum \left(p * \log_2(p)\right)^2$	This equation shows the statistical measure which can be used to get the texture of the image.
Eccentricity (ECT)	$\sqrt{\left(\dfrac{MAX_R - MIN_R}{MAX_R}\right)^2}$	This formula represents the ratio of distance between major axis and ellipse focal. Value can be 0-1.
Rectangularity	$\dfrac{Area}{MAX_R - MIN_R}$	This formula is used to identify the similarity of image shape with rectangular shape.
Elongation	$1 - \dfrac{MIN_R}{MAX_R}$	This formula is used to measure the length of the object.

2.2.3. Reconstruction Independent Component Analysis (RICA)

The RICA does not require any class label information because of its un-supervised nature. The ICA algorithm deficiencies were removed using RICA algorithm. The results yielded using RICA are more robust than ICA. This algorithm learns based on the sparse feature learning mechanism. The algorithm based on the sparse filter is capable to distinguish the various made natural signals, and these features can play a vital role in many of the ML techniques.

Consider an unlabeled data with input $\{y^{(i)}\}_{i=1}^n, y^{(i)} \in \mathbb{R}^m$, the optimization problem of standard ICA using optimization algorithms [36], kernel sparse representation [37] for estimating ICA [36,37] mathematically defied as:

$$\min_{Xn} \frac{1}{n} \sum_{i=1}^n h(Xy^i) \qquad (2.1)$$

$$Subject\ to\ \dots XX^U = I$$

Where $h(.)$ indicate nonlinear penalty function, $X \in S^{Lxm}$ is a matrix, L represent number of vectors and I is the identity matrix. Additionally, $XX^U = I$ is used for avoiding the vectors in X to become degenerate. A smooth penalty function is used to handle this situation as indicated below:

$$h(.)= log(cosh\ (.)\ \text{[38].} \qquad (2.2)$$

To completely learn the standard ICA, there are several orthonormality constraint which obstruct it. Subsequently, this drawback stops ICA from scaling to high dimensional data. To resolve this matter, the soft reconstruction cost is used in RICA. Thus, RICA after this replacement, can be characterized by following equation (2.3)

$$\min_{Xn} \frac{\lambda}{n} \sum_{i=1}^n \left(||X^U Xy^i - y^i||_2^2 + \sum_{i=1}^n \sum_{k=1}^l h(X_k y^i) \right) \quad (2.3)$$

Here parameter $\lambda > 0$ shows the tradeoff between reconstruction error and sparsity.

The penalty h can produce sparse representations only, but not invariant [38]. Thus, RICA using efficient overcomplete feature learning algorithms [39], building low level features using feature learning [40] studied by V Le et al. [39,40] swapped it by an extra L_2 pooling penalty, by promoting pooling features to cluster correlated features together. Furthermore, feature learning can be done using L_2 pooling. L_2 pooling using feature pooling [41], learning invariant features [42] studied by [41,42] is a two-layered network having square nonlinearity in the 1st layer $(.)^2$ and square root nonlinearity in

the 2nd layer $\sqrt{(.)}$ as reflected in equation (2.4)

$$h(Xy^i) = \sum_{k=1}^L \sqrt{\varepsilon + H_k \cdot \left((Xy^i) \odot (Xy^i) \right)} \quad (2.4)$$

Here H_k represents a row of spatial pooling matrix $H \in P^{L \times L}$ set to constant weights i.e.,1 for each element in matrix H, $\odot$ represents the element wise multiplication and $\varepsilon > 0$ is a small constant.

The sparse representation of the actual data can be represented using RICA. The following steps are used to compute the RICA features.
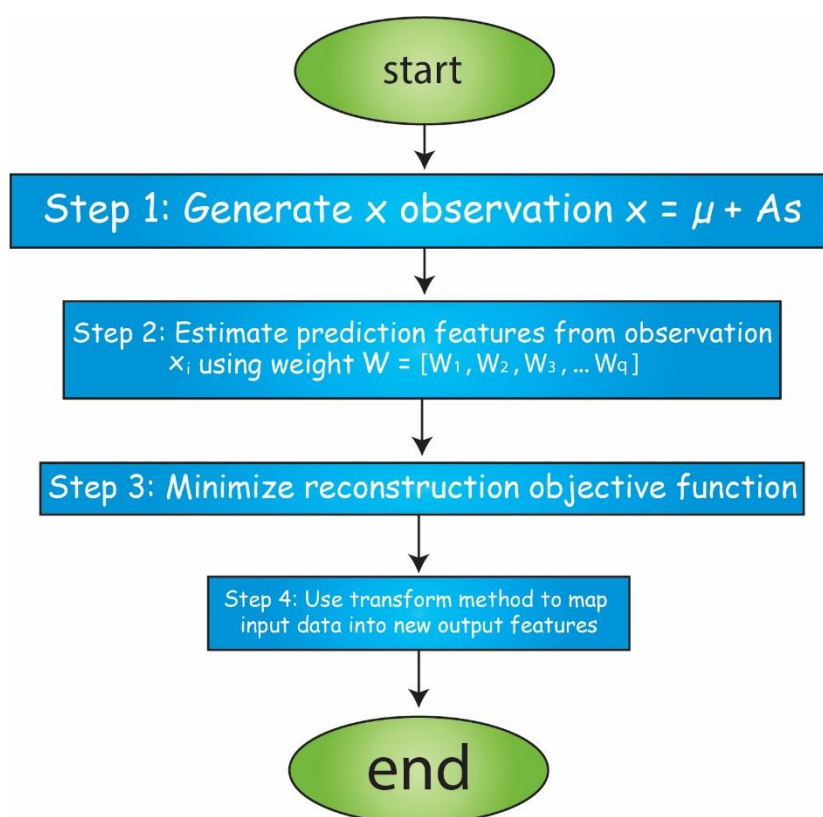
Figure 2. Steps in computing features based on RICA Algorithm.

The step-by-step procedure to compute the features using RICA algorithm is reflected in the Figure 2. The RICA feature model is obtained by applying RICA to the matrix of predictor data X containing p variables q number of features to extract from X. The RICA thus learns p by q matrix of transformation weights. The value of q can be less than or greater than the number of predictor variables to avoid from undercomplete or overcomplete feature representation. In this study, we choose q to 100 features and default values of alpha and gamma are set.

2.3. *Classification methods*

2.3.1. Support Vector Machine (SVM)

Vladimir Vapnik proposed SVM in 1979, which is a state of art algorithm used in different fields including medical diagnosis area [43], visual pattern recognition [44] and machine learning [45]. SVM is successfully used in many applications including text recognition, face expression recognition, emotion recognition, biometrics, and content-based image retrial etc. It constructs a hyperplane in the infinite dimensional space. The hyperplane helps to achieve the largest distance to any nearest training data point of any class. The lower generalization error can be obtained with the larger functional margin. To achieve this, SVM use the kernel trick. The linear and nonlinear separation with margin and slack variables in case of error examples are reflected in the Figure 3 (a, b) and Figure 4 (a, b).
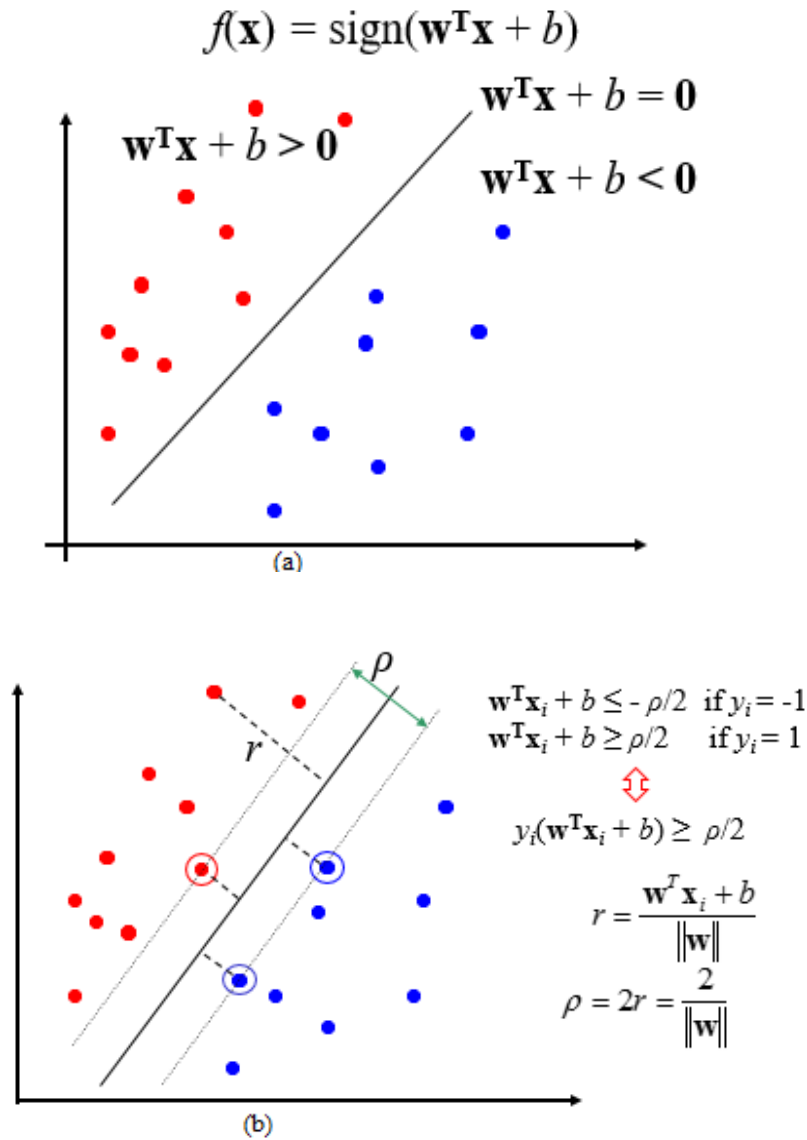
Figure 3. SVM (a) linear separation and (b) Margin.

Consider a hyperplane defined by x.w + b =0, where w is its normal. The data is linearly separated and is labelled as:

$$\{x_i, y_i\}, x_i \epsilon R^N d, y_i \epsilon \{-1, 1\}, i = 1, 2, \dots \dots, N \qquad (2.5)$$

Here y_i is the class label of two class SVM. To obtain the optimal boundary, the objective function is minimized with maximal margin i.e. $E = \|w\|^2$ subject to

$$x_i.w + b \geq 1 \; for \; y_i = +1$$

$$x_i.w + b \leq 1 \; for \; y_i = -1 \qquad (2.6)$$

Combining these into set of inequalities as

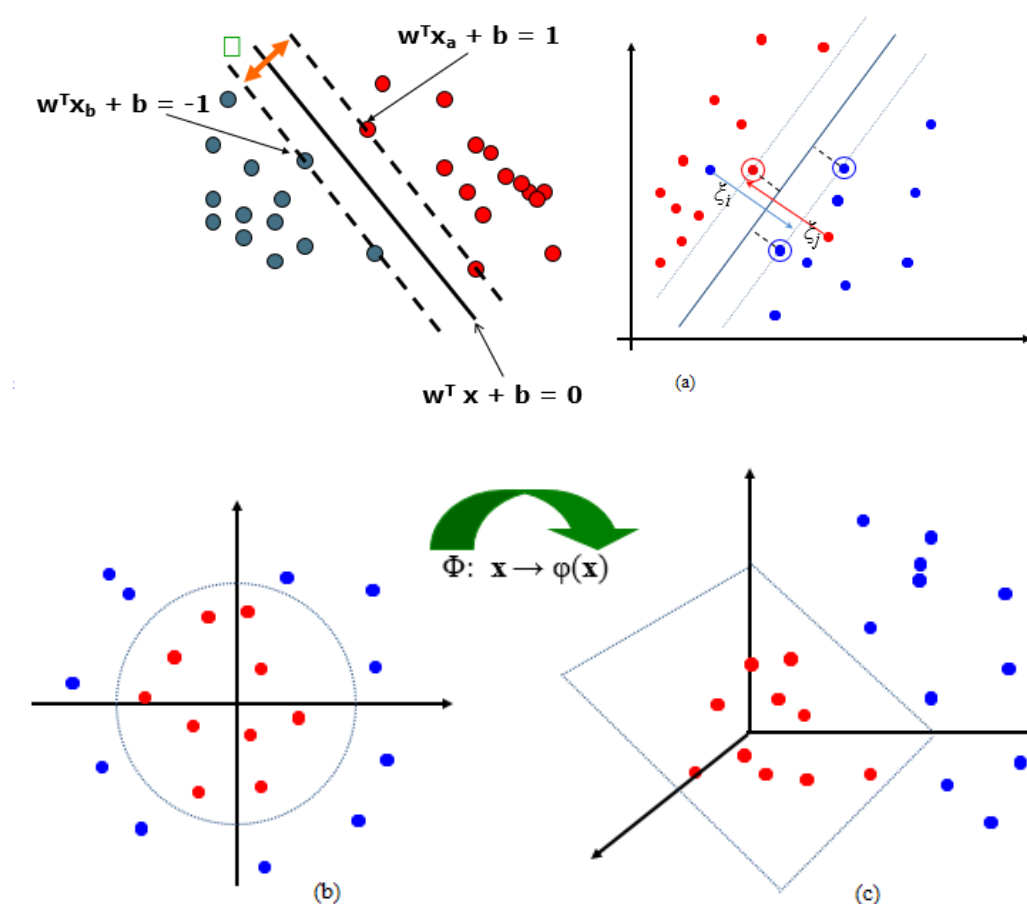$$(x_i.b + b)y_i \geq 1 \; for \; all \; i$$

Figure 4. (a) Error on margin using slack variable, (b, c) SVM non-linear separation.

Generally, the data is not linearly separable, in such cases a slack variable Ξ_i is used to indicate the amount of misclassification rate. Thus, new subjective function is then reformulated as:

$$E = \frac{1}{2}\|w\|^2 + C\sum_i L(\Xi_i) \tag{2.7}$$

Subject to

$$(x_i.b + b)y_i \geq 1 - \xi_i \ for \ all \ i$$

The first term on the right-hand side is the regularization term which gives the SVM an ability to generalize the sparse data well. The points which lie outside the margin are represented by the second term denoted by the empirical risk. The cost function is denoted by L, and hyper parameter is denoted by C, which shows a trade-off effect by minimizing the empirical risk against maximizing the margin. Linear-error cost function is most used because of its ability to detect the outliers. The dual formulation with

$$L(\Xi_i) = \Xi_i \ is$$

$$\alpha^* = max_\alpha\left(\sum_i \alpha_i + \sum_{i,j} \alpha_i\alpha_j \ y_iy_j \ x_ix_j\right) \tag{2.8}$$

Subject to

$$0 \leq \alpha_i \leq C \ and \ \sum_i \alpha_i y_j = 0$$

In which $\alpha = \{\alpha_1, \alpha_2, \alpha_3, \ldots \ldots \alpha_i, \}$is a set of Lagrange multipliers of the constraints in the primal optimization problem. The optimal decision boundary is now given by.

$$w_0 = \sum_i \alpha_i x_i y_i \qquad (2.9)$$

SVM for non-linearly separable data

The kernel function trick is recommended by the Muller et al. (2001) to deal the data with nonlinear separability. In this case the non-linear mapping from input space is made to higher dimensional feature space. The dot product between two vectors in the input space is expressed by dot product with some kernel functions in the feature space.
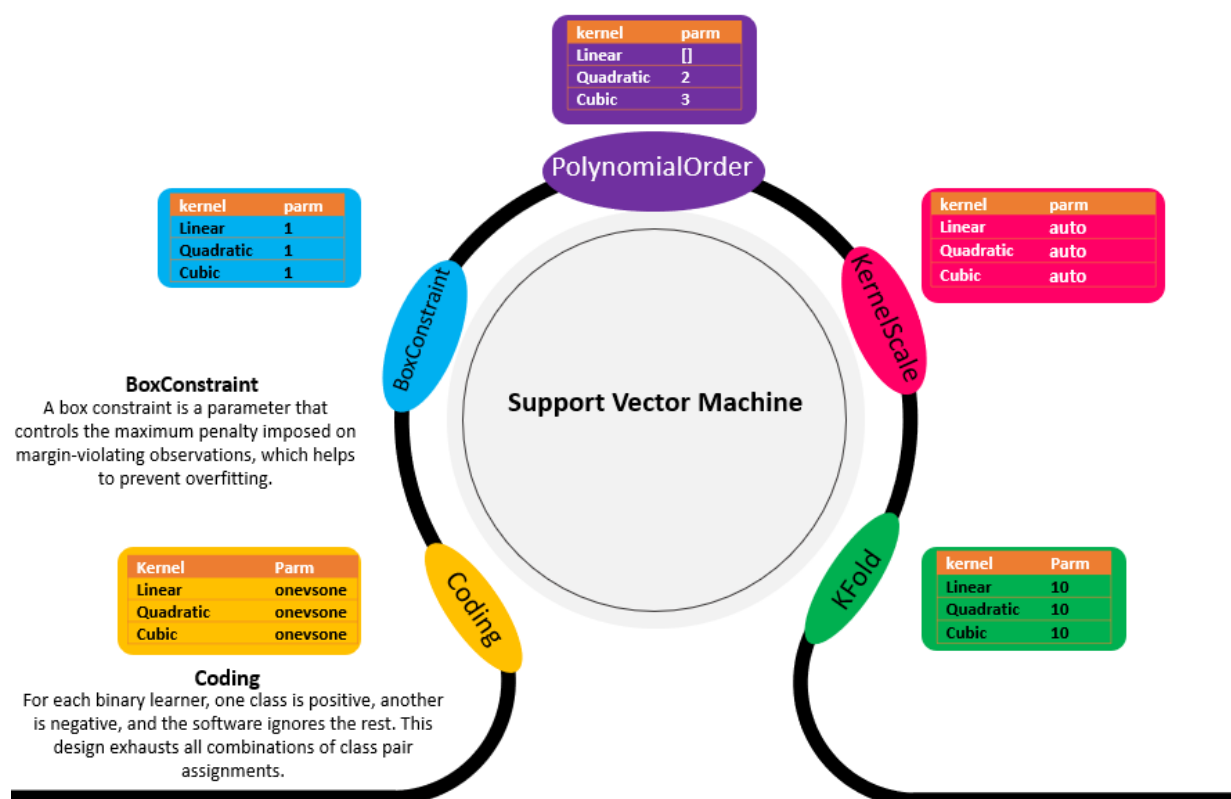

Figure 5. Parameter optimizations for SVM kernels.

The Figure 5 reflects the SVM kernels parameter optimization settings. The kernel parameters, box constraints, polynomial order (1,2,3) were used according to the default settings. As shown in the above figure, in this research work three SVM kernels (Linear, Quadratic, cubic) are used for the classification of Brain Tumor. All three SVM classifiers are trained with 10-Fold Cross-validation and Kernel Scale auto. Box Constant parameter is used to control the overfitting problem. SVM is a binary

classifier and to train on multi-class, Coding parameter oneVSone is used. In the oneVSone option, one class is treated as a positive, the other as a negative class, and all other classes are not used in training, this process repeated for all the class combinations.

The most used kernel functions are polynomial and radial base function (RBF). Mathematically, these are expressed as:

Types of Different Machine Learning Kernels with formulae

SVM Linear Kernel

$$K(x_i, y_i) = x_i.y_i + 1 \qquad (2.10)$$

SVM Quadratic Kernel

$$K(x_i, y_i) = (x_i.y_i + 1)^2 \qquad (2.11)$$

SVM Cubic Kernel

$$K(x_i, y_i) = (x_i.y_i + 1)^3 \qquad (2.12)$$

2.3.2. Linear Discriminant Analysis (LDA)

Belhumeur in 1997 [46] proposed LDA as one of the classical algorithms in the field of pattern recognition and artificial intelligence (AI). The main functionality of this algorithm is to project the high dimensional samples into low dimensional space to achieve the effect of extracting classification information and to compress the feature space dimension. LDA is successfully been employed in many of the applications such as Pathak et al. [47] applied this algorithm for removing the redundancy and inconsistency in the data. Moreover, LDA can be used for classification and dimensionality reduction, we used LDA for multi-class classification.

LDA is a simple method of classification using the generative methodology. It assumes that a Gaussian distribution is possible for each class and that every class has the same matrix of covariance. The LDA is a linear classification method with these assumptions. If they are by chance supportive of the actual data distribution, LDA is optimal in that it converges to the classifier of Bayes, when the number of data tends to infinitely (the parameter estimates, therefore, correspond to the real distribution parameters). In fact, LDA needs few computations to approximate the parameters of the classifier that amount to the estimation of the percentages and means plus the inversion of the matrix.). The LDA takes the generative method when presuming that a Gaussian distribution with probability density function generates the data of each class. The probability density function of x in population π_i is multivariate natural with mean variable μ_i and variance-covariance matrix. The formula for this usual function of probability density is:

$$p_{X|Y=y}(x|Y = y) = \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma_y|^{\frac{1}{2}}} \exp(-\frac{1}{2}(x - \mu_y)^T \Sigma_y^{-1}(x - \mu_y)) \quad (2.13)$$

And that the covariance matrix Σ_y for all labels is the same:

$$\forall y \in \mathcal{Y}, \Sigma_y = \Sigma \qquad (2.14)$$

They approximate the parameters as follows. The previous probabilities are essentially the data point fractions of each group:

$$\forall y \in \mathcal{Y}, P(Y = y) = \frac{N_y}{N}, \text{ with } N_y = \sum_{i=1}^{N} \mathbf{1}_{y_i=y} \qquad (2.15)$$

The Gaussians' means are estimated by the means of the sample.

$$\forall y \in \mathcal{Y}, \boldsymbol{\mu}_k = \frac{1}{N_y} \sum_{y_i=y} x_i \qquad (2.16)$$

And the matrix for covariance by

$$\boldsymbol{\Sigma} = \frac{1}{N - |\mathcal{Y}|} \sum_{y \in \mathcal{Y}} \cdot \sum_{y_i=y} (x_i - \boldsymbol{\mu}_y)(x_i - \boldsymbol{\mu}_y)^T \qquad (2.17)$$

2.4. *Training/Testing Data Formulation*

For training/testing data formulation, the Jack-knife 10-fold cross validation (CV) was used. The performance was evaluated using the similar metrics to detect brain tumor by applying adaptive spatial pooling methods [21], margin information and learning distance metric [48], bag-of-visual word representation methods [49], spatial layout information based methods [50] as employed and tested by [21,48–50], and CE-MRI data of 233 patients was randomly divided into 10 subsets of equal size. We also ensured that there is no overlap and equal ratios of the different type of tumors in the 10 subsets for the CE-MRI datasets. The division according to the patients ensure that images from same patient did not exist simultaneously in the training and testing set. Using 10-fold cross validation, the data is partitioned into 10 folds and 9 folds participate in training and remaining folds in testing. The samples in the test fold are purely unseen. The entire process is repeated 10 times.

K-fold Cross-validation is an effective preventative measure against overfitting. Thus, to tune the model, the dataset is split into multiple train-test bins. Using k-fold CV, the dataset is divided into k-folds. For model training, k-1 folds are involved, and rest of the folds are used for model testing. Moreover, k-fold method is helpful for fine-tuning the hyperparameters with the given original training dataset in order to determine that how the outcome of ML model could be generalized. The k-fold cross validation procedure is reflected in Figure 6 below.
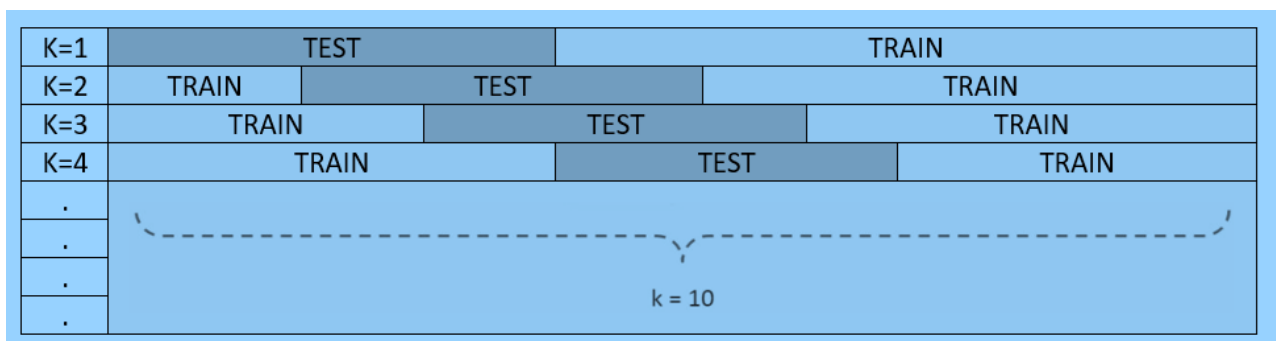

Figure 6. K-fold Cross-validation.

3.   Results and discussions

The researchers are devising automated tools to improve the prediction of brain tumor types because of the multivariate characteristics of the tumor types. Extracting the most relevant and appropriate feature is still a challenging task. In this study, we first extracted the traditional texture and morphological features from brain tumor types and computed the performance using the machine learning classification techniques such as LDA, SVM with linear, quadratic, cubic and cosine kernels. We then extracted the RICA based features based on the multivariate characteristics. These features are then used as input to these classifiers for multi-class approach. The results reveal that proposed feature extraction approach using SVM cubic yielded more appropriate results to predict the tumor types.

Table 4 shows the results of AI multiclass brain tumor types (Glioma, meningioma, pituitary) classification of texture and morphological features. The classifiers LDA and SVM with its kernel yielded moderate performance. Specifically, SVM quadratic classifiers yielded best performance with accuracy (93.11%), AUC (0.8928) followed by SVM cubic with accuracy (93.04%) and AUC (0.8895) to predict the pituitary from multiclass. The other performance metrics are reflected in the Table 3.

Table 4. Multiclass classification of Brain tumor types (Glioma, meningioma, pituitary) by extracting texture + morphological based features and employing machine learning techniques.

Class	Sens.	Spec.	PPV	NPV	FPR	Acc.	AUC
LDA							
Glioma	100%	16.09%	47.79%	100%	0.839	52.54%	0.5804
meningioma		100%		100%	0	100%	
Pituitary	39.78%	100%	100%	94.31%	0	94.52%	0.6989
SVM Linear							
Glioma	100%	54.42%	60.67%	100%	0.455	73.24%	0.7720
meningioma		100%		100%	0	100%	
Pituitary	75.48%	100%	100%	89.67%	0	92.16%	0.8774
SVM Quadratic							
Glioma	100%	62.30%	63.95%	100%	0.376	77.42%	0.8115
meningioma	52.28%	99.79%	93.02%	97.54%	0.0020	97.42%	0.7604
Pituitary	78.57%	100%	100%	90.78%	0	93.11%	0.8928
SVM Cubic							
Glioma	100%	71.25%	64.91%	100%	0.2875	81.23%	0.8562
meningioma	46.80%	98.56%	83.89%	92.04%	0.014	91.41%	0.7268
Pituitary	77.90%	100%	100%	90.79%	0	93.04%	0.8895
SVM Cosine							
Glioma	100%	75.50%	67.42%	100%	0.2449	83.74%	0.8775
meningioma	47.61%	99.29%	90.45%	93.08%	0.007	92.92%	0.7345
Pituitary	71.79%	100%	100%	85.71%	0	89.95%	0.8589

Table 5 reflect the multi-class classification results of brain tumor types (meningioma, Glioma, pituitary) based on the RICA features. The classifiers LDA and SVM with its kernel yielded highest performance. Specifically, SVM cubic classifiers yielded best performance with accuracy (99.34%), AUC (0.9892) followed by SVM quadratic with accuracy (98.10%) and AUC (0.9699) to predict the

pituitary from multiclass. To predict the meningioma from multiclass, SVM cubic yielded an accuracy (96.96%), AUC (0.9348) and to predict glioma from multiclass and accuracy (95.88%), AUC (0.9635) was obtained. The highest multi-class prediction with other classifiers was obtained by LDA followed by SVM linear, and SVM cosine.

Table 5. Multiclass classification of Brain tumor types by extracting RICA based features and utilizing ML techniques.

Class	Sensitivity	Specificity	PPV	NPV	FPR	Accuracy	AUC
LDA							
Glioma	100%	82.38%	78.95%	100%	0.1761	89.39%	0.9119
meningioma	69.18%	99.32%	95.66%	93.75%	0.0067	93.99%	0.8425
Pituitary	89.07	100%	100%	95.24%	0	96.57%	0.9453
SVM Linear							
Glioma	100%	84.64%	81.45%	100%	0.153	90.83%	0.9232
meningioma	71.50%	99.60%	97.46%	94.26%	0.004	94.68%	0.8555
Pituitary	88.32%	100%	100%	94.63%	0	96.18%	0.9416
SVM Quadratic							
Glioma	100%	91.18%	89.53%	100%	0.088	94.97%	0.9559
meningioma	84.51%	99.63%	98.31%	96.20%	0.0036	96.57%	0.9207
Pituitary	93.89%	100%	100%	97.31%	0	98.10%	0.9699
SVM Cubic							
Glioma	100%	92.27%	91.38%	100%	0.072	95.88%	0.9635
meningioma	87.34%	99.62%	98.47%	96.60%	0.0038	96.96%	0.9348
Pituitary	97.78%	100%	100%	99.07%	0	99.34%	0.9892
SVM Cosine							
Glioma	100%	87.51%	84.01%	100%	0.1248	92.46%	0.9375
meningioma	72.98%	99.76%	98.66%	93.72%	0.0024	94.45%	0.8636
Pituitary	87.27%	100%	100%	94.14%	0	95.58%	0.9336

The Figure 7 (a–e) reflects the Multi-class distribution of glioma (1426 slices), meningioma (708 slices) and pituitary (930 slices). From figure 7 (d) using SVM cubic, out of 1426 glioma, there were 1337 were predicted a glioma, 113 as meningioma and 13 as pituitary. From 708 meningioma, after prediction, there were 84 predicted as glioma, 580 as meningioma and 9 as pituitary. From 930 pituitary, there were 5 predicted as glioma, 15 as meningioma and 908 as pituitary. The distribution using other classifiers is reflected in the Figure 7 (a–e).

The researchers extracted various features extraction approaches using ML and DL methods to detect the binary class classification of brain tumor types. The highest performance based on the overall accuracy was obtained by [22] 91.28%, [51] 90.89%, [52] 86.56%, and [53] 84.19%. With the multi-class classification, the LDA yielded accuracy for pituitary (96.48%), meningioma (93.89%) and glioma (89.39%). Using SVM linear, the accuracy to detect pituitary was yielded (96.28%), an accuracy (94.45%) was obtained to detect meningioma, while to detect the glioma, and accuracy (90.76%) was yielded. By employing the quadratic kernel, the highest detection was obtained to detect pituitary with accuracy (98.07%), followed by accuracy (96.18%) to predict meningioma and accuracy (94.35%) to detect glioma.
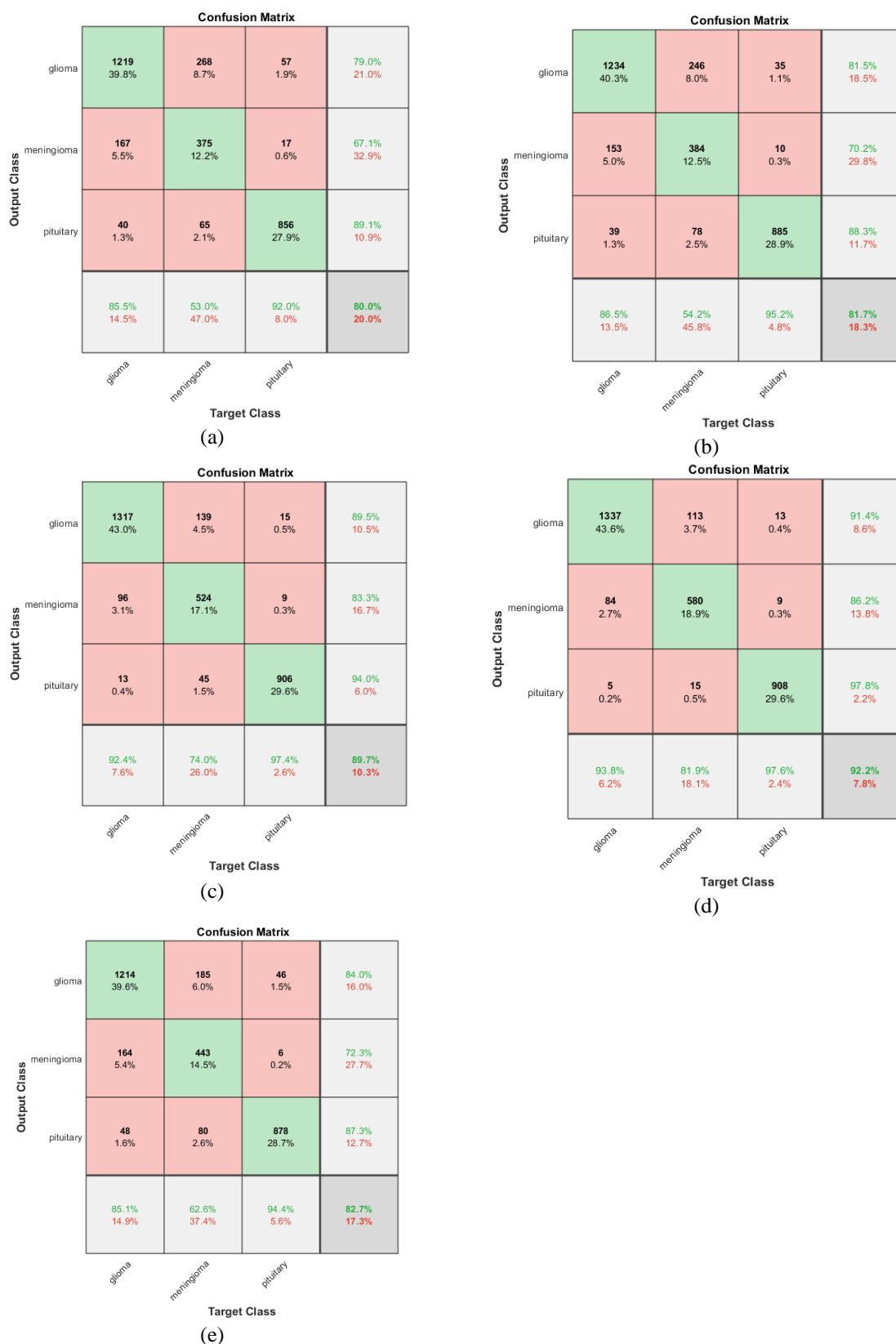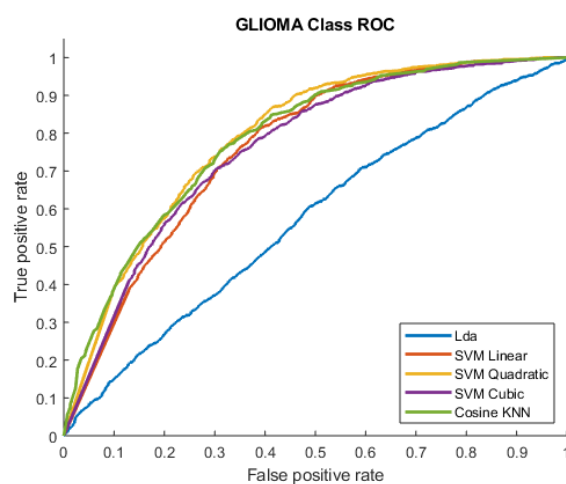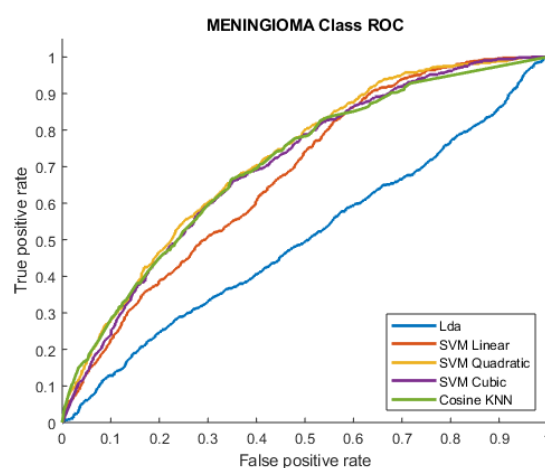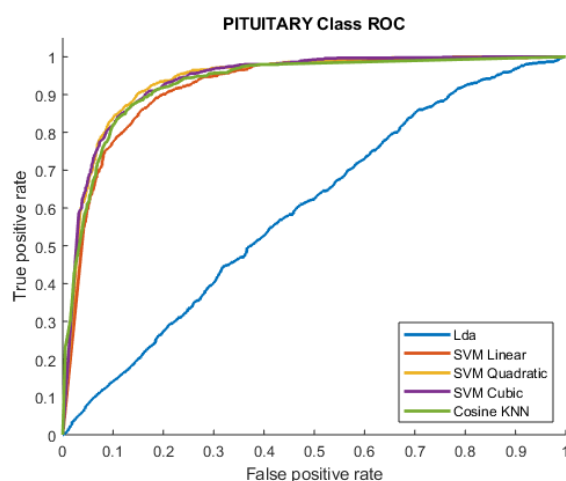
Figure 7. Confusion matrix to detect Multi-class brain tumor types (Glioma, Pituitary, Meningioma) by Computing RICA features and using machine learning techniques a) LDA, b) SVM linear, c) SVM quadratic, d) SVM cubic, e) SVM cosine.
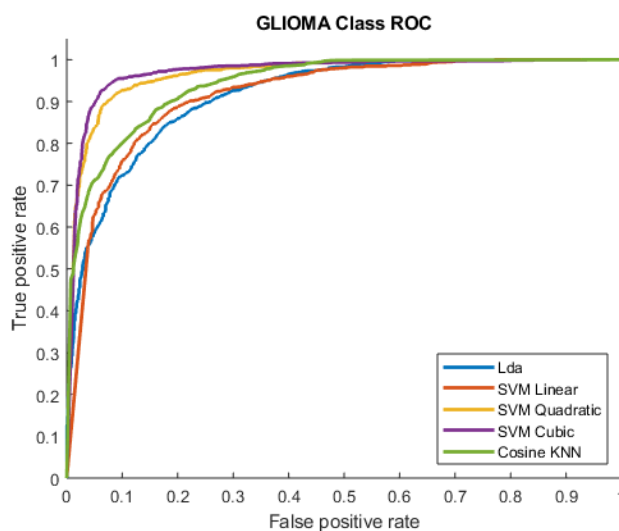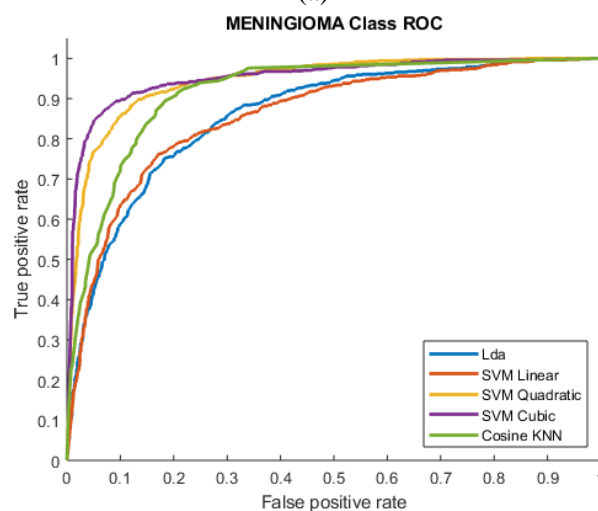
Figure 8. Area under the receiver operating curve (AUC) to detect Multi-class Brain tumor types a) Glioma, b) Meningioma, c) Pituitary using Robust Machine learning classifiers and extracting texture + morphological features.
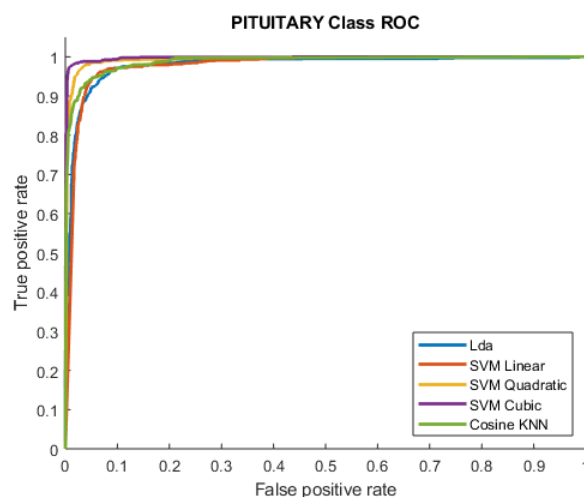
Figure 9. Area under the receiver operating curve (AUC) to detect Multi-class Brain tumor types a) Glioma, b) Meningioma, c) Pituitary using Robust Machine learning classifiers by extracting RICA features.

The Figure 8 (a–c) shows the multi-class separation in the form of the area under the receiver operating curve based on texture + morphological features extracted and employing machine learning techniques. The highest separation was obtained with AUC (0.8928) to detect pituitary using SVM quadratic followed by AUC (0.8895) to detect pituitary using SVM cubic.

The Figure 9 (a-c) reflect the Multi-class separation to distinguish a) Glioma, b) Meningioma, and c) Pituitary by computing RICA features and utilizing robust machine learning techniques. To detect the Glioma, the separation with AUC was obtained using LDA (0.9119), SVM Linear (0.9232), SVM quadratic (0.9559), SVM cubic (0.9635), and SVM cosine (0.9375). To detect the meningioma, the separation with AUC was obtained using LDA (0.8425), SVM Linear (0.8555), SVM quadratic (0.9207), SVM cubic (0.9348), and SVM cosine (0.8636). To detect the pituitary, the separation with AUC was obtained using LDA (0.9453), SVM Linear (0.9416), SVM quadratic (0.9699), SVM cubic (0.9892), and SVM cosine (0.9336).

The Table 6 presents the findings of different hand-crafted features techniques alongwith machine learning techniques to classify the brain tumor from normal and between brain tumor types using similar dataset and different datasets. Using LDA, the highest detection performance was obtained to detect pituitary with accuracy (96.57%), AUC (0.9453) followed by meningioma and glioma. Using SVM linear kernel, the highest detection performance was obtained to detect pituitary with accuracy (96.18%), AUC (0.9416). Using SVM quadratic kernel, the highest detection performance was obtained to detect pituitary with accuracy (98.10%) and AUC (0.9699). Likewise, using SVM cubic, the highest detection performance to detect the pituitary was obtained with accuracy (99.34%), AUC (0.9892). Moreover, using SVM cosine, to detect pituitary an accuracy (95.58%) and AUC (0.9336) was yielded.

To extract the diagnostic information from MR images, researchers employed several image analysis techniques using tissue characterization methods [57], texture text objects and intracranial brain tumor detection [58], and tissue characterization and intracranial brain tumor detection [59] detailed in [57–59]. The texture analysis and pattern recognition techniques were employed in these studies to characterize the types of brain tumor. Recently, [60] employed SVM to classify the gliomas and meningiomas and obtained 95% overall accuracy to distinguish these types. Moreover, [57] employed k-nearest neighbor and discriminant analysis to distinguish between oedematous and brain tumor tissues by achieving a maximum accuracy of 95%. Recently, several studies applied MR spectroscopic features such as long echo proton MRs signals [61], short echo time [62], tumor grading [63], short time multicenter study [64], and short echo metabolic patterns [65] as described in [61–65] or combination of spectroscopic and texture features to distinguish between various brain tumor types by achieving a maximum accuracy of 99% [64]. Moreover, authors with benchmark with similar dataset [37] extracted hand-crafted features by applying machine learning techniques and deep convolutional neural network methods obtained performance in terms of overall accuracy [7] 98%, [54] 96.4%, [66] 80%, [52] 86.56%, [25] 85.69%, [22] 91.28%, [55] 94.2%, [53] 91.43%, and [56] 96.67%. In the present study, we used MRI brain tumor types dataset originally provided by Cheng et al. which is used in his studies [21,22]. We compared the results with similar dataset used by other researchers such as Abiwinanda et. al. [23], Cheng et al. [22], Sajjad et al. [24], Zia et al. [25], Badža and Barjaktarović [26], Gumaei et al. [27], Swati et al. [4], and Huang et al. [28] as reflected in Table 6.

Table 6. Literature review of existing techniques and methods and with similar dataset.

Author	Feature/Methods	Performance
Machhale et al. [7]	SVM-KNN	Sensitivity: 100% Specificity: 93.75% Accuracy: 98%
Zacharaki et al. [54]	Cross-Validation Using different Classifiers (LDA,k-NN,SVM)	Sensitivity: 75% Specificity: 100% Accuracy: 96.4%
Badža and Barjaktarović [26]	CNN	Accuracy 95.40%
Gumaei et al. [27]	Regularized extreme learning machine (RELM)	Accuracy 94.23%
Swati et al. [4]	Automatic content-based image retrieval (CBIR) system	Average precision 96.13%
Huang et al. [28]	convolutional neural network based on complex networks (CNNBCN)	Accuracy 95.49%
Afshar et al. [52]	Capsule Network Method	Accuracy: 86.56%
Zia et al. [25]	Window Based Image Cropping	Sensitivity:86.26% Specificity:90.90% Accuracy: 85.69%
Sajjad et al. [24]	CNN with data augmentation	Sensitivity:88.41% Secificity:96.12% Accuracy: 94.58%
Cheng et al. [22]	Feature extraction methods: Intensity Histogram GLCM BOW Classification Methods: SVM SRC KNN	Accuracy:91.28%
Abiwinanda et. al. [23]	CNN	Accuracy: 84.19%
Anaraki et al. [55]	Genetic Algorithms	Accuracy: 94.2%
Paul et al. [53]	NN	Accuracy: 91.43%
Sachdeva et al. [56]	Segmentation and Feature extraction	Highest accuracy 96.67%
This work	RICA Based Features SVM Cubic with Multiclass classification 1) Pituitary 2) Meningioma 3) Glioma	1) Accuracy: 99.34%, AUC: 0.9892 2) Accuracy: 96.96%, AUC: 0.9348 3) Accuracy: 95.88%, AUC: 0.9635

The authors who used the similar database includes Abiwinanda et. al., Sajjad et al., Anaraki et al., Cheng et al., Swati et al., and Gumaei et al. to predict the brain tumor types such as Glioma, Meningioma and Pituitary. Abiwinanda et. al. [23] trained the CNN to predict the three most common types of brain tumor i.e. Glioma, Meningioma and Pituitary. They implemented the simple CNN architecture i.e., max-pooling, convolution, and flattening layers followed by a full connection from one hidden layer. The CNN was trained on similar dataset consisting of 3064 T-1 weighted CE-MRI images publicly available Cheng et al. [22] yielded a training accuracy of 98.51% and validation accuracy of 84.19% at best. The results are compared with similar dataset by employing the region-based segmentation algorithms yielded accuracies ranged between 71.39% to 94.68%. Sajjad et al. [24] applied CNN with and without data augmentation methods to detect the brain tumor types such as Glioma, Meningioma and Pituitary. With the original dataset, the highest performance was obtained with sensitivity (88.41%), specificity (96.12%) and accuracy (94.58%). Anaraki et al. [55] applied CNN and genetic algorithms to classify the MRI brain tumor grades types. The highest classification accuracy of 94.2% was yielded to classify brain tumor types such as Glioma, Meningioma and Pituitary tumor with improved results as computed by Paul et al. by employing Vanilla preprocessing with shallow CNN to distinguish the Glioma, Meningioma and Pituitary tumor types. Cheng et al. [22] classified the three brain tumor types such as Glioma, Meningioma and Pituitary. The classification performance was evaluated with three feature extraction methods namely gray level co-occurrence matrix (GLCM), intensity histogram and bag-of-words model Enhanced Performance of Brain Tumor Classification via Tumor Region Augmentation and Partition. The improved performacne are reflected in Table 6.

In many imaging pathologies, the texture properties along with morphological imaging features played a vital role in prediction. This may be since most of these pathologies may contain the hidden information can be best extracted from these texture and shape properties. Due to the heterogenous characteristics, aggressive nature and involvement of several factors, the brain tumor is categorized into different types (i.e. glioma, meningioma and pituitary etc.). Researchers are developing various automated tools to improve the prediction. The results yielded by extracting texture and morphological features reveal that some machine learning algorithms provided higher sensitivity while some other provided higher specificity. It can be inferred that these features still cannot be better fit to better predict the brain tumor types based on these heterogenous characteristics. While extracting RICA features improved both specificity and sensitivity substantially using SVM quadratic and cubic kernels. Thus, RICA feature characteristics may better tailor to distinguish these multiclass brain tumor types and hence improved the prediction performance.

4.   Conclusion

In this study, we used the RICA based advanced feature extraction methods from MRI scans of multi-class brain tumor types of patients. The brain tumor types properly classification is of much significance to correctly treat the brain tumor. The proposed multiclass approach yielded the highest detection rate to detect pituitary followed by meningioma and glioma type. The results revealed that proposed approach based on RICA features from brain tumor types of MRIs will be very helpful for early detection of tumor type and to treat the patients to improve the survival rate.

Limitation and future recommendations

In this study, we used multi-class classification between few brain tumors types. The data is lacking the description of distribution of each type of patient, which we will address in future. In future, we will also extend the work with other types of brain tumor and larger datasets along with more feature extraction methods. We will also employ this model for other type of medical images such as ultrasonography (ultrasound), radiography (X-ray), dermoscopic, endoscopic and histology images along with demographic information and tumor staging. Machine learning based on the feature extraction approach is hot topic of research due to less computational time as compared to the deep learning which require more computational resources. The researchers are developing different feature extraction approaches in order to improve the detection performance. We will extract more relevant features for further improving the machine learning (i.e. non-deep learning) classification results. We will also compute and compare the results of Machine learning methods using feature extraction approach with the deep convolutional neural network methods with optimization of parameters.

Conflicts of interest

The authors declare that they have no conflict of interest.

Ethical approval

Not Applicable Data were obtained from a publicly available, deidentified dataset. For this type of study formal consent is not required https://github.com/chengjun583/brainTumorRetrieval

References

1. S. B. Gaikwad, M. S. Joshi, Brain tumor classification using principal component analysis and probabilistic neural network, *Int. J. Comput. Appl.*, **120** (2015), 5–9.
2. Q. T. Ostrom, G. Cioffi, H. Gittleman, N. Patil, K. Waite, C. Kruchko, et al., CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2012–2016, *Neuro. Oncol.*, **21** (2019), v1–v100.
3. D. N. Louis, A. Perry, G. Reifenberger, A. von Deimling, D. Figarella-Branger, W. K. Cavenee, et al., The 2016 World Health Organization Classification of tumors of the central nervous system: A summary, *Acta Neuropathol.*, **131** (2016), 803–820.
4. Z. N. K. Swati, Q. Zhao, M. Kabir, F. Ali, Z. Ali, S. Ahmed, et al., Content-Based brain tumor retrieval for MR images using transfer learning, *IEEE Access*, **7** (2019), 17809–17822.
5. S. Pereira, R. Meier, V. Alves, M. Reyes, C. A. Silva, Automatic brain tumor grading from mri data using convolutional neural networks and quality assessment, in: *Understanding and interpreting machine learning in medical image computing applications*, Springer, Cham., 2018, 106–114.
6. S. Deepak, P. M. Ameer, Brain tumor classification using deep CNN features via transfer learning, *Comput. Biol. Med.*, **111** (2019), 103345.
7. K. Machhale, H. B. Nandpuru, V. Kapur, L. Kosta, MRI brain cancer classification using hybrid classifier (SVM-KNN), in: *2015 Int. Conf. Ind. Instrum. Control*, IEEE, 2015, 60–65.
8. S. Rathore, M. Hussain, M. Aksam Iftikhar, A. Jalil, Ensemble classification of colon biopsy

images based on information rich hybrid features, *Comput. Biol. Med.*, **47** (2014), 76–92.

9. S. Rathore, M. Hussain, A. Khan, Automated colon cancer detection using hybrid of novel geometric features and some traditional features, *Comput. Biol. Med.*, **65** (2015) 279–296.

10. L. Hussain, A. Ahmed, S. Saeed, S. Rathore, I. A. Awan, S. A. Shah, et al., Prostate cancer detection using machine learning techniques by employing combination of features extracting strategies, *Cancer Biomarkers*, **21** (2018), 393–413.

11. Y. Asim, B. Raza, A. Kamran, M. Saima, A.K. Malik, S. Rathore, et al., A multi-modal, multi-atlas-based approach for Alzheimer detection via machine learning, *Int. J. Imag. Sys. Tech.*, **28** (2018), 113–123.

12. A. Islam, M. F. Hossain, C. Saha, A new hybrid approach for brain tumor classification using BWT-KSVM, in: *2017 4th Int. Conf. Adv. Electr. Eng.*, IEEE, 2017, 241–246.

13. P. P. Rebouças Filho, E. de S. Rebouças, L. B. Marinho, R. M. Sarmento, J. M. R. S. Tavares, V. H. C. de Albuquerque, Analysis of human tissue densities: A new approach to extract features from medical images, *Pattern Recognit. Lett.*, **94** (2017), 211–218.

14. B. Dhruv, N. Mittal, M. Modi, Study of Haralick's and GLCM texture analysis on 3D medical images, *Int. J. Neurosci.*, **129** (2019), 350–362.

15. Q. Zheng, H. Li, B. Fan, S. Wu, J. Xu, Integrating support vector machine and graph cuts for medical image segmentation, *J. Vis. Commun. Image Represent*, **55** (2018), 157–165.

16. S. A. Taie, W. Ghonaim, Title CSO-based algorithm with support vector machine for brain tumor's disease diagnosis, in: *2017 IEEE Int. Conf. Pervasive Comput. Commun. Work. (PerCom Work.)*, IEEE, 2017, 183–187.

17. M. K. Abd-Ellah, A. I. Awad, A. A. M. Khalaf, H. F. A. Hamed, Classification of brain tumor MRIs using a kernel support vector machine, in: *International Conference on Well-Being in the Information Society,* Springer, Cham. 2016,151–160.

18. H. Alquran, I. A. Qasmieh, A. M. Alqudah, S. Alhammouri, E. Alawneh, A. Abughazaleh, et al., The melanoma skin cancer detection and classification using support vector machine, in: *2017 IEEE Jordan Conf. Appl. Electr. Eng. Comput. Technol.*, IEEE, 2017, 1–5.

19. S. Wang, S. Du, A. Atangana, A. Liu, Z. Lu, Application of stationary wavelet entropy in pathological brain detection, *Multimed. Tools Appl.*, **77** (2018), 3701–3714.

20. Y. Zhang, J. Yang, S. Wang, Z. Dong, P. Phillips, Pathological brain detection in MRI scanning via Hu moment invariants and machine learning, *J. Exp. Theor. Artif. Intell.,* **29** (2017), 299–312.

21. J. Cheng, W. Yang, M. Huang, W. Huang, J. Jiang, Y. Zhou, et al., Retrieval of brain tumors by adaptive spatial pooling and fisher vector representation, *PLoS One*, **11** (2016), e0157112.

22. J. Cheng, W. Huang, S. Cao, R. Yang, W. Yang, Z. Yun, et al., Enhanced performance of brain tumor classification via tumor region augmentation and partition, *PLoS One.,* **10** (2015), e0140381.

23. N. Abiwinanda, M. Hanif, S. T. Hesaputra, A. Handayani, T. R. Mengko, Brain tumor classification using convolutional neural network, in: *World Congr. Med. Phys. Biomed. Eng. 2018*, Springer, Singapore, 2019, 183–189.

24. M. Sajjad, S. Khan, K. Muhammad, W. Wu, A. Ullah, S. W. Baik, Multi-grade brain tumor classification using deep CNN with extensive data augmentation, *J. Comput. Sci.*, **30** (2019), 174–182.

25. R. Zia, P. Akhtar, A. Aziz, A new rectangular window based image cropping method for generalization of brain neoplasm classification systems, *Int. J. Imag. Syst. Technol.*, **28** (2018), 153–162.

26. M. M. Badža, M. Č. Barjaktarović, Classification of brain tumors from MRI images using a convolutional neural network, *Appl. Sci.*, **10** (2020), 1999.

27. A. Gumaei, M. M. Hassan, M. R. Hassan, A. Alelaiwi, G. Fortino, A hybrid feature extraction method with regularized extreme learning machine for brain tumor classification, *IEEE Access*, **7** (2019), 36266–36273.

28. Z. Huang, X. Du, L. Chen, Y. Li, M. Liu, Y. Chou, et al., Convolutional neural network based on complex networks for brain tumor image classification with a modified activation function, *IEEE Access*, **8** (2020), 89281–89290.

29. L. Hussain, I. A. Awan, W. Aziz, S. Saeed, A. Ali, F. Zeeshan, et al., Detecting congestive heart failure by extracting multimodal features and employing machine learning techniques, *Biomed. Res. Int.*, **2020** (2020), 1–19.

30. L. Hussain, W. Aziz, S. Saeed, I. A. Awan, A. A. Abbasi, N. Maroof, Arrhythmia detection by extracting hybrid features based on refined Fuzzy entropy (FuzEn) approach and employing machine learning techniques, *Waves Rand. Complex Media.*, **30** (2020), 656–686.

31. D. S. Guru, Y. H. Sharath, S. Manjunath, Texture features and KNN in classification of flower images, *Int. J. Comput. Appl.*, (2010), 21–29.

32. A. N. Esgiar, R. N. Naguib, B. S. Sharif, M. K. Bennett, A. Murray, Microscopic image analysis for quantitative measurement and feature identification of normal and cancerous colonic mucosa, *IEEE Trans. Inf. Technol. Biomed.*, **2** (1998), 197–203.

33. A. N. Esgiar, R. N. G. Naguib, B. S. Sharif, M. K. Bennett, A. Murray, Fractal analysis in the detection of colonic cancer images, *IEEE Trans. Inf. Technol. Biomed.*, **6** (2002), 54–58.

34. M. Masseroli, A. Bollea, G. Forloni, Quantitative morphology and shape classification of neurons by computerized image analysis, *Comput. Methods Programs Biomed.*, (1993), 89–99.

35. Y. M. Li, X. P. Zeng, A new strategy for urinary sediment segmentation based on wavelet, morphology and combination method, *Comput. Methods Programs Biomed.,* **84** (2006), 162–173.

36. A. Hyvärinen, E. Oja, Independent component analysis: Algorithms and applications, *Neural Networks*, **13** (2000), 411–430.

37. Y. Xiao, Z. Zhu, Y. Zhao, Kernel reconstruction ICA for sparse representation, *IEEE Trans. Neural Networks Learn. Syst.*, **26** (2015), 1222–1232.

38. J. Hurri, P. O. Hoyer, Natural Image Statistics, *A probabilistic approach to early computational vision*, Springer Science & Business Media, **39** (2009).

39. Q. V. Le, A. Karpenko, J. Ngiam, A. Y. Ng, ICA with reconstruction cost for efficient overcomplete feature learning, *Adv. Neural. Inform. Process Syst.,* **24** (2011), 1017–1025.

40. Q. V. Le, M. A. Ranzato, M. Devin, G. S. Corrado, A. Y. Ng, Building high-level features using large scale unsupervised learning, In *2013 IEEE international conference on acoustics, speech and signal processing*, IEEE, (2013), 8595–8598

41. Y. Boureau, A theoretical analysis of feature pooling in visual recognition, In *Proceedings of the 27th international conference on machine learning (ICML-10)*, (2010), 111–118.

42. Y. Lecun, Learning invariant feature hierarchies, In *European conference on computer vision*, Springer, Berlin, Heidelberg, (2012), 496–505

43. A. P. Dobrowolski, M. Wierzbowski, K. Tomczykiewicz, Multiresolution MUAPs decomposition and SVM-based analysis in the classification of neuromuscular disorders, *Comput. Methods Programs Biomed.*, **107** (2012), 393–403.

44. J. Schmidhuber, Deep learning in neural networks: An overview, *Neural Networks*, **61** (2015), 85–

117.

45. H. Papadopoulos, V. Vovk, A. Gammerman, Guest editors' preface to the special issue on conformal prediction and its applications, *Ann. Math. Artif. Intell.*, **74** (2015), 1–7.

46. N. Kambhatla, T. K. Leen, Dimension reduction by local principal component analysis, *Neural Comput.*, **9** (1997), 1493–1516.

47. A. Pathak, B. Vohra, K. Gupta, Supervised learning approach towards class separability-linear discriminant analysis, in: *2019 Int. Conf. Intell. Comput. Control Syst.*, IEEE, (2019), 1088–1093.

48. W. Yang, Q. Feng, M. Yu, Z. Lu, Y. Gao, Y. Xu, et al., Content-based retrieval of brain tumor in contrast-enhanced MRI images using tumor margin information and learned distance metric, *Med. Phys.*, **39** (2012), 6929–6942.

49. M. Huang, W. Yang, M. Yu, Z. Lu, Q. Feng, W. Chen, Retrieval of brain tumors with region-specific bag-of-visual-words representations in contrast-enhanced MRI images, *Comput. Math. Methods Med.*, **2012** (2012), 1–17.

50. M. Huang, W. Yang, Y. Wu, J. Jiang, Y. Gao, Y. Chen, et al., Content-based image retrieval using spatial layout information in brain tumor T1-weighted contrast-enhanced mr images, *PLoS One*, **9** (2014), e102754.

51. P. Afshar, K. N. Plataniotis, A. Mohammadi, Capsule networks for brain tumor classification based on mri images and coarse tumor boundaries, in: *ICASSP 2019 - 2019 IEEE Int. Conf. Acoust. Speech Signal Process.*, IEEE, (2019), 1368–1372.

52. P. Afshar, A. Mohammadi, K. N. Plataniotis, Brain tumor type classification via capsule networks, in: *2018 25th IEEE Int. Conf. Image Process.*, IEEE, (2018), 3129–3133.

53. J. S. Paul, A. J. Plassard, B. A. Landman, D. Fabbri, Deep learning for brain tumor classification, in: *Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging.*, International Society for Optics and Photonics, 10137 (2017), 1013710.

54. E. I. Zacharaki, S. Wang, S. Chawla, D. Soo, R. Yoo, E. R. Wolf, et al., Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme, *Magn. Reson. Med.*, **62** (2009), 1609–1618.

55. A. Kabir Anaraki, M. Ayati, F. Kazemi, Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms, *Biocybern. Biomed. Eng.*, **39** (2019), 63–74.

56. J. Sachdeva, V. Kumar, I. Gupta, N. Khandelwal, C. K. Ahuja, Segmentation, feature extraction, and multiclass brain tumor classification, *J. Digit. Imag.*, **26** (2013), 1141–1150.

57. R. A. Lerski, K. Straughan, L. R. Schad, D. Boyce, S. Blüml, I. Zuna, VIII. MR image texture analysis—An approach to tissue characterization, *Magn. Reson. Imaging.*, **11** (1993), 873–887.

58. S. Herlidou-Même, J. Constans, B. Carsin, D. Olivie, P. Eliat, L. Nadal-Desbarats, et al., MRI texture analysis on texture test objects, normal brain and intracranial tumors, *Magn. Reson. Imaging.*, **21** (2003), 989–993.

59. L. R. Schad, S. Blüml, I. Zuna, IX. MR tissue characterization of intracranial tumors by means of texture analysis, *Magn. Reson. Imaging.*, **11** (1993), 889–896.

60. A. Devos, A.W. Simonetti, M. van der Graaf, L. Lukas, J. A. K. Suykens, L. Vanhamme, et al., The use of multivariate MR imaging intensities versus metabolic data from MR spectroscopic imaging for brain tumour classification, *J. Magn. Reson.*, **173** (2005), 218–228.

61. L. Lukas, A. Devos, J. A. K. Suykens, L. Vanhamme, F. A. Howe, C. Majós, et al., Brain tumor classification based on long echo proton MRS signals, *Artif. Intell. Med.,* **31** (2004), 73–89.

62. A. Devos, L. Lukas, J. A. K. Suykens, L. Vanhamme, A. R. Tate, F. A. Howe, et al., Classification of brain tumours using short echo time 1H MR spectra, *J. Magn. Reson.*, **170** (2004), 164–175.

63. Y. Huang, P. J. G. Lisboa, W. El-Deredy, Tumour grading from magnetic resonance spectroscopy: A comparison of feature extraction with variable selection, *Stat. Med.,* **22** (2003), 147–164.

64. A. R. Tate, C. Majós, A. Moreno, F. A. Howe, J. R. Griffiths, C. Arús, Automated classification of short echo time in in vivo 1 H brain tumor spectra: A multicenter study, *Magn. Reson. Med.*, **49** (2003), 29–36.

65. Y. D. Cho, G. H. Choi, S. P. Lee, J. K. Kim, 1H-MRS metabolic patterns for distinguishing between meningiomas and other brain tumors, *Magn. Reson. Imaging.*, **21** (2003), 663–672.

66. Y. Pan, W. Huang, Z. Lin, W. Zhu, J. Zhou, J. Wong, et al., Brain tumor grading based on Neural Networks and Convolutional Neural Networks, in: *2015 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, IEEE, (2015), 699–702.