

Received April 6, 2022, accepted April 29, 2022, date of publication May 11, 2022, date of current version May 16, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3172303

Machine Learning Assisted Methodology for Multiclass Classification of Malignant Brain Tumors

ANKIT VIDYARTHI¹, RUCHI AGARWAL², DEEPAK GUPTA³,
RAHUL SHARMA⁴, DIRK DRAHEIM⁴, AND PRAYAG TIWARI⁵

¹Department of CSE and IT, Jaypee Institute of Information Technology, Noida 201309, India

²Department of CS, JIMS Engineering Management Technical Campus, Greater Noida 201303, India

³Department of CSE, Maharaja Agrasen Institute of Technology, Delhi 110096, India

⁴Information Systems Group, Tallinn University of Technology, 12616 Tallinn, Estonia

⁵Department of Computer Science, Aalto University, 02150 Espoo, Finland

Corresponding author: Prayag Tiwari (prayag.tiwari@aalto.fi)

This work was supported in part by the Academy of Finland under Grant 336033 and Grant 315896, in part by the Business Finland under Grant 884/31/2018, and in part by the EU H2020 under Grant 101016775.

ABSTRACT Analysis of malignant and non-malignant brain tumors is done using a computer-aided diagnosis system by practitioners worldwide. Radiologists refer computer-assisted techniques to draw conclusions using image modalities and inferences. Pedagogically, various machine learning approaches have been used, which usually focus on the classification of imaging modality into two categories, either normal and abnormal images or differentiating between benign and malignant tumors. Still, the work requirement is to classify these multi-class malignant tumors into their specific class with better precision. The proposed work focuses on distinguishing between the types of high-grade malignant brain tumors. This study is performed on real-life malignant brain tumor datasets having five classes. The proposed methodology uses the vast feature set from six domains to capture most of the hidden information in the extracted region of interest. Later, relevant features are extracted from the feature set pool using a new proposed feature selection algorithm named the Cumulative Variance method (CVM). Next, the selected features are used for model training and testing using K-Nearest Neighbour (KNN), multi-class Support Vector Machine (mSVM), and Neural Network (NN) for predicting multi-class classification accuracy. The experiments are performed using the proposed feature selection algorithm with three classifiers. The mean average classification accuracy achieved by using the proposed approach is 88.43% (KNN), 92.5% (mSVM), and 95.86% (NN), respectively. The comparative analysis of the proposed approach with other existing algorithms like ICA, and GA suggest that the proposed approach gains an increase of accuracy around 2% (KNN), 3% (SVM), and 4% (NN). The experimentation results concluded that the proposed approach is found better with NN classifier with an accuracy of 95.86% using diversified features.

INDEX TERMS Feature extraction, feature selection, malignant brain tumors, region of interests (ROI's), k-nearest neighbour (KNN), multi-class support vector machine (m-SVM), neural network.

I. INTRODUCTION

A tumor is a medical term for abnormal cell growth inside the brain. Brain tumors in the human body comprise a diverse group of neoplasms that vary in their behavior depending on various factors, such as the cell of origin, site of occurrence, morphology, and pattern of spread. The human brain is the

The associate editor coordinating the review of this manuscript and approving it for publication was Donato Impedovo¹.

most sensitive part of the body, which controls muscle movements and interprets sensory information like sight, sound, touch, taste, pain, etc. Any tumor can affect such sensory details and muscle movements or even result in more hazardous situations, including loss of life. Since the position of the tumor is not designated, it can occur in any part of the brain. Depending upon the place of origination, the tumor can be categorized into primary and secondary tumors [1]. If the tumor has originated inside the skull, it is known as

a primary brain tumor; otherwise, if the tumor's origination point is somewhere else in the body and has later moved to the brain, such tumors are called secondary tumors.

The locality of the tumor is not specific to any pre-defined place, so it can affect any of the brain lobes where it originates. Clinical methodologies for the treatment of brain neoplasms place trust in a variety of imaging modalities such as CT scan, MRI, fMRI, PET, and so on, with each imaging modality serving a distinct role in disease diagnosis. Among the various imaging modalities, Magnetic Resonance Imaging (MRI) is the most widely used and radiology relies on imaging in clinical diagnosis at multiple stages of treatment. MRI helps in the visualization of the soft tissues more effectively than any other imaging technique. The literature suggests that the MRI modality gives better statistically based analysis results in classification purposes than other imaging techniques.

The MR imaging system generates several voxels of different modalities, which are T1-Weighted, T2-Weighted, eT1-weighted, eT2-Weighted, and FLAIR. Among the various modalities suggested, the T1-weighted modality is widely used. These images are made post-contrast-enhanced using a contrast material named gadolinium. This contrast material is injected into the patient's body using an injection having 0.15-0.20 mMol/kg dose. The radiologists then visualize the post-contrast-enhanced T1-weighted MR image using a computer-aided system to find out the tumors or any other abnormality in brain MR images.

These imaging techniques do not lead to perfect identification of the type of neoplasm as they give a closer idea to the characterization of tumor only as neoplasm tissues are often heterogeneous in spatial and imaging systems. Furthermore, malignant brain neoplasm exhibits a wide range of characterizations and variations in imaging systems. The reference standard for brain neoplasm characterization is currently based on pathologic analysis based on biopsy. Some machine-learning-based analyses aid in gaining an understanding of brain neoplasm abnormalities [2]. Various algorithmic concepts have been proposed to analyze any given MR image based on statistical parameters and are based on statistical results that the machine will learn and predict the class of abnormality.

Based on the literature study, it was found that most of the literature studies are towards binary classification problem where categorization is between normal and abnormal classes or in between benign and malignant. These literature studies use machine learning approaches for the classification with the extraction of only textural features. While there exists some work where spectral features were used but they lack in the classification accuracy. Besides this for multi-class problems the claimed accuracy was not found promising may be because of using only textual features for representation and model learning. Motivated by this, the objective of current study is towards exploring the diversified feature set by incorporating large number of features for image representation. Next, these features are filtered out by selecting the most

appropriate features which are relevant and non-redundant in nature. At the end, the selected features are used for the model training and explore the power of computation using three state-of-art classifiers.

The rest of this paper is organized as follows: The related work towards the use of a machine learning-based diagnosis system for brain tumor imaging is presented in Section II. Section III discusses the proposed methodology for tumor classification. Section IV provides the experimental results. Section V gives the conclusion of the paper followed by discussion in Section VI.

II. RELATED WORK

Recent literature after the year 2016 mainly focuses on the use of Deep Learning and Transfer learning to classify tumor images. But succumbing to the limitation of the scope of the presented work towards the machine learning aspect, the authors have included only those papers which are on machine learning and deliberately exclude deep learning.

For the classification of brain tumors in [3], the author proposed the use of a genetic algorithm for feature extraction, which used fuzzy c-means for segmentation purposes. Intensity-based shape features were identified using Fourier [4] analysis for classification of the breast tumors. A more robust feature extraction model is defined in [5], which is based on the concept of the GARCH series where wavelet-based parameters were used to calculate the feature vector followed by PCA and LDA for the selection of relevant features. Another approach for generating intensity-based features using Discrete Wavelet Transformation (DWT) was proposed for the classification of brain tumors using Bayesian Neural Network (BNN) [6].

For tumor classification, three-domain features are extracted: spatial domain, frequency domain, and intensity-based features [7]. These feature vectors represent the overall imaging statistics that are used by the classifier for its categorization. For classification, the literature suggests that there are two types of classifiers used, i.e., linear and non-linear. Wavelet-based feature extraction was used primarily with non-linear classifiers and neural network variants [5], [8].

In [9], the author describes the comparative study for the classification of brain tumors which was done from 2009 to 2013. The comparative research provides the detail of the work done for the tumor classification using different feature extraction and selection techniques and compares the accuracy of these approaches. In [51], the authors proposed model to detect glioblastoma tumor using deep belief network.

In [11], the author gives the idea of the use of intensity, shape, and texture as spatial domain features for the segmentation of posterior-fossa tumors. The imaging features' effectiveness is shown with a feature selection mechanism for the segmentation of tumor regions from MR images. A decision forest-based tumor segmentation using symmetric texture and symmetric intensity-based features is shown and described [15]. Another approach to frequency-based feature

classification using Discrete Wavelet Transformation (DWT) is described in [12].

Classification based on artificial neural networks has been widely used by various authors for tumor classification. In a similar approach, a supervised KNN and feed-forward Artificial Neural Network (ANN) hybrid classifier was used for MRI classification. Some approaches rely on the probabilistic neural network model PNN with textural features, which is also used for brain tumor characterization [13]. A modification based on the iteration of ANN was used in [14], which improved the performance rate of abnormal to normal brain MR image classification where ANN iteration-free was used, which ultimately improved the convergence rate besides yielding accurate results. The performance of these networks was analyzed in the context of abnormal brain image classification.

An ANN and its variant-based classification were described in [16]–[19], where the GLCM-based texture features along with some intensity-based features were used [16]. The dimensionality of the feature vector was further reduced using PCA. Using ANN and its variants for classification, various authors used a variable dataset consisting of 60–428 MR images of 256×256 dimensions. It was observed that for a small dataset of 60 or 80 images, the classifier gave an accuracy of nearly 91% [16]. But as the size of the dataset increased up to 273, the classification accuracy decreased to 64% and varied in a range of 64%–94% using Bayesian ANN [19]. While using a huge dataset of 428 MR images, the author of [16] claims an accuracy of 85%.

Instead of bi-class brain tumor classification on MR images, some multi-class classifications for brain tumors were proposed by the author of [25]–[27]. The experiments were performed on the dataset of 98 images [25], which resulted in the individual accuracy factor of MTS-92%, GBM-41%, GLI-91% approx. Likewise, the authors of [26], [27] used the MR dataset of 75 brain tumor images, including tumors of type Metastases, Meningioma (MEN), and Gliomas, which were classified using the Least Square Feature Transformed Probabilistic Neural Network (LSFTPNN). This classifier works in two domains, i.e., in the first phase, the features are transformed using the Least Square function method, and then the features are given as an input to the probabilistic neural network classifier. The experimental results showed MTS-87.5%, GLI-97%, MEN-95% accuracy.

Instead of MR imaging, many studies have suggested the use of MR spectroscopy for brain tumor classification [20]–[22]. To differentiate between benign and malignant brain neoplasm specifically, spectroscopic and conventional MR imaging were used in [20] by the application of a decision tree algorithm. In addition, spectroscopic and perfusion MRI were used in [22] to assess the inherent heterogeneity of brain neoplasms based on tumor and peri-tumor regions of interest (ROIs). Some recent works exhibit the use of the ensemble approach either on classifiers [33] or on feature extraction [34] for the classification of brain tumor images.

These studies have used various feature extraction methods to extract relevant features. Other works on the use of extracting relevant statistical features with a family of neural networks such as BPNN [35] and probabilistic Neural Network [36] for tumor classification were also used.

The literature study shows that few studies focused on high-grade malignant brain tumors, and the experiments performed resulted in low accuracy, even on smaller datasets of MR images. The comparative study of these works is presented in Table 1. Some malignant tumors like CNC and IVMM have not been taken into consideration yet. In this paper, an attempt has been made to classify the five types of high-grade malignant tumors with normal regions, which may help radiologists better analyze the subject matter in MR images.

III. PROPOSED APPROACH

The proposed approach adopted in this work is the extension of our previously published work [37]. The previous work was based on the usage of frequency-based features using Gabor and Wavelet Transformation. While the presented extended work uses a large set of features pool incorporating both spatial and spectral domain features for the classification of malignant brain tumors, the anticipated approach provides detailed information on the usage of machine learning approaches for the classification of high-grade malignant tumors in brain MR images. To identify the degree of malignancy and grade of tumor based on the World Health Organization (WHO) grading system, a diverse feature set including features vector of Intensity-based features, Frequency-based features, Texture features, Shape features, Contour based features, and Moments based features are used. This diverse feature set helps to gain the benefit of analyzing the malignant image in both the spatial and frequency domains, and thereby includes both forms of image analysis parameters.

Machine learning (ML) is an artificial intelligence branch that focuses on the construction and study of systems that can learn from data, use a diverse feature set for their learning, and predict the degree of malignancy. ML is a five-step process that helps to learn and build a model for its prediction. These steps are standard dataset collection which is taken into consideration, pre-processing of the dataset, feature extraction, feature selection, and classification [50]. All the model building steps are equally important and have their specific role, which is discussed in detail below:

A. STANDARD DATA COLLECTION/DATA SOURCE

The proposed system examined over 110 patients with high-grade malignant brain tumors. From these patients, a dataset of 660 malignant tumor images is generated. Among the generated dataset, an overall of 1160 Regions of Interest (ROI's) is extracted, which are taken into consideration. These ROIs are classified into 760 malicious regions and 400 normal regions. The examined dataset was collected over the duration from October 2013 to October 2014 [37]. All the patients are identified initially at the time of MRI, followed

TABLE 1. Comparison of previous studies on classification of brain tumors.

Author- year	Brain Tumor classes/ Images per Class	Image Dataset	Feature Extraction	Feature Selection	Classifier	Accuracy Obtained (%)
Megha.P. Arakeri-Springer, 2013	1-Primary (42) 2- Malignant (40)	82	Shape+Texture (GLCM)	PCA+KD Tree indexing	SVM, ANN, K-NN	97
Jainysachdeva, vinodkr-Springer, 2013	Astocytoma, Glioblastoma, Medulloblastoma, meningioma, Metastatic	428	GLCM+LoG+IB F+RILBP	PCA	ANN+PCA-ANN	85.23
N.HemaRajni, R. Bhavani-Springer, 2013	Normal (20) Abnormal (60)	80	DWT	PCA	ANN,K-NN, SVM	91, 93, 95
D. sridhar, Murali Krishna- IEEE, 2013	Primary, Secondary	20	DCT	DCT	PNN	100
Padma Nanthagopal-Springer, 2012	Normal (40) Benign (37) Malignant (43)	120	DWT+WCT	GA	PNN	96
Padma Nanthagopal-Springer, 2012	Normal (40) Benign (37) Malignant (43)	120	DWT+WCT	GA+PCA	SVM	96
Prof.Vikas Gupta- IEEE, 2012	Class I,II,III,IV	60	GLCM	--	ANN	89
P. Rajendran, M. Madheswaran-IEEE, 2012	Nornal (151), Benign (22), Malignant (27)	200	Shape features	ASM	Decision Tree	90
DrMohdFauzi Bin Othman-IEEE, 2012	Normal, Abnormal	15	DWT	--	Kernel-SVM	N-98 AB- 67
Amer Al-Badarneh-IEEE/ACM 2012	1- Normal (10) 2- Abnormal (60)	70	DWT	PCA	ANN+KNN	N- 96.77 AB- 97.56
Evangelia I. Zacharaki-Springer, 2011	Meningioma,Glioma Grade II,III,IV,Metastatis	100	Age,shape,intensity	PCA	J48,KNN,VFI, SVM,Naive Bayes	85, 93.5, 95, 93
GinniGarg, 2021	Benign, malignant	2556	SWT, PCA, GLCM		KNN-RF-DT	97.3
Jaeyong K, MDPI, 2021	Normal, Abnormal	253-3000	Deep Model	Top Three	NB, AdaBoost, k-NN, RF, SVM	73 - 93
Ismael M.R., IEEE, 2018	Meningioma, Glioma, and Pituitary tumor	3,064	Statistical	DWT, Gabor	BPNN	91.9
Varuna S.N, Springer, 2018	Normal, Abnormal	25	GLCM+DWT	--	Probabilistic NN	100
Emrah Irmak, Springer, 2021	Normal, Glioma, Meningioma, pituitary and metastatic	3950	CNN Model			95-97
Francisco Javier Diaz-Pernas, MDPI, 2021	Meningioma, Glioma, and pituitary tumor	3064	CNN Model			97.3
XiaoqingGu, Neuroscience, 2021	Meningioma, Glioma, and Pituitary tumor	3064	Convolution Dictionary Learning + AlexNet			91-96
Mesut T., Springer 2021	Meningioma, Glioma, and Pituitary tumor	3064	BrainMRNet including hypercolumn technique, attention modules, and residual blocks			96-98

by a biopsy after surgery. The collected dataset consists of 5 class types of malignant tumors which are named Central NeuroCytoma (CNC), GlioblastomaMultiforme (GBM), Gliomas (GLI), Intra Ventricular Malignant Mass (IVMM), Metastasis (MTS), and 1 class type of normal regions. Thus, the generated dataset has a total of 6 classes which are used as an input to the classifier.

The experimental images are generated using a 3.0 Tesla GE MRI Scanner at SMS Medical College Jaipur, Rajasthan, India. All the patients whose images are used for dataset generation are imaged using the same imaging system and environment variables. The obtained images have the following specifications: 3D axial T1-Weighted, T2-weighted, eT1-Weighted, eT2-Weighted images, and Fluid Attenuated Inversion Recovery (FLAIR), each having a size of $256 \times 256 \times 3$. In the proposed work, a post-contrast-enhanced T1-weighted axial imaging modality is used for examination purposes, having a dimensionality of $256 \times 256 \times 3$ [37].

B. PRE-PROCESSING

The input MR image is pre-processed to remove any kind of noise generated during image generation. Noise removal helps us to discard unwanted signals which can lead to errors during processing. For noise removal, a median filter is used as it preserves edges while removing noise [37].

A median filter is then used to scan pixel by pixel of the image and replace each entry with the median value of the neighborhood pixels. In the proposed work, a 3×3 box window is used as a median filter. Next, the filtered image is converted into an 8-bit grey level to reduce the bands of the image. An original 3 band color image is converted to a corresponding intensity image ranging between 0 and 255. After preprocessing, the input image becomes a noise-free, intense image of size 256×256 , which serves as an input to the next step.

After converting the image into the grey level, the image is preprocessed for sharpening. Since the soft tissues in the MR image are a concern, the input grey image is sharpened to visualize the soft tissues. Sharpening of an image can be done using high-pass filters, which help to pass high-frequency components and discard low-frequency components. Therefore, the filtered image generated after a high pass filter results in a sharpened image where soft tissues and edges are visualized. In this paper, a high pass Gaussian filter is used for the sharpening of MR images.

The next part of pre-processing is skull stripping. Since the concerned area of the proposed work is to classify high-grade malignant tumors in brain images, which always has a high-intensity value as compared with other parts of the brain, it needs to be removed. Malignant tumors are mainly associated with soft tissues of the brain and the skull part in the ongoing processed image gives unnecessary information and increases complexity.

After skull stripping from MR images, the next step is the extraction of the region of interest from the MR images.

In any MR image, some regions may be identified as malicious and differentiated from others. These regions are under suspicion for radiologists during the analysis of the MR image using the computer-aided machine. From any MR image, several regions are marked by the radiologists, who are termed as “malicious” and “normal” regions. In this paper, Region of Interest (ROI's) extraction is done using an automated approach that selects the abnormal identified malicious region of interest and normal regions using free flow marking of the boundary region. The selected malicious and normal regions of interest are extracted from sharpened skull removed images.

Among 110 malignant brain tumor suspects, a dataset of 660 MR images is prepared. From this dataset, regions marked as malignant and normal by the radiologist are extracted. This resulted in the form of 1160 ROI's having 760 malignant regions and 400 normal regions. Thus, the obtained dataset has 1160 regions which are going to be classified using the machine learning approach in 6 classes.

C. FEATURE EXTRACTION

This subsection provides detailed information about the varieties of features used using machine learning for predicting the degree of malignancy. This diverse feature vector includes features like Intensity-based features, Frequency-based features, Texture based features, Shape-based features, Contour features, and Moments based features. The generated feature set results in gaining the benefit of processing a single image in a variety of its representation forms. Since an image can be represented and processed on digital machines either in the frequency domain or in the spatial domain, both the representation forms have significance over each other. In the proposed work, both image representation forms are taken into consideration and a feature set is obtained over those representation forms which have the properties of the two domains.

1) FREQUENCY-BASED FEATURES

Imaging machines use signal values for storing and reconstructing brain images. Signal processing is thus an important factor for feature analysis. The input malignant image is thus mapped to the signal domain using Discrete Wavelet Transform (DWT), which provides enough information for analysis and synthesis of the original signal while reducing computation time significantly. To use DWT for brain MR images, we implement the 2D variant of the analysis and synthesis filter bank, which results in forming an image in four bands, i.e.,

LL, LH, HL, HH, as shown in Figure 1 using row-wise and column-wise DWT transformations on an input image. From the resultant signal obtained using 2D DWT, several features are obtained like Contrast, Homogeneity, Entropy, Dissimilarity, and Energy. The phase offsets (θ) which are taken into consideration for 2D DWT are 0° , 45° , 90° , and 135° .

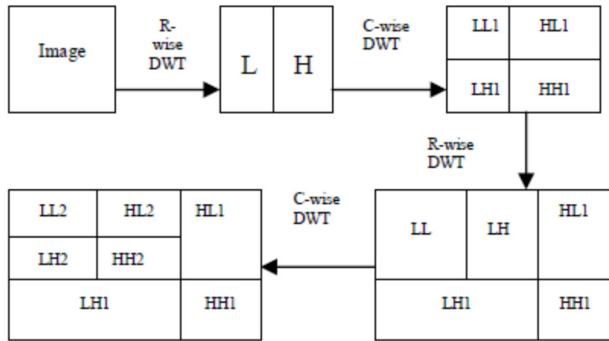


FIGURE 1. 2D DWT block diagram.

The above features can be represented as given below:

$$\begin{aligned} \text{Contrast} &= \sum_{i=0}^{c-1} \sum_{j=0}^{n-1} (i-j)^2 g(i,j) \\ &\times \sum_{i=0}^{c-1} \sum_{j=0}^{n-1} (i-j)^2 g(ij) \\ &\times \sum_{i=0}^{c-1} \sum_{j=0}^{n-1} (i-j)^2 g(i,j) \end{aligned} \quad (1)$$

$$\text{Homogeneity} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \frac{1}{1 + (i-j)^2} g(i,j) \quad (2)$$

$$\text{Entropy} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} g(i,j) (-\ln(g(i,j))) \quad (3)$$

$$\text{Dissimilarity} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} g(i,j) \text{mod}(i-j) \quad (4)$$

$$\text{Energy} = \sqrt{\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} g(i,j)^2} \quad (5)$$

where, i, j are the coordinates for the normalized co-occurrence matrix space are, $g(i, j)$ is the element of the i, j coordinate's space and c is the dimension of the normalized co-occurrence matrix. Among every offset, all the above five features are calculated which results in forming a feature vector of size 20.

The next feature extraction mechanism is also based on the wavelet formed using the Gabor filter. The frequency and orientation representation of the Gabor filter is found to be appropriate for texture representation and discrimination. In image analysis, the image is processed by the Gabor filter, which results in the generation of Gabor descriptors for that image. The Gabor descriptor of a linear low pass filter is defined as:

$$G_{\theta, \sigma, \varphi}(x, y) = \exp\left(-\frac{A^2 + \gamma^2 B^2}{2\sigma^2}\right) \cos\left(\frac{2\pi}{\lambda}A + \varphi\right) \quad (6)$$

where, $A = x \cos(\theta) + y \sin(\theta)$,

$B = -x \sin(\theta) + y \cos(\theta)$

In the above Equation (6), represents the wavelength of the sinusoidal factor, θ represents the orientation of the Gabor function, φ is the phase offset, σ is the standard deviation of

the Gaussian factor and γ is the spatial aspect ratio to specify the ellipticity of Gabor function.

To calculate the texture feature from an image the outputs of the symmetric ($\varphi = 0$) and anti-symmetric ($\varphi = \pi/2$) Gabor kernel are combined using the distance metric and 2D linear convolution. The mathematical model of such is given as:

$$\begin{aligned} G_{\lambda, \theta}(x, y) \\ = \sqrt{(I(x, y) \otimes g_{\lambda, \theta, \varphi=0}(x, y))^2 + (I(x, y) \otimes g_{\lambda, \theta, \varphi=\pi/2}(x, y))^2} \end{aligned} \quad (7)$$

Textural features of an image are calculated by using a set of Gabor filters with different frequencies and orientations. We use five wavelengths, i.e., $\lambda = \{2\sqrt{2}, 4, 4\sqrt{2}, 8, 8\sqrt{2}\}$ and four orientations in $[0, \pi)$ i.e. $\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ which are taken at an equal interval of 45° . The special aspect ratio ($\gamma = 1$) is selected for computation. Thus the above statement concludes that for a single wavelength factor and particular orientation, we can obtain five textural features for an image. Thus the total feature set obtained is of size 100. Now the above feature set obtained is only dependent on the single standard deviation value which is taken into consideration. We used two iterative values of standard deviation, i.e., $\sigma = (1.5, 2.5)$, and finally, the obtained feature vector size becomes 200 features in total for texture analysis using Gabor.

The above two feature extraction techniques, named Gabor wavelet and Discrete Wavelet Transformation, helped to form a diverse feature set by combining both feature domain vectors, resulting in the formation of a large feature vector of 220 features.

2) TEXTURE BASED FEATURES

The gray level extracted ROI's based MR dataset is processed for feature extraction using Gray Level co-occurrence Matrix (GLCM). Several features are added to the feature set based on the computation over GLCM. The computation in this paper is based on the selected four offset values i.e. $\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$. For each offset value, various features are extracted, which is discussed as:

Let $V_\theta(i, j)$ be the co-occurrence value of the gray level matrix at intensity pair (i, j) on specific offset θ . The extracted features description are given as:-

$$\begin{aligned} \text{Contrast} \\ = \sum_{i,j} \text{mod}(i-j)^2 \cdot V_\theta(i,j) \end{aligned} \quad (8)$$

$$\begin{aligned} \text{Correlation} \\ = \sum_{i,j} \frac{(1 - \mu_i)(1 - \mu_j) \cdot V_\theta(i,j)}{\sigma_i \sigma_j} \end{aligned} \quad (9)$$

$$\begin{aligned} \text{Homogeneity} \\ = \sum_{i,j} \frac{V_\theta(i,j)}{1 + \text{mod}(i-j)} \end{aligned} \quad (10)$$

$$\begin{aligned} \text{Energy} \\ = \sum_{i,j} V_\theta(i,j)^2 \end{aligned} \quad (11)$$

$$\begin{aligned} \text{Angular Second moment} \\ = \sum_{i,j} V_{\theta}(i,j)^2 \end{aligned} \quad (12)$$

$$\begin{aligned} \text{Inverse Difference moment} \\ = \sum_{i,j} \frac{1}{1+(i-j)^2} V_{\theta}(i,j) \end{aligned} \quad (13)$$

$$\begin{aligned} \text{Variance} \\ = \sum_{i,j} [(i-\mu)^2 V_{\theta}(i,j)] \end{aligned} \quad (14)$$

$$\begin{aligned} \text{Entropy} \\ = - \sum_{i,j} V_{\theta}(i,j) \log(V_{\theta}(i,j)) \end{aligned} \quad (15)$$

Let N_g be defined as the no. of distinct gray levels in an image obtained using histogram of an image and $V_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} V_{\theta}(i,j)$ and $i+j=k$, i.e. $k=2,3,4 \dots 2N_g$ then,

$$\text{Sum average} = \sum_{i=2}^{2N_g} i \vartheta_{x+y}(i) \quad (16)$$

$$\text{Sum variance} = \sum_{i=2}^{2N_g} (i-\mu)^2 \vartheta_{x+y}(i) \quad (17)$$

$$\text{Sum entropy} = - \sum_{i=2}^{2N_g} \vartheta_{x+y}(i) \log(\vartheta_{x+y}(i)) \quad (18)$$

$$\text{Diff entropy} = - \sum_{i=0}^{2N_g-1} \vartheta_{x-y}(i) \log(\vartheta_{x-y}(i)) \quad (19)$$

The above feature extraction mechanism generates 12 different features for a particular offset, which results in a total of 48 features based on the four different offsets which are added to the feature pool.

3) INTENSITY-BASED FEATURES

In the intensity-based features extraction system, the gray level intensity image is modelled using a histogram of an image. The histogram is the process in which each eminent-intensity pixel is mapped to the total number of times it appears in an image. If $h(k)$ be the histogram representation of the k^{th} intensity value of an image, then $h(k)$ is defined as:

$$h(k) = N_k$$

where, N is the number of times the k^{th} intensity value appears in an image.

For each ROI's dataset of MR image, several features are extracted using a histogram of an image. If N is the total number of pixels in an image and $I(i,j)$ is the intensity value of an image at indexed (i,j) then these features are described as:

$$\text{Mean} = \frac{\sum_{i,j} I(i,j)}{N} \quad (20)$$

$$\text{Standard Deviation (SD)} = \sqrt{\frac{\sum_{i,j} (I(i,j) - m)^2}{N}} \quad (21)$$

$$\text{Skewness (SW)} = \frac{1}{N} \frac{\sum_{i,j} (I(i,j) - m)^3}{SD^3} \quad (22)$$

$$\text{Kurtosis (KR)} = \frac{1}{N} \frac{\sum_{i,j} (I(i,j) - m)^4}{SD^4} \quad (23)$$

$$\text{Entropy (E)} = - \sum_{k=0}^{l-1} h(k) \ln h(k) \quad (24)$$

The above intensity-based feature extraction mechanism generates 5 features that are added to the feature set pool.

4) SHAPE-BASED FEATURES

The shape-based feature extraction mechanism has a great impact on image retrieval systems. The main idea of these features is to search out the image based on specific shape descriptions. For the classification of the tumor, some basic shape descriptor features are calculated. These basic shape features are described as:

$$\text{Circularity} = 1 - \frac{4\pi a}{P} \quad (25)$$

where P is perimeter and a is an area of an object

$$\text{Bending Energy} = \left(\frac{1}{P}\right) * \sum_{i=1}^{N-2} C_i^2 \quad (26)$$

where, curvature of the contour at the i^{th} position can be approximated as:

$$C_i = (2x_i - x_{i-1} - x_{i+1})^2 + (2y_i - y_{i-1} - y_{i+1})^2$$

x_i, y_i represent the i^{th} coordinate location describing an object's contour.

$$\text{Circularity ratio} = \frac{\sigma_r}{\mu_r} \quad (27)$$

where, σ_r and μ_r are the mean and standard deviation of the radial distance from the centroid (g_x, g_y) of the shape to the boundary points (x_i, y_i) i.e.

$$\mu_r = \frac{1}{N} \sum_{i=1}^{N-1} d_i; \quad \sigma_r = \sqrt{\frac{1}{N} \sum_{i=1}^{N-1} (d_i - \mu_r)^2}$$

and

$$\begin{aligned} d_i &= \sqrt{(x_i - g_x)^2 + (y_i - g_y)^2} \\ \text{Rectangularity} &= \frac{A_s}{A_r} \end{aligned} \quad (28)$$

where, A_s is the area of a shape and A_r is the area of the minimum bounding rectangle.

$$\text{Convexity} = \frac{\Phi}{C} \quad (29)$$

Convexity is defined as the ratio of perimeters of the convex hull Φ over that of the original contour C .

The above feature extraction mechanism helps to get shape-based features of the extracted ROI's from brain MR images and thus adds five shape-based features to the feature set pool.

5) CONTOUR BASED FEATURES

Let, x_i , and y_i represent the i^{th} coordinate location describing an object's contour and z_i be the Euclidean distance of this point to the Centroid of the object. Then, the p^{th} CSM is defined as:

$$m_p = \frac{1}{N} \sum_{i=1}^N (Z_i)^p \quad (30)$$

Similarly, the p^{th} central CSM is defined as:

$$M_p = \frac{1}{N} \sum_{i=1}^N [Z_i - m_1]^p \quad (31)$$

Here in this work we select the following low-order moments to form a set of shape features:

$$F_1 = \frac{\sqrt[2]{M_2}}{m_1}, \quad F_2 = \frac{\sqrt[3]{M_3}}{m_1} \quad (32)$$

$$F_3 = \frac{\sqrt[4]{M_4}}{m_1}, \quad F_4 = F_3 - F_1 \quad (33)$$

The difference between F_3 and F_1 can give information about shape roughness which may not be represented by both F_1 and F_3 . The above feature extraction mechanism generates 4 contour point features which are added to the feature set pool.

6) MOMENT BASED FEATURES

A moment is a specific quantitative measure of the shape or set of points. If the points represent probability density, then the *zeroth* moment is the total probability (i.e. one), the first moment is the mean, the second moment is the variance, and the third moment is the skewness [29]. In image analysis, image moments are useful in describing objects after segmentation. Thus, after extracting the ROI's from the malignant brain MR image, moment based features help to describe the tumor type in image classification. Some of the basic image moments features are described as:

$$\text{Raw moment: } M_{i,j} = \frac{\sum_{i,j} x^i y^j I(x,y)}{\sum_{i,j} I(x,y)} \quad (34)$$

where, $I(x, y)$ is the pixel intensity at indexed (x, y) .

Central moments: three level central moments are given as:

if $\bar{x} = \frac{M_{10}}{M_{00}}, \bar{y} = \frac{M_{01}}{M_{00}}$ and $\mu_{01} = \mu_{10} = 0$ then,

I order moment: $\mu_{00} = M_{00}, \mu_{11} = M_{11} - \bar{x}M_{01}$

$$\mu_{02} = M_{02} - \bar{y}M_{01}, \mu_{20} = M_{20} - \bar{x}M_{01} \quad (35)$$

II order moment: $\mu_{21} = M_{21} - 2\bar{x}M_{11} - \bar{y}M_{20} + 2\bar{x}^2M_{01}$

$$\mu_{12} = M_{12} - 2\bar{y}M_{11} - \bar{x}M_{02} + 2\bar{y}^2M_{01} \quad (36)$$

III order moment: $\mu_{03} = M_{03} - 3\bar{y}M_{02} + 2\bar{y}^2M_{01}$

$$\mu_{30} = M_{30} - 3\bar{x}M_{20} + 2\bar{x}^2M_{10} \quad (37)$$

Rotation invariant moments: seven rotation invariant moment features are given as:

If, $\vartheta_{i,j} = \frac{\mu_{ij}}{\mu_{00}^{(1+\frac{i+j}{2})}}$ then,

$$F_1 = \vartheta_{20} + \vartheta_{02} \quad (38)$$

$$F_2 = (\vartheta_{20} - \vartheta_{02})^2 + 4\vartheta_{11}^2$$

$$F_3 = (\vartheta_{30} - 3\vartheta_{12})^2 + (3\vartheta_{21} - \vartheta_{03})^2 \quad (39)$$

$$F_4 = (\vartheta_{30} + \vartheta_{12})^2 + (\vartheta_{21} + \vartheta_{03})^2 \quad (40)$$

$$F_5 = (\vartheta_{30} - 3\vartheta_{12})(\vartheta_{30} + \vartheta_{12})[(\vartheta_{30} + \vartheta_{12})^2 - 3(\vartheta_{21} + \vartheta_{03})^2] + (3\vartheta_{21} - \vartheta_{03})(\vartheta_{21} + \vartheta_{03})^2 \times [3(\vartheta_{30} + \vartheta_{12})^2 - (\vartheta_{21} + \vartheta_{03})^2] \quad (41)$$

$$F_6 = (\vartheta_{20} - \vartheta_{02}) \left[(\vartheta_{30} + \vartheta_{12})^2 - (\vartheta_{21} + \vartheta_{03})^2 \right] + 4\vartheta_{11}(\vartheta_{30} + \vartheta_{12})(\vartheta_{21} + \vartheta_{03}) \quad (42)$$

$$F_7 = (3\vartheta_{21} - \vartheta_{03})(\vartheta_{30} + \vartheta_{12}) \left[(\vartheta_{30} + \vartheta_{12})^2 - 3(\vartheta_{21} + \vartheta_{03})^2 \right] - (\vartheta_{30} - 3\vartheta_{12})(\vartheta_{21} + \vartheta_{03})[3(\vartheta_{30} + \vartheta_{12})^2 - (\vartheta_{21} + \vartheta_{03})^2] \quad (43)$$

The above moment-based image feature extraction technique helps to extract the features set, which are rotation and translation invariant. These features help the radiologist to diagnose the malignant ROI extracted from brain MR images even after rotation and translation of the images on a variety of dimensional scales. This mechanism helps to generate 16 features that are added to the feature set pool of machine classification. The overall summary of all the feature extraction techniques with the number of extracted features is shown in Table 2.

TABLE 2. Summary of feature extraction techniques.

Type of Features	No. of Features extracted	No. of Orientations	Total no. of features
DWT	5	4	20
Gabor	5	20*2=40	200
Texture	12	4	48
Intensity	5	1	5
Shape	5	1	5
Contours	4	1	4
Moments	16	1	16
Total = 20+200+48+5+5+4+16 = 298			

The feature set pool now consists of a large number of features from a variety of feature extraction domains. This huge, diverse feature set pool is used for machine learning for classification.

D. FEATURE SELECTION

This subsection describes the feature selection mechanism, which is an important phase used in the machine learning system. Feature selection is also known as variable selection, subset selection, or even by the name of attribute selection. Feature selection is defined as a process of selecting a relevant set of features or attributes from the large feature set vector, which is used for building a machine learning model. The generated feature set can have some features which are irrelevant and redundant. These features have a lower impact on generating machine learning models, but identifying the irrelevant features is a major concern for model building.

In regard to machine learning, features are considered redundant if the selected feature provides no more information than the previously selected features. While irrelevant features are those features that provide no useful information in any context, they are also those features that provide no useful information in any context. For filtering out redundant and irrelevant features, feature selection algorithms are divided into two major sub-domains, i.e., feature ranking methods and feature subset selection methods [24].

Feature ranking methods are generally used for high-dimensional datasets. This method processes the initially generated features and computes the ranking score of each feature. These ranking score features are now sorted in decreasing order of their rank and the top-ranked features are selected for the classification. But the major problem associated with this approach is the selection of redundant features in the selected feature set. As if the features are redundant, then they have nearly the same feature value, which results in the formation of a higher ranking for both features, and these redundant features are selected. This problem occurs due to the absence of finding the correlation between the selected features.

On the other hand, feature subset selection mechanism spotlight on selecting a subset (s') from the overall feature set (S) such that $s \in S$. The subset (s') must form greater significance than the rest of the other features subset. For selecting the relevant subset, various methods are available as a wrapper method, Filter based method, and embedded methods [24].

In the proposed work, three experiments are done based on different feature selection algorithms, i.e., Cumulative Variance based feature selection (CVM), Genetic Algorithm (GA), and Independent Component Analysis (ICA). CVM is our previous proposed feature selection approach whose description is found in [28]. Other algorithms like GA and ICA are state-of-the-art algorithms. The basic description of the algorithms is given below:

1) CUMULATIVE VARIANCE METHOD FOR FEATURE SELECTION (CVM)

In this algorithm, CVM utilizes the advantages of both the feature selection approaches, i.e., feature ranking and subset selection. Firstly, the redundant and irrelevant features are eliminated from the feature vector using the feature ranking approach that is based on variance analysis among the features. The ranking of the features is dependent on the cumulative variance percentage. Thus, based on the above-proposed approach, we utilized the benefit of feature ranking. In the second step, the most desirable subset of features is selected among the ranked feature vectors using a statistical t-test approach as defined below:

$$t = \frac{\overline{x1} - \overline{x2}}{\sqrt{\frac{(N1-1)v1+(N2-1)v2}{N1+N2-2} \left(\frac{1}{N1} + \frac{1}{N2} \right)}} \tag{44}$$

where, $\overline{x1}$, $\overline{x2}$ are the sample means and $v1$, $v2$ are the sample variance of a particular feature subset. $N1$, $N2$ is the number of features in each subset.

The proposed cumulative variance and t-test based algorithm is given below:

2) Detailed Description of CVM

In feature selection, the input is the vector 'V', i.e., $V = v1, v2, \dots, vn$ returned by the feature extraction step.

Algorithm 1 CVM Based Feature Selection

Input: Feature vector (V)

Output: Relevant feature subset (RFS)

Procedure:

For each vector $v_i \in V$

$\mu_i \leftarrow Col_mean(v_i)$

$D_i \leftarrow (v_i - \mu_i)$

end for

$C \leftarrow Find_cov(D)$

$Ev \leftarrow Find_eigenVec(C)$

$Var \leftarrow Find_var(Ev)$

$SEv \leftarrow Sort(Var)$

$Cv \leftarrow cumulative(SEv)$

$RFS \leftarrow T_test(Cv)$

return (RFS)

The vector is of size $1 * n$ where each v_i is of dimension p , and $n =$ no. of data elements, $p =$ no. of features. Initially, we find out the mean of each column vector which is represented by the μ_i given in Equation 45. Then the data normalization is done by subtracting the mean from each column vector. For this normalized column vector, a covariance matrix is generated using Equation 46.

$$\mu_i = \frac{1}{n} \sum_1^n v_i \tag{45}$$

$$Cov = \left(\frac{1}{n-1} \right) D * D^T \tag{46}$$

where D^T is the transposed of the normalized vector D . The generated covariance matrix is of size $p * p$. An Eigenvector is calculated from the covariance matrix using Equation 47 as given below:

$$Ev^{-1} Cov Ev = D \tag{47}$$

where Ev is the Eigenvector matrix, Cov is the covariance matrix and D is the diagonal Eigenvalues matrix. After getting Eigenvectors, the variance among the vectors is computed for each Eigenvector as given in Equation 48.

$$Var = \frac{\sum_{i=1}^n X^i - \bar{X}^2}{(n-1)} \tag{48}$$

Such that, $x^i \in Ev$ (Eigenvectors).

All the extracted variances are now sorted in decreasing order of magnitude and a test is applied as such. The statistical t-test helps to find out the subsets using Equation 44. The subsets that can pass the t-test will be considered as relevant features, and the rest are all rated as irrelevant at once.

The above selection mechanism will automatically find the number of relevant features and return a vector with a lower number of features as compared with the original vector. Features that pass this significance level are selected, and the rest of the features are dropped. This reduced feature set is now used in the classification step.

3) GENETIC ALGORITHM

A Genetic Algorithm (GA) is constructed to evolve the evolutionary process. GA helps by allowing the current population to reproduce and generate the next level of children. The basic steps for implementing the GA are described below:

- Initially, set the no. of children in each generation (G), mutation probability (MP), and stopping criteria (SC).
- Initialize the random set of binary chromosomes
- loop
- do for each chromosome
- Train the model and calculate its fitness.
- for reproduction, $i = 1$ to $G/2$
 - do select 2 chromosomes based on fitness
 - do the crossover over them
 - do mutation using probability (MP) and generate new child chromosomes
- end till stopping criteria (SC) achieved

The basic of the GA algorithm, which is used in this work for selecting features, is described in [31].

4) INDEPENDENT COMPONENT ANALYSIS (ICA)

ICA is defined as finding a linear transformation that maximizes non-Gaussianity of statistically independent components [30] given by a matrix M as given in Equation below i.e., so that the random variable $y_i, i = 1, 2, \dots, n$ are as independent as possible.

$$\begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_{n-1}(t) \\ y_n(t) \end{bmatrix} = M \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{m-1}(t) \\ x_m(t) \end{bmatrix} \quad (49)$$

where t is the time or sample index, M is some unknown matrix. ICA now consists of estimating both M and $x_i(t)$ when we only observe the $y_i(t)$. The matrix M is used to project the feature vector into independent components. The method used for feature selection using ICA in this work is well defined in [30].

E. CLASSIFICATION

This subsection deals with the final processing step for model building using machine learning. The last phase of the machine learning model is the classification of the dataset. Classification is the procedure for organizing the input patterns into equivalent classes and providing labels for them. The selection of a suitable classifier requires consideration of many factors, like classification accuracy, algorithmic performance, computational resources, etc. Classification is performed by using the selective feature set using the feature selection mechanism. These selective features are used for learning and prediction of the dataset by the model, called training and testing of the machine model. Training and testing are the two main domains for any classification model. In classification, the accuracy mainly depends upon how well

the classification network is trained. A better-trained classifier model always has the advantage of predicting outputs and generating class labels that are free from over-fitting and under-fitting conditions.

In this work, several machine learning-based classification algorithms like K-Nearest Neighbour (KNN), multi-class Support Vector Machine (m-SVM), and Neural Network (NN) is used for classification purposes. Various experiments are conducted based on the selected features set using CVM, GA, and ICA methods. Each of the classifiers has experimented with each of the selected feature set using the above three defined feature selection methods. All three of the above classifiers are multi-class classifiers, which means that classifiers can classify the given data pattern into two or more classes. Also, all the classifiers are from different domains, like KNN, which is a linear, lazy-learning classifier, m-SVM is a non-linear classifier, and NN is a probabilistic classifier based on the initial selection of the weights, activation function, learning rate, and the number of hidden layers.

1) K-NEAREST NEIGHBOUR (KNN) CLASSIFIER

A supervised nearest neighbor-based classifier named K-NN is used for the result analysis of the proposed approach. Dissimilarity is used as criteria for neighborhood selection. The square of Euclidean distance is a metric used in dissimilarity selection. KNN is based on the closest training samples in feature space thus, KNN classification is performed by starting with the $k = 2$ nearest neighbors and regularly increasing the value of K until classification accuracy no longer improves. There is no proper justification present for the initial selection of the best value of 'k'. This 'k' value must be defined in a specific range, i.e.

$$0 < k \leq n; \quad \text{Where, } n = \text{no. of items in a dataset} \quad (50)$$

If $k = 1$, then the algorithm will search for the very first nearest neighbor who is closest and return the class label of the unknown sample as the same as its nearest neighbor. While if $k = m$, where $m \in [2, n]$ then the algorithm will search for the first m closest neighbors from the sample and return the class label of the neighbor who has the maximum majority of class labels among the 'm' nearest samples.

One of the main and critical parameters which should be kept in consideration before using KNN is a selection of the value of 'k', i.e., nearest neighbors for any particular application domain. In general, the value of the nearest neighbor should always be greater or equal to the number of unique class labels present in the training dataset. Also, the value of k must be kept odd in number in respect of making the algorithm bias-free.

2) MULTI-CLASS SUPPORT VECTOR MACHINE (M-SVM) CLASSIFIER

The Support Vector Machine (SVM) is a popular binary non-linear classifier that is used in the majority of pattern classification problems. SVM is used in problems where a binary distinction is required. In problems where more than

two classes are present for classification, the multi-class SVM (m-SVM) classifier is used. For multi-class classification using SVM, several methods are used, which are given as:

- Ranking based multi-class Classification
- One –Vs-Other Classification

In the ranking-based mechanism, one SVM-based decision function is used to classify all the instances according to their particular class label. While in other one-vs.-other classification mechanisms, including one-against-all classification, in which one class instance is separated from all the other instances of the other classes, and pairwise classification, in which two instances are taken from classes which are separated by other instances of the classes.

In this paper, a one-against-all classification mechanism is used to classify the instances of various brain tumor dataset images using m-SVM. The experiment is performed initially to identify the number of correct instances using one-against-all m-SVM. Then the incorrect classified images are again being processed to find out the number of incorrect classified images for each type of tumor dataset.

3) NEURAL NETWORK (NN) CLASSIFIER

A neural network classifier is a well-known state-of-art classifier that is based on the neural configuration of the brain. From the classification perspective, they process one input sample at a time and learn the network by comparing the obtained class label of the sample to the original supervised sample known previously. The error rate obtained, the difference between the actual result and the obtained result, feeds backward to update the weighting factor. The whole process loops till the error rate comes within the limits defined initially.

A neural network consists of several modules:

- A set of inputs, (X_i) and initial weights, (W_i)
- An activation function, (ϕ) that sums the weight
- An output, (Y_i)

An input layer of the network includes the data values, selected features through a feature selection algorithm in our case, which is given as an input to the next layer known as the hidden layer. Each of the input layers is connected with each of the hidden layers using some initial weights associated with them. The last layer is the output layer in which each node of the layer represents the class label. In this paper, a three-layered neural network architecture is used, having a variable number of nodes depending upon the type of feature selection algorithm used. The number of selected features represents the number of input nodes in the neural network. The number of hidden layers is used in the range of 15-50, again depending upon the feature selection algorithm.

F. 2-FOLD CROSS-VALIDATION

During the training part of the above-mentioned classifier in the proposed work, a 5*2 fold cross-validation is performed on the malignant brain tumor MR image ROI's dataset.

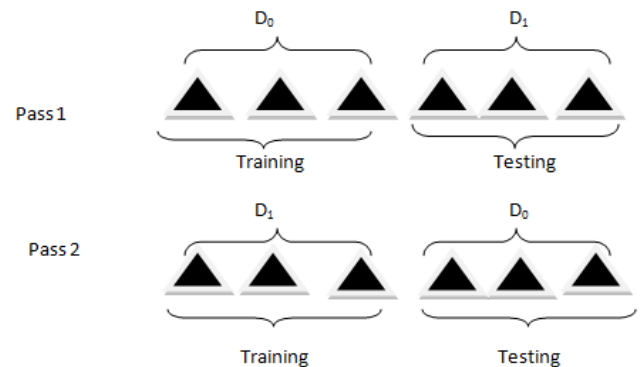


FIGURE 2. 2-fold cross-validation approach.

Cross-validation helps in removing bias between the dataset images. The purpose of cross-validation is to characterize a dataset to learn the model in the training phase to limit problems such as over-fitting and give an insight into how the model will simplify to an independent data set.

Here, in this work, a whole malignant brain neoplasm dataset 'D' is divided into two equal subsets i.e. D_0 and D_1 . Firstly, the training of the classification model is being done using D_0 and the testing of the model using D_1 . Then the training and testing datasets are swapped, i.e., training of the model using D_1 and testing using D_0 . Figure 2 shows the 2-fold cross-validation approach. Similarly, the process iterates 5 times with two classification accuracy results on each run. The datasets D_0 and D_1 are selected randomly on every run, which helps the model learn about almost all the dataset images. This also results in the removal of the biasing among the dataset images. The final accuracy of the dataset is claimed as the mean of all the accuracies obtained in each run.

IV. RESULT AND ANALYSIS

This subsection describes the experimental results and the analysis carried out based on experimental results. In the proposed work, a diverse feature set based on frequency-based features, texture features, intensity features, shape features, contour features, and moment features are used for extracting significant features from the MR imaging dataset, which consists of 5 types of malignant brain tumors named Central NeuroCytoma (CNC), GlioblastomaMultiforme (GBM), Gliomas (GLI), Intra Ventricular Malignant Mass (IVMM), and Metastasis (MTS). There are 110 malignant brain tumor suspects are taken into consideration. From these patients, a dataset of 660 malignant tumor images is generated. Among the generated dataset, an overall of 1160 Regions of Interest (ROI's) are extracted, which are taken into consideration. These ROIs are classified into 760 malicious regions and 400 normal regions (NR). The numbers of samples taken into consideration for an experiment are given as CNC-133, GBM-160, GLI-155, IVMM-152, MTS-160, and NR-400.

TABLE 3. Feature selection Using CVM, ICA, GA.

Input Dataset	CVM	ICA	GA
Central NeuroCytoma	115	55	50
GlioblastomaMultiforme	108	47	47
Gliomas	110	53	46
Intra Ventricular Malignant Mass	121	61	50
Metastasis	109	58	48
Normal	120	61	50

Here, 5 × 2 fold cross-validations are applied to the dataset where the dataset is divided equally for training and testing. In each fold, randomly dataset is divided and then the model is trained by using training dataset. Later, in another fold, this splitting is reversed, and thus model is now trained with testing dataset, which now act as training. The whole methodology is repeated by applying three different feature selection mechanisms (CVM, GA, and ICA) and three different classifiers (KNN, mSVM, and NN).

The results of various feature selection algorithms are shown in Table 3. The first column shows the type of MR image which is taken into consideration. The second column represents the achieved results for various feature selection algorithms based on the input feature set given as an input. The first algorithm, CVM, is based on the parameter of cumulative variance and t-test with a 95% confidence level, while the ICA algorithm is based upon finding the independent components based on the information present in features. GA is based upon the fitness value of the parameter and number of iterations which is given manually during the execution of the program. Here, the number of iterations on which GA runs is 100. Table 2 shows the detailed results of each algorithm with the number of selected features out of a total number of features (298). These selected features are used for machine learning and classification.

A. EVALUATION METRICS

To evaluate the performance of the classifier, several statistical measurement parameters named precision, recall, accuracy, and F-measure are used for evaluating the performance of brain tumor classification. Precision for a class is defined as the fraction of the total number of images that are correctly classified to the total number of images that are classified into the class (sum of True Positives (TP) and False Positives (FP)). The recall is the fraction of the total number of correctly classified images to the total number of images that belong to a class (the sum of True Positives and False

TABLE 4. Statistical performance evaluation metrics.

Metric	Equation
Precision (P)	$\frac{TP}{TP + FP}$
Recall (R)	$\frac{TP}{TP + FN}$
Accuracy (A)	$\frac{TP + TN}{TP + TN + FP + FN}$
F-measure (F)	$\frac{2RP}{R + P}$

Negatives (FN)). F-measure is the combination of both precision and recall. The F-measure is used to report the performance of classifiers for tumor classification. The statistical evaluation of this is shown in Table 4.

For a particular class or an individual class performance analysis through each classifier, the accuracy of each class is identified. The classification for a particular class is defined as the percentage of correctly classified samples over the whole class. Mathematically, it is given as:

$$Accuracy = \frac{\sum_i TP(i)}{\sum_i Class(i)} \times 100 \tag{51}$$

In the proposed work, three experiments are performed based on the selected features through three feature selection algorithms. The first experiment is done using a KNN classifier, the second with a multi-class SVM classifier, and the third with the neural network. The corresponding experimental results are shown in Table 5, Table 6, and Table 7. Each of the tables represents the confusion matrix obtained during the experimental run.

In the confusion matrix, each type of malignant tumor is denoted as T1-Central NeuroCytoma, T2-Glioblastoma Multiforme, T3-Gliomas, T4-Intra Ventricular Malignant Mass, T5-Metastasis, and T6-Normal Regions.

B. EXPERIMENT 1

The first experiment is performed with a KNN classifier in conjunction with three feature selection algorithms, i.e., CVM, GA, and ICA, over the whole dataset of 1160 malignant brain MR ROI images. The dataset has six classes named Central NeuroCytoma (CNC), GlioblastomaMultiforme (GBM), Gliomas (GLI), Intra Ventricular Malignant Mass (IVMM), Metastasis (MTS), and Normal regions (NR). The numbers of samples taken into consideration for the experiment are given as CNC-133, GBM-160, GLI-155, IVMM-152, MTS-160, and NR-400. This whole dataset is now divided into the training set, which has 80% of the images, and the testing set, which includes the remaining 20% of the images from the dataset. Over these samples and three feature selection algorithms, the experiment is performed on several values of 'k'. The best accuracy achieved is at k=7, which is shown in Table 5. The accuracy for each particular class for various feature selection algorithms is CNC-92%,

TABLE 5. Confusion matrix for KNN algorithm with an accuracy.

Predicted Class	Supervised Class prediction (Radiologist)																	
	CVM						ICA						GA					
	T1	T2	T3	T4	T5	T6	T1	T2	T3	T4	T5	T6	T1	T2	T3	T4	T5	T6
T1	123	7	1	0	9	13	112	10	6	11	3	17	118	3	1	7	11	17
T2	3	138	4	0	1	17	7	130	1	7	9	14	3	130	3	5	11	17
T3	0	2	142	7	7	9	3	5	137	5	9	20	6	11	144	6	4	14
T4	4	0	1	140	4	19	4	9	7	125	2	15	3	6	4	132	0	17
T5	1	6	3	2	137	5	2	6	3	1	132	15	0	4	0	2	131	12
T6	2	7	4	3	2	337	5	0	1	3	5	319	3	6	3	0	3	323
Accuracy	92	86	91	92	85.6	84	84	81	88	82	82.5	79.75	88.7	81	93	86	81	80.75
Avg. Accuracy	88.43						82.87						85					

TABLE 6. Confusion matrix for mSVM algorithm with an accuracy.

Predicted Class	Supervised Class prediction (Radiologist)																	
	CVM						ICA						GA					
	T1	T2	T3	T4	T5	T6	T1	T2	T3	T4	T5	T6	T1	T2	T3	T4	T5	T6
T1	127	4	2	3	5	9	119	6	4	3	6	9	122	8	4	3	0	8
T2	3	143	2	1	2	11	3	136	4	6	3	11	3	140	4	5	3	11
T3	0	3	144	1	2	9	3	5	137	2	5	8	3	4	139	2	2	9
T4	1	3	4	144	2	4	2	7	3	136	3	9	2	4	3	139	2	8
T5	1	4	0	2	148	6	2	3	6	5	141	13	3	3	3	3	150	7
T6	1	3	3	1	1	361	4	3	1	0	2	350	0	1	2	0	3	357
Accuracy	95.4	89.3	92.9	94.7	92.5	90.25	89	85	88.3	89.4	88.1	87.5	91.7	87.5	89.6	91.4	93.7	89.2
Avg. Accuracy	92.5						87.88						90.52					

GBM-86%, GLI-91%, IVMM-92%, MTS-85.6%, and NR-84% when using CVM as a feature selection algorithm. The average accuracy gained is 88.43%, which is greater than other feature selection algorithms like ICA (82.87%) and GA (85.07%) as shown in Table 5. The experimental result shows that the order performance of three feature selection algorithms is CVM > GA > ICA.

C. EXPERIMENT NO. 2

The second experiment is performed with a non-linear multi-class SVM classifier (mSVM). The dataset is kept constant throughout this work as discussed in experiment 1. When multi-class SVM is used as a classifier with three feature selection algorithms, then an increase in the accuracy of particular classes is observed (as shown in Table 6). The classification accuracy achieved for various classes using CVM

as a feature selection approach is CNC-95.4%, GBM-89.3%, GLI-92.9%, IVMM-94.7%, MTS-92.5%, and NR-90.25%, which is greater than for MTS using GA, which is presented clearly in Table 6. Still, it is seen that the overall average classification accuracy of the mSVM classifier is still higher for CVM-92.5% than for other feature selection algorithms like ICA-87.88% and GA-90.52%. There is no change seen in the order of performance of the three feature selection algorithms, i.e., CVM > GA > ICA.

D. EXPERIMENT NO. 3

The third experiment is performed with a neural network (NN) classifier. The experimental result is presented in Table 7. The accuracy achieved using the NN classifier with the CVM feature selection algorithm is CNC-92.7%, GBM-96.2%, GLI-95.4%, IVMM-96.7%, MTS-95%, and

TABLE 7. Confusion matrix for NN algorithm with an accuracy.

Predicted Class	Supervised Class prediction (Radiologist)																	
	CVM						ICA						GA					
	T1	T2	T3	T4	T5	T6	T1	T2	T3	T4	T5	T6	T1	T2	T3	T4	T5	T6
T1	130	2	3	1	2	4	123	6	4	3	4	9	125	0	2	1	1	7
T2	0	154	0	0	3	7	0	142	7	6	7	12	1	147	4	1	0	12
T3	1	2	148	0	1	3	4	2	137	6	7	11	1	4	142	7	3	9
T4	2	1	2	147	2	4	3	3	4	132	1	8	3	7	3	140	4	3
T5	0	1	1	3	152	5	3	4	2	2	141	10	2	1	2	3	148	6
T6	0	0	1	1	0	377	0	3	1	3	0	350	1	1	2	0	5	363
Accuracy	97.7	96.2	95.4	96.7	95	94.2	92	88	88	86	88	87.5	93.9	91.8	91.6	92.1	92.5	90.7
Avg. Accuracy	95.86						88.25						92.10					

NR-94.2%. The average accuracy gained is 95.86%, which is more than any other feature selection algorithm presented in Table 7. The order performance of the feature selection algorithm is still the same as shown in earlier experiments.

Based on the experimentation results it is identified that the proposed approach is providing satisfactory results as compared with other methods. One of the main reasons for this is towards use of diversified feature set. In this work, six different feature domains are used to extract the large set of features in the feature pool. It includes an overall of 298 features having texture, spectral, shape, and moment information. These features are further filtered out using the proposed CVM method, which uses the cumulative variance of the extracted features. Thus, using the standard Joliffe’s B4 method threshold, the filtered features are relevant and informative to capture different information.

E. EXPERIMENT NO. 4

The proposed methodology is also experimented with one of the multiclass brain tumor dataset having 3 classes i.e. Meningioma, Glioma, and pituitary tumor. This dataset is collected from the online platform which consists of 3064 brain MR scans [49]. The dataset is having T1-weighted image slices of 233 individuals with the distribution as Gliomas-1426, Meningiomas-708, and pituitary-930 images. In the literature, there exists many of the research contributions on the similar dataset [45]–[48]. These research contributions uses the CNN based deep models for classification. The same is being shown in Table 1 with the predicted accuracy.

In this experimentation setup, the proposed methodology is experimented on the similar dataset having 3 classes of brain tumors using machine learning environment. The experimentation results of the proposed method gives significant

accuracies when using with the three machine learning classifiers. The gained in the classification accuracies are given as Gliomas-{89.63% with KNN, 92.84% with m-SVM, and 95.45% with NN}, Meningiomas-{91.66% with KNN, 92.33% with m-SVM, and 94.66% with NN}, and pituitary-{90.53% with KNN, 91.67% with m-SVM, and 94.46% with NN} and the same is shown in Table 9. The experimentation results were found significant enough as compared with other feature selection algorithms.

V. DISCUSSION

This paper uses three feature selection algorithms, i.e., proposed CVM, ICA, and GA, with three well-known state-of-the-art classifiers, KNN, mSVM, and NN. The experimental result of each classifier with every feature selection algorithm is presented in Table 5, Table 6, and Table 7. The experimental results show that the CVM feature selection algorithm gives better accuracy results in respect of each class of malignant brain tumor prediction concerning other feature selection algorithms like ICA and GA. The overall order of the feature selection algorithm in respect of accuracy result is found as CVM > GA > ICA. Also, it has been noticed that the performance level of the classifier is also varied. The order of the classifier in NN > mSVM > KNN as the highest average accuracy achieved using NN is 95.86% using CVM for malignant tumor class label prediction. Table 8 will summarize the analysis result of the three-feature selection algorithm with three classifiers. Each row in Table 8 represents the feature selection algorithm and each column represents the classifier. The values in bold displayed in the first row of Table 7 represent that the CVM method gets the best accuracy concerning other algorithms. In contrast, the last column bold indicated values represent that NN gives the best classification results among all classifiers.

TABLE 8. Overall summarized results of three feature selection algorithms.

Feature Selection Algorithm	Classifier Accuracy (%)		
	KNN	mSVM	NN
CVM	88.43	92.5	95.86
GA	85.07	90.52	92.10
ICA	82.87	87.88	88.25

TABLE 9. Experimentation results on 3-class brain tumor dataset [49].

Tumor class	Classifier Accuracy (%) (with CVM)			Classifier Accuracy (%) (with GA)			Classifier Accuracy (%) (with ICA)		
	KNN	mSVM	NN	KNN	mSVM	NN	KNN	mSVM	NN
Gliomas	89.63	92.84	95.45	78.55	79.46	80.15	84.12	87.52	90.54
Meningiomas	91.66	92.33	94.66	76.48	76.57	79.97	85.82	88.78	90.03
Pituitary	90.53	91.67	94.46	78.55	77.85	79.97	84.12	86.66	89.97

These experimental results could help the radiologist perform better analyses of the malignant brain tumor suspect using a computer-aided system with the help of various feature extraction techniques described in this paper. The extracted features are free from rotation, scaling and include the properties of both spatial and frequency domains. Over this, three feature selection algorithms and three classifiers are presented, which help a radiologist better analyze malignant tumors. In the future, frameworks based on hybrid decision support techniques may also be explored for the better brain tumor prediction.

The main advantage of the proposed methodology is to find the most relevant and non-redundant features by using the computationally efficient proposed feature selection algorithm called CVM. It helps to gain the best in-class accuracies when experimented with three state-of-art classifiers like KNN, mSVM, and Neural Network. This gives the new insight in the machine learning framework for problem solving. On the other hand, the limitation of the proposed work is its scalability. The proposed methodology is not tested with some of the open accessible imaging datasets having abnormality characteristics. However, major the extension of the proposed methodology with deep learning models can also be one of the limitations that need to be taken care of in near future.

VI. CONCLUSION

In this paper, a study was performed on 110 patients having high-grade malignant brain tumors like Central Neuro-Cytoma (CNC), GlioblastomaMultiforme (GBM), Gliomas (GLI), Intra Ventricular Malignant Mass (IVMM), and Metastasis (MTS). From these 110 patients, the T1-Weighted

post-contrast image specification dataset was generated, containing 660 malignant tumor images. Among the generated dataset, 1160 Regions of Interest's (ROI's) are extracted, which are taken into consideration. These ROIs are classified into 760 malicious regions and 400 normal regions. The whole dataset is categorized into six broad domains, including five classes for high-grade malignant tumors and one normal class. Over the entire dataset, a 5 x2 cross-validation mechanism is used to partition the dataset into training (50%) and testing (50%) subsets of images.

The main contributions in this work is defined in two folds i.e. the designing of the new feature selection algorithm called CVM and the experimentation of the multiclass brain imaging dataset using benchmark models in machine learning environment. The proposed CVM algorithm is based on variance analysis and thus will retain the features having high cumulative variance up to the threshold limit 99.5%. While in the experimentation part, the selected features are used to train the machine learning classification models and the comparative result analysis is presented by using three benchmark classifiers and three feature selection algorithms.

The experiments were performed by incorporating the diversified features set containing six different domain features. It results in extracting 298 features from the dataset used for classification purposes in the machine learning environment. These features were filtered out using the proposed CVM feature selection algorithm based on cumulative variance. Later, the model training uses all the filtered features using three state-of-art classifiers, i.e., KNN, mSVM, and NN. The average gain in the accuracy is 88.43%, 92.5%, and 95.86%, respectively, using the CVM method. Based on

the comparative result, it was also found that the proposed approach is outperforming as compared with other feature selection algorithms.

In the future, the study can be extended towards the use of the deep learning models for the classification of the brain tumor images. In the literature there are significant works on the use of the deep learning models for the binary classification of the brain tumor images. However, the limited work was published on multiclass classification using deep learning. This work can be further extended towards the use of new deep learning model for the multiclass classification of the brain tumor types. Moreover, the work can also extendable enough to include the more number of images in the dataset. Also, the algorithm and the models can be experimented on some gold standard benchmark datasets in future.

ACKNOWLEDGMENT

The authors are thankful to the Sawai Man Singh (SMS) Medical College for providing them the original brain tumor images in DICOM format. They would also like to thank Dr. Sunil Jakhar, Department of Radiology, for helping them with providing the MR images and verifying the results.

REFERENCES

- [1] *About Brain Tumors: A Primer for Patients and Caregivers*. Accessed: 2015. [Online]. Available: <https://www.abta.org/secure/about-brain-tumors-a-primer.pdf>
- [2] S. Wang and R. M. Summers, "Machine learning and radiology," *Int. J. Med. Image Anal.*, vol. 16, no. 5, pp. 933–951, 2012.
- [3] P. R. Velthuizen, O. L. Hall, and P. L. Clarke, "MRI feature extraction using genetic algorithm," in *Proc. 18th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 3, Nov. 1996, pp. 1138–1139.
- [4] H. J. Lee, K. Y. Seong, H. C. Chang, and J. Park, "Fourier-based shape feature extraction technique for computer-aided B-mode ultrasound diagnosis of breast tumor," in *Proc. Annu. Int. Conf. Eng. Med. Biol. Soc. (EMBC)*, Sep. 2012, pp. 6551–6554.
- [5] H. Kalbkhani, M. G. Shayesteh, and B. Zali-Vargahan, "Robust algorithm for brain magnetic resonance image (MRI) classification based on GARCH variances series," *Biomed. Signal Process. Control*, vol. 8, no. 6, pp. 909–919, Nov. 2013.
- [6] C. Arizmendi, A. Vellido, and E. Romero, "Classification of human brain tumours from MRS data using discrete wavelet transform and Bayesian neural networks," *Expert Syst. Appl.*, vol. 39, no. 5, pp. 5223–5232, Apr. 2012.
- [7] M. Kaushik, R. Sharma, and A. Vidyarthi, "Analysis of spatial features in CBIR system," *Int. J. Comput. Appl.*, vol. 54, no. 17, pp. 11–15, Sep. 2012.
- [8] M. Saritha, K. P. Joseph, and A. T. Mathew, "Classification of MRI brain images using combined wavelet entropy based spider web plots and probabilistic neural network," *Pattern Recognit. Lett.*, vol. 34, no. 16, pp. 2151–2156, Dec. 2013.
- [9] A. Vidyarthi and N. Mittal, "Comparative study for brain tumor classification on MR/CT images," in *Proc. 3rd Int. Conf. Soft Comput. Problem Solving (SocPros)*, 2013, pp. 889–897.
- [10] S. D. Guru and S. Manjunath, "Texture feature and KNN in classification of flower images," *Int. J. Control Automat. Special Issue Recent Trends Image Process. Pattern Recognit.*, vol. 1, pp. 21–29, Aug. 2010.
- [11] S. Ahmed and K. M. Iftikharuddin, "Efficacy of texture, shape and intensity feature fusion for posterior-fossa tumor segmentation in MRI," *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 2, pp. 206–213, Mar. 2011.
- [12] E. S. A. El-Dahshan, T. Hosny, and A. B. M. Salem, "Hybrid intelligent techniques for brain tumor classification," *Int. J. Digit. Signal Process.* vol. 20, no. 2, pp. 433–441, 2010.
- [13] P. Georgiadis, D. Cavouras, I. Kalatzis, A. Daskalakis, G. C. Kagadis, K. Sifaki, M. Malamas, G. Nikiforidis, and E. Solomou, "Improving brain tumor characterization on MRI by probabilistic neural networks and non-linear transformation of textural features," *Comput. Methods Programs Biomed.*, vol. 89, no. 1, pp. 24–32, Jan. 2008.
- [14] D. Jude Hemanth, C. K. S. Vijila, A. I. Selvakumar, and J. Anitha, "Performance improved iteration-free artificial neural networks for abnormal magnetic resonance brain image classification," *Neurocomputing*, vol. 130, pp. 98–107, Apr. 2014.
- [15] A. Bianchi, J. V. Miller, E. T. Tan, and A. Montillo, "Brain tumor segmentation with symmetric texture and symmetric intensity-based decision forests," in *Proc. IEEE 10th Int. Symp. Biomed. Imag.*, Apr. 2013, pp. 748–751.
- [16] J. Sachdeva, V. Kumar, I. Gupta, N. Khandelwal, and C. K. Ahuja, "Segmentation, feature extraction, and multiclass brain tumor classification," *J. Digit. Imag.*, vol. 26, no. 6, pp. 1141–1150, Dec. 2013.
- [17] N. H. Rajni and R. Bhavani, "Automatic classification of CT brain images using ANN, KNN and SVM," *Int. J. Artif. Intell. Soc.*, vol. 29, no. 1, pp. 97–102, 2013.
- [18] V. Gupta and K. S. Sagale, "Implementation of classification system for brain cancer using backpropagation network and MRI," in *Proc. Nirma Univ. Int. Conf. Eng. (NUICONE)*, Dec. 2012, pp. 1–4.
- [19] C. Arizmendi, A. Vellido, and E. Romero, "Binary classification of brain tumours using a discrete wavelet transform and energy criteria," in *Proc. IEEE 2nd Latin Amer. Symp. Circuits Syst. (LASCAS)*, Feb. 2011, pp. 1–4.
- [20] Q. Wang, E. K. Liacouras, E. Miranda, U. S. Kanamalla, and V. Megalooikonomou, "Classification of brain tumors using MRI and MRS data," *Proc. SPIE* vol. 6514, pp. 231–238, Mar. 2007.
- [21] M. A. Weber, S. Zoubaa, M. Schlieter, E. Jüttler, H. B. Huttner, K. Geletneký, C. Ittrich, M. P. Lichy, A. Kroll, J. Debus, and F. L. Giesel, "Diagnostic performance of spectroscopic and Perfusional MRI for distinction of brain tumors," *J. Neurol.*, vol. 66, no. 12, pp. 1899–1906, 2006.
- [22] Y.-D. Cho, G.-H. Choi, S.-P. Lee, and J.-K. Kim, "1H-MRS metabolic patterns for distinguishing between meningiomas and other brain tumors," *Magn. Reson. Imag.*, vol. 21, no. 6, pp. 663–672, Jul. 2003.
- [23] J. Iivarinen and A. Visa, "An adaptive texture and shape based defect classification," in *Proc. 14th Int. Conf. Pattern Recognit.*, vol. 1, 1998, pp. 117–122.
- [24] R. M. Haralick, K. Shanmugam, and H. Dinstein, "Texture features for image classification," *IEEE Trans. Syst. Man Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973.
- [25] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn.*, vol. 3, pp. 1157–1182, Mar. 2003.
- [26] M. A. Weber, S. Zoubaa, M. Schlieter, E. Jüttler, H. B. Huttner, K. Geletneký, C. Ittrich, M. P. Lichy, A. Kroll, J. Debus, F. L. Giesel, M. Hartmann, and M. Essig, "Diagnostic performance of spectroscopic and perfusion MRI for distinction of brain tumors," *Neurology*, vol. 66, no. 12, pp. 1899–1906, Jun. 2006.
- [27] Y.-D. Cho, G.-H. Choi, S.-P. Lee, and J.-K. Kim, "1H-MRS metabolic patterns for distinguishing between meningiomas and other brain tumors," *Magn. Reson. Imag.*, vol. 21, no. 6, pp. 663–672, Jul. 2003.
- [28] S. Higano, X. Yun, T. Kumabe, M. Watanabe, S. Mugikura, A. Umetsu, A. Sato, T. Yamada, and S. Takahashi, "Malignant astrocytic tumors: Clinical importance of apparent diffusion coefficient in prediction of grade and prognosis," *Radiology*, vol. 241, no. 3, pp. 839–846, Dec. 2006.
- [29] A. Vidyarthi and N. Mittal, "Utilization of shape and texture features with statistical feature selection mechanism for classification of malignant tumors in MR images," in *Proc. 48th DGBMT Annu. Conf. (BMT)*, vol. 59, 2014, pp. 155–159.
- [30] *Moments in Mathematics*. Accessed: 2015. [Online]. Available: http://en.wikipedia.org/wiki/Moment_mathematics
- [31] M. N. Prasad, A. Sowmya, and I. Koch, "Feature subset selection using ICA for classifying emphysema in HRCT images," in *Proc. 17th Int. Conf. Pattern Recognit. (ICPR)*, 2004, pp. 515–518.
- [32] C. L. Huang and C. J. Wang, "A GA-based feature selection and parameter optimization for support vector machine," *J. Expert Syst. Appl.*, vol. 31, no. 2, pp. 231–240, 2006.
- [33] G. Garg and R. Garg, "Brain tumor detection and classification based on hybrid ensemble classifier," 2021, *arXiv:2101.00216*.
- [34] K. Jaeyong, U. Zahid, and G. Jeonghwan, "MRI-based brain tumor classification using ensemble of deep features and machine learning classifiers," *Sensors*, vol. 21, no. 6, p. 2222, 2021.

- [35] M. R. Ismael and I. Abdel-Qader, "Brain tumor classification via statistical features and back-propagation neural network," in *Proc. IEEE Int. Conf. Electro Inf. Technol. (EIT)*, May 2018, pp. 0252–0257.
- [36] N. V. Shree and T. N. R. Kumar, "Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network," *Brain Informat.*, vol. 5, no. 1, pp. 23–30, 2018.
- [37] A. Vidyarthi and N. Mittal, "Performance analysis of Gabor-wavelet based features in classification of high grade malignant brain tumors," in *Proc. 39th Nat. Syst. Conf. (NSC)*, Dec. 2015, pp. 1–6.
- [38] M. P. Arakeri and G. R. M. Reddy, "An intelligent content-based image retrieval system for clinical decision support in brain tumor diagnosis," *Int. J. Multimedia Inf. Retr.*, vol. 2, no. 3, pp. 175–188, Sep. 2013.
- [39] D. Sridhar and I. V. M. Krishna, "Brain tumor classification using discrete cosine transform and probabilistic neural network," in *Proc. Int. Conf. Signal Process. Image Process. Pattern Recognit.*, Feb. 2013, pp. 92–96.
- [40] A. P. Nanthagopal and R. S. Rajamony, "Classification of benign and malignant brain tumor CT images using wavelet texture parameters and neural network classifier," *Int. J. Comput. Vis.*, vol. 16, no. 1, pp. 19–28, 2012.
- [41] P. Rajendran, M. Madheswaran, and K. Naganandhini, "An improved pre-processing technique with image mining approach for the medical image classification," in *Proc. 2nd Int. Conf. Comput., Commun. Netw. Technol.*, Jul. 2012, pp. 1–7.
- [42] M. F. B. Othman, N. B. Abdullah, and N. F. B. Kamal, "MRI brain classification using support vector machine," in *Proc. IEEE 4th Int. Conf. Model. Simul. Appl. Optim. (ICMSAO)*, Apr. 2011, pp. 1–4.
- [43] A. Al-Badarnah, H. Najadat, and A. M. Alraziqi, "A classifier to detect tumor disease in MRI brain images," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining*, Aug. 2012, pp. 784–787.
- [44] E. I. Zacharaki and V. G. Kanas, "Investigating machine learning techniques for MRI based classification of brain neoplasm," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 6, no. 6, pp. 821–828, 2011.
- [45] E. Irmak, "Multi-classification of brain tumor MRI images using deep convolutional neural network with fully optimized framework," *Iranian J. Sci. Technol., Trans. Elect. Eng.* vol. 45, no. 3, pp. 1015–1036, 2021.
- [46] F. J. Díaz-Pernas, M. Martínez-Zarzuola, M. Antón-Rodríguez, and D. González-Ortega, "A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network," *Healthcare*, vol. 9, no. 2, p. 153, 2021.
- [47] X. Gu, Z. Shen, J. Xue, Y. Fan, and T. Ni, "Brain tumor MR image classification using convolutional dictionary learning with local constraint," *Front. Neurosci.* vol. 15, May 2021, Art. no. 679847.
- [48] M. Toğaçar, B. Ergen, and Z. Cömert, "Tumor type detection in brain MR images of the deep model developed using hypercolumn technique, attention modules, and residual blocks," *Med. Biol. Eng. Comput.*, vol. 59, no. 1, pp. 57–70, 2021.
- [49] J. Cheng, *Brain Tumor Dataset*. Accessed: Mar. 28, 2022. [Online]. Available: <https://doi.org/10.6084/m9.figshare.1512427.v5>
- [50] P. Tiwari and M. Melucci, "Binary classifier inspired by quantum theory," in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, no. 1, pp. 10051–10052, Jul. 2019.
- [51] A. V. Reddy, C. Krishna, P. K. Mallick, S. K. Satapathy, P. Tiwari, M. Zymbler, and S. Kumar, "Analyzing MRI scans to detect glioblastoma tumor using hybrid deep belief networks," *J. Big Data*, vol. 7, no. 1, pp. 1–17, 2020.

• • •