IMAGE & SIGNAL PROCESSING

Glioma Tumor Grade Identification Using Artificial Intelligent Techniques

Ahammed Muneer K. V.[1] · V. R. Rajendran[2] · Paul Joseph K.[1]

Abstract

Computer aided diagnosis using artificial intelligent techniques made tremendous improvement in medical applications especially for easy detection of tumor area, tumor type and grades. This paper presents automatic glioma tumor grade identification from magnetic resonant images using Wndchrm tool based classifier (Weighted Neighbour Distance using Compound Heirarchy of Algorithms Representing Morphology) and VGG-19 deep convolutional neural network (DNN). For experimentation, DICOM images are collected from reputed government hospital and the proposed intelligent system categorized the tumor into four grades such as low grade glioma, oligodendroglioma, anaplastic glioma and glioblastoma multiform. After preprocessing, features are extracted, optimized and then classified using Windchrm tool where the most significant features are selected on the basis of Fisher score. In the case of DNN classifier, data augmentation is also performed before applying the images into the deep learning network. The performance of the classifiers are analysed with various measures such as accuracy, precision, sensitivity, specificity and F1-score. The results showed reasonably good performance with a maximum classification accuracy of 92.86% for the Wndchrm classifier and 98.25% for VGG-19 DNN classifier. The results are also compared with similar recent works and the proposed system is found to have better performance.

Keywords MRI · Artificial intelligence · Wndchrm · DNN · Glioma grades

Introduction

Today cancer has become a very deadly disease affecting people of all age groups. Biomedical researches in recent decade has shown greater development with accurate results for classifying brain tumor into benign and malignant or cancerous tumors. According to the survey by WHO [1], the brain tumors are classified into different types namely, glioma, meningioma, metastatis, sarcoma etc. Currently, one of the most recent research direction is to further categorize the brain tumor type into four different grades. Recent advances in medical image processing with the application of computer aided diagnosis (CAD), using magnetic resonance images (MRI) facilitates easy detection of tumor portion. Nevertheless, identification and classification of brain tumor types and grades is still a challenging problem. In this paper, we present a novel method based on CAD system for categorizing glioma brain tumor into four different grades.

Many researches have been done for brain tumor categorization. A review of recent literatures is given here. E. A. El Dahshan et al. in [2] introduced a method for classifying brain tumor using hybrid clustering approach. U. Javed et al. [3] used supervised classifier such as SVM for classifying the texture features from MR brain images into different types. M. Nazir in [4] classified the tumor into various types using artificial neural network approach.

✉ Ahammed Muneer K. V.
 ahammedcet@gmail.com

 V. R. Rajendran
 vrajendranclt@yahoo.com

 Paul Joseph K.
 paul@nitc.ac.in

[1] Department of Electrical Engineering, National Institute of Technology Calicut, 673601, Calicut, India

[2] Department of Radiology, Government Medical College Kozhikode, 673008 Calicut, India

Monica Subashini et al. [5] identifies astrocytoma type brain tumor grades using latest image processing techniques. Geethu Mohan et al. [6] performed a survey on MR images for classifying brain tumors into four grades. Ivan I. Maximov et al. [7] used diffusion MRI to differentiate the different levels of glioma malignancy grades. Kevin et al. [8] used invariant textures on grey scale images for grade categorization. A novel method introduced to extract a larger feature set is the Weighted Neighbor Distance using Compound Hierarchy of Algorithms Representing Morphology (Wndchrm) based classifiers. Lior Shamir et al. [9] proposed the new open source classifier called Wndchrm. This tool uses various command utilities for extracting, selecting and classifying the images into their respective classes. Orlov N et al. in [10] used compound features alone from complete Wndchrm feature set to perform multi-classification. B. S. Shajee Mohan et al. in [11] analysed various supervised and semi-supervised classifiers for biological images in content based image retrieval system. The image segmentation can also be performed before the Wndchrm based classification. In this work, hybrid segmentation using combined k-means and fuzzy-c-means is used to segment out the tumor portion effectively [12, 13].

The recent development in deep learning have made it prominent in biomedical applications in various image modalities like MRI, CT etc. Heba Mohsen et al. in [14] proposed a simple deep nueral network for categorizing brain tumors into four types. Harshitha Sharma et al. [15] proposed a novel deep nueral network architecture and compared with existing basic deep convolutional neural networks (DNN) for categorizing gastric carcinoma. Xiaohog W. Gao et al. [16] used computed tomography (CT) images instead of MRI and performed classification based on deep learning networks. Recently, Vivanti. R. Joskowicz. L, Lev-Cohain. N, et al. [17] presented a new automatic liver tumor segmentation method for follow up CT studies using deep neural networks. Maryam M Najafabadi et al. and Eli Gibson et al. in [18, 19] proposed deep learning techniques for big data and medical applications. Automation of brain abnormality classification on MRI images using deep transfer learning technique with ResNet34 model is reported in [20]. A comparative study on the application of multi-resolution analysis for automated brain abnormality detection is done in [21].

Saed Khawaldeh et al. [22] proposed non-invasive method for grade categorization using convolutional neural networks (CNN). Lina Chato et al. [23] used a linear discriminant method for deep learning categorization. V.P. Gladis Pushpa Rathi et al. [24] used a kernel based clustering method and DNN for categorizing the tumor and non-tumor MRI. Ye. C.Z, Yang. J, Geng. D. Y. et al.

in [25] presented a fuzzy rule based neural network to extract degree of malignancy between low grade and high grade brain glioma. Recently, Pang. S., Du. A., Orgun M. A. et al. [26] proposed novel fused convolutional neural network architecture for biomedical image classification making use of combined shallow and deep feature layers. The methodology comprising of CNN and DNN can also be effectively used for different medical images corresponding to various applications such as diagnosis of glaucoma using fundus images [27], age related macular degeneration detection [28], automatic detection of thoracolumbar fractures [29] and segmentation of optic disc, fovea, retinal vasculature, exudates, haemorrhages etc. [30, 31].

Though the work presented in [22] and [24] are focussed on glioma grade categorization, the classification accuracy depends upon the algorithm and methodology we use and also on the data set under experimentation. The major contributions of the proposed work are given as follows:

1. We considered clinically annotated data set for our experimentation and hence the confidence in the training set used is very high.
2. Wndchrm is a powerful tool for biological image analysis. We considered it for feature extraction, selection and classification.
3. We also considered deep learning network such as VGG-19 based on transfer learning technique, where the network model and parameters are fine tuned to suit for glioma grade categorization.
4. We compared the performance of the two presented methodologies and finally compared the performance of proposed intelligent systems with the recent works that used similar architecture and dataset.

The organization of this paper is given as follows. Section "Methods" describes about the methodology giving in detail about the proposed work, "Results" gives details about the experimental analysis and results. Section "Discussion" discusses about the performance analysis of the proposed methodology and the comparison with the related works. Finally, "Conclusions" concludes this paper with future scope.

Methods

This section describes the data set used, methodology adopted and the performance analysis of the classifiers used. The classifiers proposed for glioma tumor grade categorization are Weighted Neighbour Distance using Compound Heirarchy of Algorithms Representing Morphology (Wndchrm tool) and VGG-19 deep convolutional neural network (DNN).

Data set and experimental set up

The glioma type MRI brain tumor data set was collected from Government Medical College, Calicut, India [32]. It comprises MR Images of Axial T2 series in DICOM format. They are collected from around 20 proven case subjects containing both male and female belonging to age group of 30 to 60 years. The images belong to four grades such as, low grade glioma (Grade I), oligodendroglioma (Grade II), anaplastic glioma (Grade III) and glioblastoma mutiform (Grade IV). Figure 1 shows the raw MR images collected from the hospital. The MRI machine and image specifications are given in Table 1. The data set split up details is given in Table 2.

Schematic overview

The schematic overview of the proposed methodology for glioma tumor grade categorization is given in Fig. 2. It includes pre-processing of raw MRI data, two intelligent classifiers and their performance analysis. The experimentation proceeds in two directions. The wndchrm classifier uses the skull stripped and segmented images, where the DNN

Table 1 MRI specifications

Parameter	Value
Magnetic filed strength	1.5 T
Field of view (FOV)	26 cm
Image resolution	256×256
Slice thickness	5 mm

architecture uses the skull stripped and augmented images. Finally, the performance of the two classifiers are analysed.

Preprocessing stage

Prior to classification, the noisy and degrading signals are removed from original image through pre-processing. Generally, various preprocessing step like skull stripping, image filtering and enhancement, segmentation, image augmentation etc. are applied depending on the methodology and application. In our experimentation, skull-stripping and segmentation of tumor portion is being used as the preprocessing step before Wndchrm based classification. Similarly,

Fig. 1 Input DICOM MR image with glioma tumor [32] **a** Grade I **b** Grade II **c** Grade III **d** Grade IV
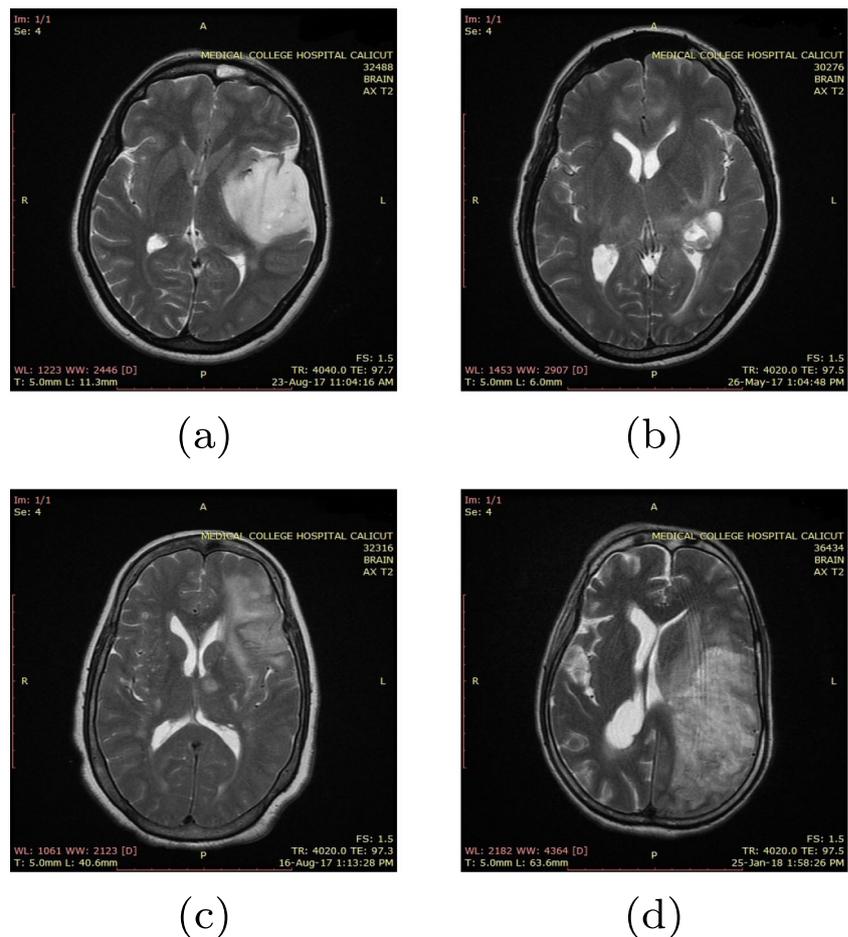

(a)

(b)

(c)

(d)

Table 2 Glioma tumor data set split as training and test

	Grade I	Grade II	Grade III	Grade IV	Total
Testing	39	51	31	47	168
Training	91	118	72	108	389
Total	130	169	103	155	557

the skull-stripping and the data augmentation are performed before DNN based classification.

Skull stripping is performed using Brain Suite tool which uses brain surface extractor algorithm (BSE). Here the skull area including eye and dura is removed and the brain surface is isolated. The segmentation is done using combined k-means and fuzzy-c-means (FCM) clustering algorithms [12, 13]. Here the visual perception of the segmented portion is good. k-means is a type of hard clustering algorithm, in which every single point specifically belongs to a particular cluster. Primarily, the cluster number is fixed as three, since raw MRI comprises of brain tissue, tumor portion and the background. In k-means algorithm, J being the objective function and is minimized as per the Eq. 1.

$$J = \sum_{j=1}^{K} \sum_{i=1}^{N} \left\| \mathbf{x}_i - c_j \right\|^2 \tag{1}$$

Here, K is the number of clusters and N is the number of data points. In FCM clustering algorithm, each data point can belong to different groups, where the cluster probability distribution governs the degree of membership. Here also, the objective function J is minimized as given in Eq. 2

$$J = \sum_{i=1}^{N} \sum_{j=1}^{K} \delta_{ij} \left\| \mathbf{x}_i - c_j \right\|^2 \tag{2}$$

Here, N and K are the total data points and the number of clusters respectively. Also, c_j is the center vector of cluster j and δ_{ij} is the degree of membership for the ith data point of $\mathbf{x}_i$ in cluster j. With relative to the data point $\mathbf{x}_i$, the fuzziness membership to the cluster j is calculated as follows.

$$\delta_{ij} = \frac{1}{\sum_{k=1}^{K} \left(\frac{\left\| \mathbf{x}_i - c_j \right\|}{\left\| \mathbf{x}_i - c_k \right\|} \right)^{\frac{2}{m-1}}} \tag{3}$$

Where m is referred to as the fuzziness coefficient. The centroid c_j for each clusters is calculated as per the Eq. 4.

$$c_j = \frac{\sum_{i=1}^{N} \delta_{ij}^m * \mathbf{x}_i}{\sum_{i=1}^{N} \delta_{ij}^m} \tag{4}$$

Data augmentation is specifically used along with the deep convolutional neural network classifier (VGG-19). This become useful when a large number of dataset is not available for experimentation. The CNN based classifier requires thousands of images and hence to increase the data size, various augmentation steps like resizing, reflection, rotation and translation are carried out.

Methodology 1: Wndchrm tool based classifier

Wndchrm is an open software available to perform classification in biomedical application. Along with the classification, tool also performs feature extraction and selection using various command utilities. Feature extraction is the method of obtaining most informative values from an image. In this methodology, the feature extraction is carried out using Wndchrm tool. This tool uses 11 different algorithms for obtaining around 2059 feature set. Here features are extracted using various transforms such as Fourier, wavelet, Chebyshev and its transforms. The 11 type of features extracted are given in Table 3.

Wndchrm tool extracts a very large feature set but, not all features are informative. Thus there is a need for proper feature selection. The command for parameter selection and optimization of features is detailed in the result section. The tool uses Fisher score ranking for the arrangement of different features in descending order. The distance between a feature vector $\mathbf{x}_t$ and class c is given by

$$d(\mathbf{x}_t, c) = \frac{\sum_{i=1}^{N_c} \left[\sum_{j=1}^{|\mathbf{x}|} F_j (x_{tj} - x_{ij})^2 \right]^p}{N_c} \tag{5}$$
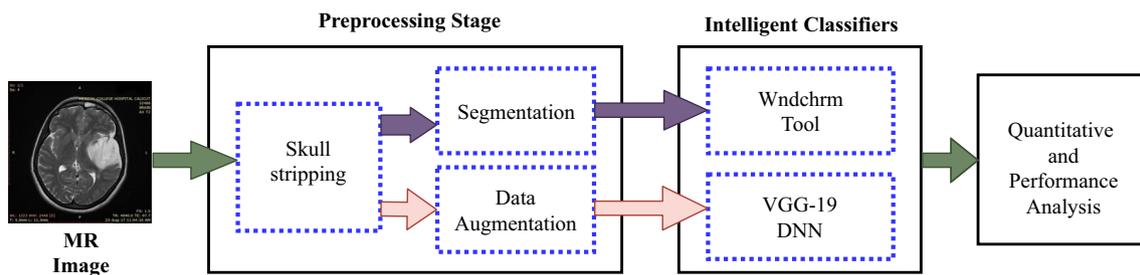

Fig. 2 Block Diagram of the Proposed System

Table 3 Wndchrm feature types

No:	Types of features
1.	Radon Transform Features
2.	Chebyshev Statistics
3.	Gabor Texture Features
4.	Multiscale Histogram
5.	First 4 Moments
6.	Tamura Texture Features
7.	Edge Statistics Features
8.	Object Statistics Features
9.	Zernike Features
10.	Haarlick Texture Features
11.	Chebyshev Fourier Features

where N_c is the size of training set for class c, x_{tj} is the jth feature of $\mathbf{x}_t$, x_{ij} is the jth feature of $\mathbf{x}_i \in c$, $|\mathbf{x}|$ is the number of features in $\mathbf{x}$, F_j is the Fisher score for jth feature and p is the exponential parameter set to -5 which is empirically determined. Initially, the tool is trained using training data set. Further the images under test are mapped into feature space. The similarity measure $m_{\mathbf{x}_t,c}$ between $\mathbf{x}_t$ and class c is computed as

$$m_{\mathbf{x}_t,c} = \frac{1}{d(\mathbf{x}_t, c) \sum_{i=1}^{M} \frac{1}{d(\mathbf{x}_t, C_i)}} \tag{6}$$

where $C_1, C_2, \cdots C_M$ represents M number of classes. $\left\{ m_{\mathbf{x}_t, C_k} \right\}_{k=1}^{M} = [k_1, k_2, \cdots k_M]$. k_1 is the similarity measure between $\mathbf{x}_t$ and first class C_1. Accordingly the similarity score for all classes are computed. Generally $\mathbf{x}_t$ is arranged to the class label as

$$k = \underset{i}{\arg\max} \; (k_i) \tag{7}$$

In our experimentation, number of classes is 4. Algorithm can be summarized as follows.

Algorithm 1 Wndchrm based intelligent system.

1: Prepare the training data set (N_c) for each class C_1, C_2, C_3, C_4.
2: Obtain the test data set (**T**).
3: Consider the finite test example $\mathbf{x}_t \in \mathbf{T}$.
   Compute $d(\mathbf{x}_t, c)$ as per Eq. 5.
   Compute similarity measure $m_{\mathbf{x}_t,c}$ as per Eq. 6.
   Determine the class label for $\mathbf{x}_t$ using Eq. 7.
4: Repeat the step 3 for all examples in **T**.
5: Compute performance metrics to evaluate the proposed system.

Methodology 2: VGG-19 DNN classifier

Another methodology proposed here is to classify tumor grades using VGG-19 DNN. A deep CNN otherwise called DNN architecture uses various layers namely, the input, hidden and output layers. The VGG-19 CNN architecture contains 12 convolution layers. The architecture uses higher receptive field stacked kernels and is very popular in biomedical application. The Fig. 3 illustrates its stacked structure suitable for glioma grade identification. The input layer of VGG-19 architecture accepts the image and defines its size using three parameters such as the height, width and channel size as $227 * 227 * 1$. In hidden layer the image is passed through various sub-layers such as convolution layer, normalization layer, rectification linear unit (ReLU) layer and pooling layer. In convolutionl layer, the weight matrix is calculated by fixing various parameters such as stride, filter size, padding, filter depth and image input size. The image after convolution gives features. The size of the output feature map is calculated using the Eq. 8 given below.

$$outputsize = \left[\frac{N + 2P - F}{S} \right] + 1 \tag{8}$$

where, N is the input image size, P is the padding used, S refers to the stride value and F is the filter size. The normalization layer performs regularization techniques which normalises the original image into its shifted and normalised version using a parameter called epsilon whose value is taken as approximately -10^{-8}. The ReLU layer defines an activation function which is mathematically represented using Eq. 9 below.

$$f(x) = \begin{cases} x & , x \geq 0 \\ 0 & , x < 0 \end{cases} \tag{9}$$

In order to reduce the spatial size of the image, pooling is performed using maximum pooling layer. In this layer initially the image is divided into various rectangular regions through down sampling and maximum value of each region is computed. Finally these features are fed to output layer for classification. In the output layer, the input image x is multiplied with the weight vector $\mathbf{w}$ and is added to a bias value b which gives the output class y as in Eq. 10 below.

$$y = w^T \mathbf{x} + b \tag{10}$$

This layer takes features obtained from previous layers and matched them to the output class. The output layer is followed by a softmax layer which converts the similarity into a normalised score. As a result, the probability of an image belonging to a particular class is predicted. The classification layer takes this probability and outputs the respective class to which the image belongs.

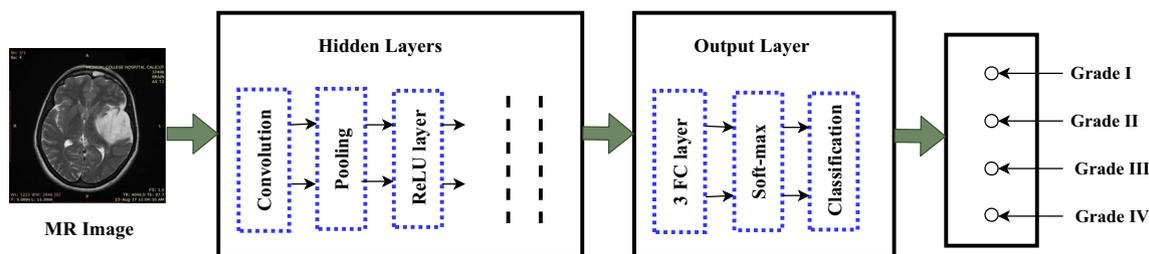   In this DNN method of classification, we used the transfer learning method for classifying the brain tumor

Fig. 3 VGG-19 CNN structure comprising of 12 convolution layers

glioma type dataset into four grades using the pre-defined VGG-19 DNN architecture. In transfer learning technique, pretrained network is fine tuned to classify new image collection. It is much faster and easier compared to training a new network as the learned features are applied to a new task that uses a fewer number of training images. The algorithm 2 gives the summary of the proposed deep learning based artificial intelligent system.

Algorithm 2 Proposed DNN based intelligent system.

1: Acquire and preprocess the MRI data set.
2: Split the dataset into 70:30 ratio as training and validation set.
3: Optimize augmentation parameters.
4: Determine the most feasible learning rate and fix the learning model.
5: Perform the transfer learning for proposed DNN model using the training data set.
6: Repeat the step 3 to 5 for other data set and analyze the performance of the model.

Results

The proposed paper categorize the glioma type brain tumor into four different grades namely low grade glioma (Grade I), oligodendroglioma (Grade II), anaplastic glioma (Grade III) and glioblastoma mutiforme (Grdae IV). Various pre-processing steps were carried out. Feature extraction and selection were performed on wndchrm based classifier. Results of various sections are given below.

Pre-processing results

In this step, initially skull-stripping was performed using Brain Suite tool. Figure 4 shows the skull-stripped images for the four grades. The segmented portion of the low grade glioma tumor using hybrid clustering method is shown in Fig. 5.

Wndchrm based classification

Initially, all tumor images belonging to each category are loaded and trained using the command ./wndchrm train [*options*] ⟨*datafile*⟩ ⟨*featurefile*⟩, where the *datafile* is the folder containing input images, *featurefile* is the features extracted as a text file and *options* is the different switching provisions that can be given. For example, the option: *-m* facilitates the generation of feature file for each class contained in the data set folder. Further, the testing is done using the command ./wndchrm test [*options*] ⟨*feature file*⟩ ⟨*report file*⟩, where the *report file* contains output details of the classifier. Test images are categorized effectively by computing the Fisher scores and arranging the weights to the image features. The option *-f* is used to get the most optimized and significant feature values. In our experimentation, five feature sets were analysed by changing the parameter values of *-f* from 0.5 to 0.25 and is given in Table 4.

The classification was performed using the command ./wndchrm classify ⟨*featurefile*⟩ ⟨*sampleimage*⟩, where *sampleimage* is the .tiff images to be classified. The classification accuracy yielded for the corresponding feature numbers is given in Table 5. It was observed that, the highest accuracy of 92.86% was obtained for 154 feature set. The most dominant feature type obtained were Haarlick texture features and first four moments as they measured using grey level matrix. The confusion matrix generated and various performance indices evaluated are given in Tables 6 and 7 respectively. The details about the performance metrics are given in Appendix.

DNN based classification

The DNN architecture used is VGG-19 for glioma tumor grade categorization. During training, the various parameters are to be defined like momentum constant, number of epoch, learning rate and minimum batch size. Momentum constant setting removes the oscillatory behaviour of the gradient descent. Epoch value is chosen in such a way that,

Fig. 4 Skull-stripped MR image
with glioma tumor [32] **a**
Grade-I **b** Grade-II **c** Grade-III **d**
Grade-IV

(a)

(b)

(c)

(d)
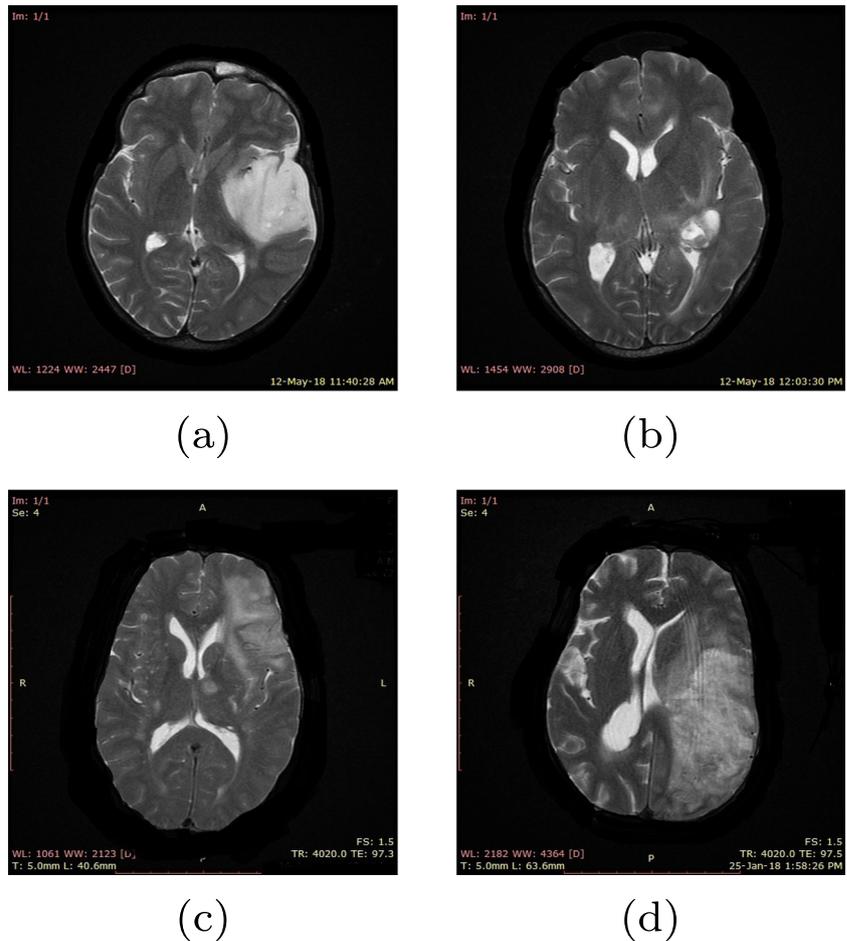
the training is continued till the error is lower than a threshold. Learning rate is chosen by considering the trade off between the convergence time and the classification accuracy. In our experimentation, those values are set as 0.09, 4, 0.0001, and 10 respectively. Prior to classification, data augmentation was performed and the augmenting parameters were optimised. This is done by changing the pixel value for various augmentation parameters such as resize, reflection and rotation. Table 8 shows the accuracy values obtained for various pixel range. It can be observed from the table that the highest accuracy values for VGG-19 network is obtained for the pixel range $(-30)°$ to $(30)°$.

Fig. 5 **a** Contrast enhanced
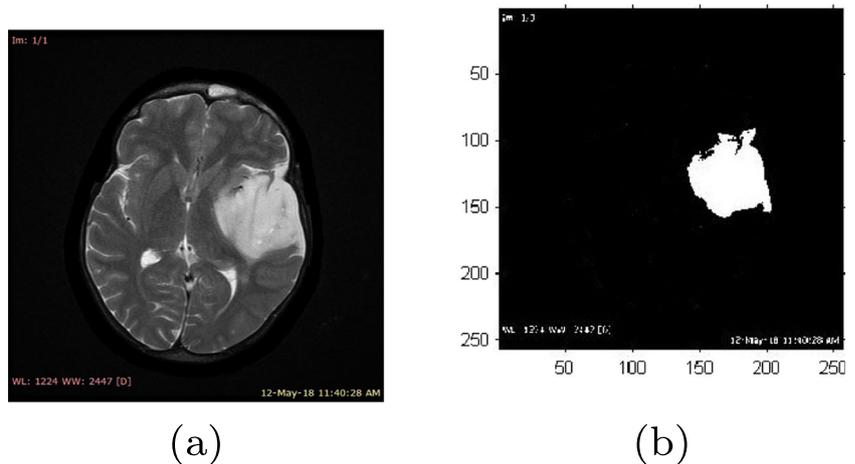Giloma grade I **b** Extracted
tumor portion

(a)

(b)

Table 4 Feature analysis for five experimentation cases

No:	Feature type	Set I	Set II	Set III	Set IV	Set V
1	Chebyshev Fourier features	2	7	4	16	6
2	Chebyshev statistics	6	5	18	0	43
3	First 4 moments	10	11	55	84	87
4	Edge statistics features	2	5	2	3	15
5	Object statistics features	3	8	15	25	50
6	Gabor texture features	1	2	0	0	8
7	Haarlick texture features	11	29	19	24	66
8	Multiscale histogram	10	14	18	23	39
9	Radon transform features	4	17	12	14	24
10	Tamura texure features	5	8	2	8	13
11	Zernike features	5	2	9	8	8
	Total	59	108	154	205	359

Table 5 Performance for different feature set

Set	No: of features	Accuracy
1	59	38.43
2	108	79.89
3	**154**	**92.86**
4	205	84.25
5	359	54.87

The raw corresponding to the highest accuracy values are highlighted

Table 6 Confusion matrix obtained for the Wndchrm classifier

Glioma grades	Grade I	Grade II	Grade III	Grade IV
Grade I	38	0	1	0
Grade II	1	47	1	2
Grade III	2	1	24	4
Grade IV	0	0	0	47

Table 7 Performance Indices of Wndchrm classifier for 557 images

Parameter	Value in %
Accuracy	92.86
Precision	92.89
Sensitivity	91.75
Specificity	97.48
F1-Score	92.32

Table 8 Accuracy values obtained using various pixel range for a set of augmenting parameters

Pixel range (in degrees)	Accuracy (in %)
0 to 30	56.99
0 to 60	68.47
0 to 90	55.86
0 to 180	70.59
−30 to 30	**82.54**
−60 to 60	74.21
−90 to 90	59.42
−180 to 80	72.54

The raw corresponding to the highest accuracy values are highlighted

Table 9 Accuracy details obtained for different learning rates

Learning rate	Accuracy (in %)
0.1	17.52
0.001	78.57
0.0001	**94.64**
0.00001	67.24
0.000001	48.29

The raw corresponding to the highest accuracy values are highlighted

Table 10 Confusion matrix obtained for the DNN classifier using 168 test images

Glioma grades	Grade I	Grade II	Grade III	Grade IV
Grade I	39	0	0	0
Grade II	1	48	1	1
Grade III	3	1	26	1
Grade IV	0	1	0	46

Table 11 Performance Indices of DNN classifier for 168 test images

Parameter	Value in %
Accuracy	94.64
Precision	94.71
Sensitivity	92.72
Specificity	98.13
F1-Score	93.71

Table 12 Confusion matrix obtained for the DNN classifier using 228 test images

Glioma grades	Grade I	Grade II	Grade III	Grade IV
Grade I	59	0	0	1
Grade II	0	60	1	0
Grade III	0	1	51	0
Grade IV	1	0	0	54
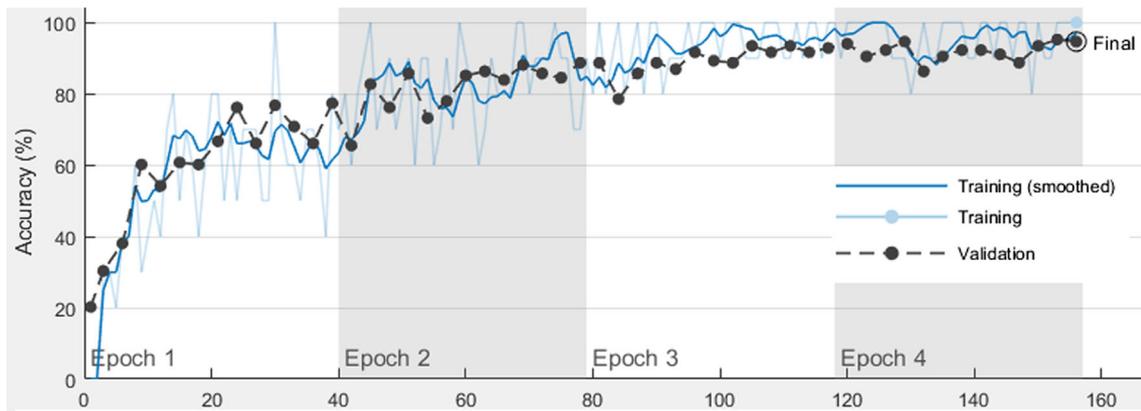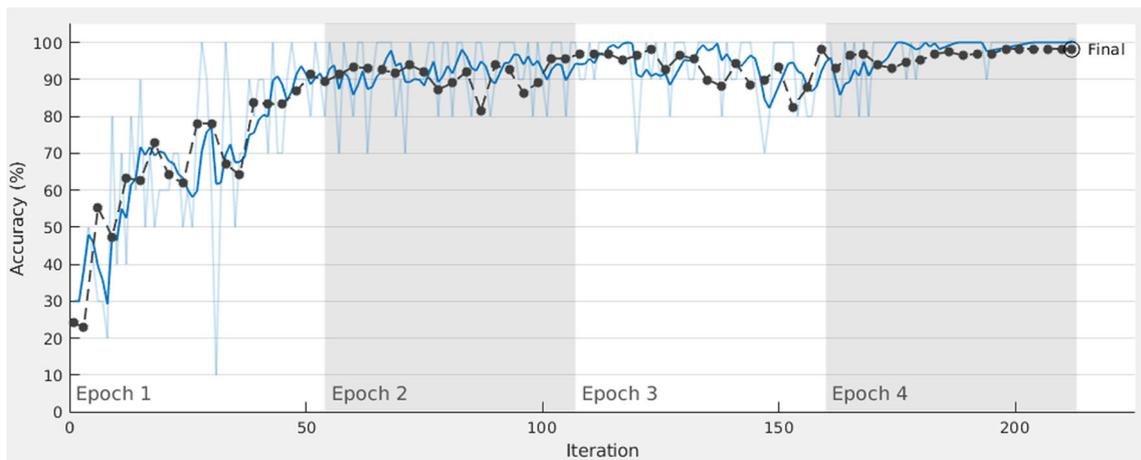
Fig. 6 Training graph obtained VGG-19 DNN system **a** for 557 images: 168 test data **b** for 761 images: 228 test data

This pixel range is given to the architecture and the learning rate is changed accordingly to obtain the maximum accuracy. Table 9 gives the different accuracy values obtained for the architectures for varying learning rates. It can be observed that, the maximum accuracy of 94.64% for VGG-19 architecture was obtained for learning rate value 0.0001. Table 10 shows the confusion matrix generated for the DNN classifier that used 557 input images corresponds to 168 test images.

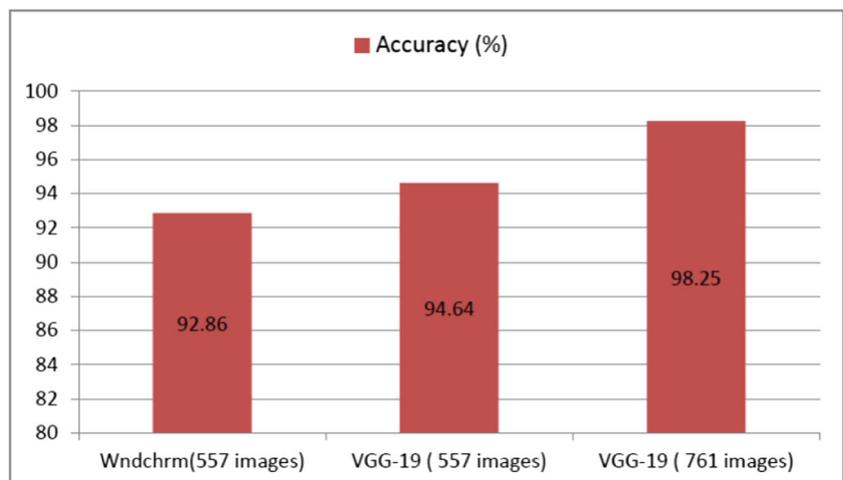**Fig. 7** Success rate comparison of the proposed intelligent classifiers

Table 13 Performance comparison with similar works

Author & Reference	Architecture/method	Output Class	Maximum Accuracy (%)
Khawaldeh S et al. [22]	AlexNet	3	91.16
Rathi, Palani [24]	LDA + DNN	2	91.66
Proposed work-1	Wndchrm	4	**92.86**
Proposed work-2	VGG-19 DNN	4	**98.25**

The raw containing results of the proposed works are highlighted

Various performance indices were measured such as accuracy, precision, sensitivity, specificity and F1-score. The Table 11 shows the performance values obtained for various metrics for the DNN architecture that used 557 input images (168 test images).

The experiment is repeated for 761 images (228 test images) keeping the same set of parameters, which gave an accuracy of 98.25%. The confusion matrix for the second case is given in Table 12. The Fig. 6 illustrates about the training graph (accuracy vs. epoch) obtained for VGG-19 architecture for the two cases of experimentation.

Discussion

In this paper, the glioma type brain tumor was categorized into four different grades using Wndchrm and DNN based intelligent systems. The DNN used VGG-19 architecture. The VGG-19 based intelligent classifier showcased better performance than the Wndchrm based system. From Fig. 7, it could be seen that accuracy of DNN classifier gave an increase of 3.61% for additional 204 input images. Figure 7 shows the graphical illustration of classification accuracy obtained for the two proposed methodologies.

Table 13 gives the performance comparison of similar architecture available in the literature using similar type of data set. In [22], they used a modified version of AlexNet for classifying MR brain images into three classes like healthy brain, low grade glioma and high grdae glioma. But, our CAD intelligent system is designed for four class problem, showed an increase in accuracy of 1.7% for Wndchrm classifier and 7.09% for VGG-19 DNN. Also, [24] classified MR brain images into tumorous or non-tumorous category. Here, the segmentation of the images have been done using probabilistic clustering methods and the feature selection is done with linear discriminant analysis (LDA). The deep neural network (DNN) classifier resulted a maximum accuracy of 91.66%. When compared to this methodology, our proposed work showed an increase in accuracy of 1.2% and 6.59% for Wndchrm and VGG-19 based intelligent systems respectively.

The computationally intensive training part of the intelligent systems can be done in an offline mode. The testing will be done in real time which is less computationally intensive. In this experiment, we considered images from a single source. Images can be collected from different hospitals to generalize the reliability and effectiveness of the proposed intelligent systems. Hence the proposed method can be used for real time application after sufficient trial and benchmarking. As a part of future work, we can do data augmentation cum classification based on Generative Adversarial Networks (GANs) for further improvement of the system performance. Thus, the system could be used for providing diagnostic suggestion for clinical application.

Conclusions

In this paper, automatic classification of glioma tumor grades using two artificial intelligent systems based on, Weighted Neighbour Distance using Compound Hierarchy of Algorithms Representing Morphology tool (Wndchrm) and VGG-19 deep learning networks is successfully performed and compared. The implemented medical decision support system used MR glioma tumor images of clinically proven cases from Government medical college hospital Kozhikode, India. The proposed system yielded good classification accuracy and categorized the images into four grades such as low grade glioma (Grade I), oligodendroglioma (Grade II), anaplastic glioma (grade III) and glioblastoma mutiforme (Grade IV).

According to the performance measures evaluated, both the proposed systems gave better specificity measure with a value higher than 97%. It was observed that VGG-19 DNN based classifier yielded an increase in accuracy of 2.28% compared to the Wndchrm tool based classifier. It is also observed that when the number of input images get increased, the accuracy is also increased where the DNN classifier gave a maximum accuracy of 98.25%. The better performance of the DNN intelligent system substantiates its significant role in clinical applications. The performance of the proposed methodologies is also compared with similar recent works. The experimental results obtained here validate the use of new biomedical feature extraction and optimization tool and the potential application of the deep learning networks for real data classification.

Compliance with Ethical Standards

Conflict of Interest The authors have declared no conflict of interest.

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors. The data only is collected from the government medical college, Kozhikode, India and is in accordance with the ethical standards of the institution.

Appendix: Performance metrics

In order to measure the optimality of the proposed systems, various performance indices are used. The various measures calculated are accuracy, sensitivity and specificity, precision and F1-score. All the parameters are derived from confusion matrix. The parameters are evaluated using the following values:

True Positive (TP):   abnormal brain accurately recognized as abnormal.

True Negative (TN):   normal brain accurately recognized as normal.

False Positive (FP):   normal brain inaccurately recognized as abnormal.

False Negative (FN):   abnormal brain inaccurately recognized as normal.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall(Sensitivity) = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$F1 - Score = \frac{2 * Precision * Recall}{Precision + Recall}$$

References

1. Louis, D. N., Perry, A., Reifenberger, G., Von Deimling, A., Figarella-Branger, D., Cavenee, W. K., Ohgaki, H., Wiestler, O. D., Kleihues, P., and Ellison, D. W., World Health Organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol.* 131:803–820, 2016.
2. El-Dahshan, E. A., Hosny, T., and Salem, A. B. M., Hybrid intelligent technique for MRI brain images classification. *Digit. Signal Process.* 20:433–441, 2010.
3. Javed, U., Riaz, M. M., Ghafoor, A. et al., MRI brain classification using texture features, fuzzy weighting and support vector machine. *Prog. Electromagn. Res. B* 53:73–88, 2013.
4. Nazir, M., and Wahid, F., A simple and intelligent approach for brain MRI classification. *J. Intell. Fuzzy Syst.* 28:1127–1135, 2015.
5. Subashini, M., Sahoo, S. K., and Sunil, V., A non-invasive methodology for the grade identification of astrocytoma using image processing and artificial intelligence techniques. *Expert Syst. Appl.* 43:186–196, 2016.
6. Mohan, G., and Subashini, M. M., MRI based medical image analysis: Survey on brain tumor grade classification. *Biomed. Signal Process. Control* 39:139–161, 2018.
7. Maximov, I. I., Tonoyan, A. S., and Pronin, I. N., Differentiation of glioma malignancy grade using diffusion MRI. *Phys. Med.* 40:24–32, 2017.
8. Hsieha, K. L.-C., Chena, C.-Y., and Lod, C.-M., Quantitative glioma grading using transformed gray-scale invariant textures of MRI. *Comput. Biol. Med.* 83:102–108, 2017.
9. Shamir, L., Orlov, N., and Eckley, D. M., Wndchrm—an open source utility for biological image analysis. *Source Code Biol. Med.* 3:1–13, 2008.
10. Orlov, N., Shamir, L., Macura, T., Johnston, J., Eckley, D. M., and Goldberg, I. G., WNDCHARM: multi-purpose image classification using compound image transforms. *Pattern Recognit. Lett.* 29:1684–1693, 2008.
11. Shajee Mohan, B. S., and Shanthini, K. S., Performance analysis of classifiers with feature selection and optimization in CBIR system for biological images. In: *Proceedings of the 2nd International Conference on Intelligent Systems and Image Processing*, pp. 216–222, 2014.
12. Madhukumar, S., and Santhiyakumari, N., Evaluation of k-means and fuzzy-c-means segmentation on MR images of brain. *Egypt. J. Radiol. Nucl. Med.* 46:475–479, 2015.
13. Ahammed Muneer, K. V., and Paul Joseph, K., Performance analysis of combined k-means and fuzzy-c-means segmentation of MR brain images. *Lect. Notes Comput. Vis. Biomech.* 28:830–836, 2018.
14. Mohsen, H., and El-Dahshan, E.-S. A., Classification using deep learning neural networks for brain tumors. *Future Comput. Inf. J.* 5:1–4, 2017.
15. Sharmaa, H., and Zerbeb, N., Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology. *Comput. Med. Imaging Graph.* 61:2–13, 2017.
16. Gao, X. W., Hui, R., and Tian, Z., Classification of CT brain images based on deep learning networks. *Comput. Methods Programs Biomed.* 138:49–56, 2017.
17. Vivanti, R., Joskowicz, L., Lev-Cohain, N., Ephrat, A., and Sosna, J., Patient-specific and global convolutional neural network for robust automatic liver tumor delineation in follow-up CT studies. *Med. Biol. Eng. Comput.* 56:1699–1713, 2018.
18. Najafabadi, M. M., and Villanustre, F., Deep learning applications and challenges in big data analytics. *J. Big Data* 2:1–21, 2015.
19. Gibson, E., Li, W. et al., NiftyNet: a deep-learning platform for medical imaging. *Comput Methods Programs Biomed.* 158:113–122, 2018.
20. Talo, M., Baloglu, U. B., Yildirim, O., and Acharya, U. R., Application of deep transfer learning for automated brain abnormality classification using MR images. *Cogn. Syst. Res.* 54:176–188, 2019.
21. Gudigar, A., Raghavendra, U., San, T. R., Ciaccio, E. J., and Acharya, U. R., Application of multiresolution analysis for automated detection of brain abnormality using MR images: a comparative study. *Future Gener. Comput. Syst.* 90:359–367, 2019.

22. Khawaldeh, S., Pervaiz, U. et al., Noninvasive grading of glioma tumor using magnetic resonance imaging with convolutional neural networks. *J. Appl. Sci.* 8:1–17, 2017.

23. Chato, L., and Latifi, S., Machine learning and deep learning techniques to predict overall survival of brain tumor patients using MRI images. In: *IEEE 17th International Conference on Bioinformatics and Bioengineering*, pp. 9–14, 2017.

24. Gladis Pushpa Rathi, V. P., and Palani, S., Brain tumor detection and classification using deep learning classifier on MRI images. *J. App. Sci. Eng. Technol.* 10:77–187, 2015.

25. Ye, C. Z., Yang, J., Geng, D. Y., Zhon, Y., and Chen, N. Y., Fuzzy rules to predict degree of malignancy in brain glioma. *Med. Biol. Eng. Comput.* 40:145–152, 2002.

26. Pang, S., Du, A., Orgun, M. A., and Yu, Z., A novel fused convolutional neural network for biomedical image classification. Med. Biol. Eng. Comput. https://doi.org/10.1007/s11517-018-1819-y, 2018.

27. Raghavendra, U., Fujita, H., Bhandary, S. V., Gudigar, A., Tan, J. H., and Acharya, U. R., Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images. *Inf. Sci.* 441:41–49, 2018.

28. Tan, J. H., Bhandary, S. V., Sivaprasad, S., Hagiwara, Y., Bagchi, A., Raghavendra, U., Rao, A. K., Raju, B., Shetty, N. S., Gertych, A., Chua, K. C., and Acharya, U. R., Age-related Macular Degeneration detection using deep convolutional neural network. *Future Gener. Comput. Syst.* 87:127–135, 2018.

29. Raghavendra, U., Bhat, N. S., Gudigar, A., and Acharya, U. R., Automated system for the detection of thoracolumbar fractures using a CNN architecture. *Future Gener. Comput. Syst.* 85:184–189, 2018.

30. Tana, J. H., Acharyaa, U. R., Bhandaryd, S. V., Chuaa, K. C., and Sivaprasad, S., Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network. *J. Comput. Sci-neth* 20:70–79, 2017.

31. Tan, J. H., Fujita, H., Sivaprasad, S., Bhandary, S. V., Rao, A. K., Chua, K. C., and Acharya, U. R., Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network. *Inf. Sci.* 420:66–76, 2017.

32. Data available at Govt, Medical College Kozhikode, India. https://www.govtmedicalcollegekozhikode.ac.in, 2017.