Research Article

An Effective and Novel Approach for Brain Tumor Classification Using AlexNet CNN Feature Extractor and Multiple Eminent Machine Learning Classifiers in MRIs

Alok Sarkar,1 Md. Maniruzzaman,1 Mohammad Ashik Alahe,1 and Mohiuddin Ahmad2

1Electronics and Communication Engineering Discipline, Khulna University, Khulna, 9208, Bangladesh
2Department of Electrical and Electronic Engineering, Khulna University of Engineering & Technology, Khulna, 9203, Bangladesh

Correspondence should be addressed to Alok Sarkar; alokece99@gmail.com

Received 27 October 2022; Revised 28 January 2023; Accepted 27 February 2023; Published 8 March 2023

Academic Editor: Akhilesh Pathak

Copyright © 2023 Alok Sarkar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A brain tumor is an uncontrolled malignant cell growth in the brain, which is denoted as one of the deadliest types of cancer in people of all ages. Early detection of brain tumors is needed to get proper and accurate treatment. Recently, deep learning technology has attained much attraction to the physicians for the diagnosis and treatment of brain tumors. This research presents a novel and effective brain tumor classification approach from MRIs utilizing AlexNet CNN for separating the dataset into training and test data along with extracting the features. The extracted features are then fed to BayesNet, sequential minimal optimization (SMO), Naïve Bayes (NB), and random forest (RF) classifiers for classifying brain tumors as no-tumor, glioma, meningioma, and pituitary tumors. To evaluate our model’s performance, we have utilized a publicly available Kaggle dataset. This paper demonstrates ROC, PRC, and cost curves for realizing classification performance of the models; also, performance evaluating parameters, such as accuracy, sensitivity, specificity, false positive rate, false negative rate, precision, f-measure, kappa statistics, MCC, ROC area, and PRC area, have been calculated for four testing options: the test data itself, cross-validation fold (CVF) 4, CVF 10, and percentage split (PS) 34% of the test data. We have achieved 88.75%, 98.15%, 86.25% and 100% of accuracy using the AlexNet CNN+BayesNet, AlexNet CNN+SMO, AlexNet CNN+NB, and AlexNet CNN+RF models, respectively, for the test data itself. The results imply that our approach is outstanding and very effective.

1. Introduction

The existence of any kind of abnormalities in the brain may lead the human nervous system as well as health in great danger [1]. Today, brain tumor is denoted as the most severe ones, which results from the development of uncontrolled destructive cells in the human brain. According to the World Health Organization and the American Brain Tumor Association reports [2], it can be categorized as glioma, meningioma and pituitary tumors based on cell structure and locations.

It is necessary to completely identify the types of brain tumors for proper diagnosis and treatment. The radiologist utilizes various clinical imaging techniques to visualize the structure of the brain to diagnose tumors. The most popular imaging techniques are ultrasound imaging, X-ray imaging, positron emission tomography, computed tomography, and magnetic resonance imaging (MRI). Among these, the MRI provides precise information on a given medical image, gives definite information on the shape of the cerebrum structure, and the location of inconsistencies in brain tissues. However, it is often chaotic, time-consuming, and error-prone if medical specialists manually detect brain tumors by examining MRIs [3]. To diminish these problems, the computer-aided detection and diagnostic approach has been developed for the last few decades [4–11]. In recent years, various efforts have been taken to elevate a robust and precise technique to classify brain tumors automatically by
using machine learning in MRIs [12–14]. Nevertheless, because of high inter- and intrashape, texture, and variation of contrast, this still remains a challenging task. Conventionally machine learning techniques rely on artisanal features, which limit the robustness of the outcomes, while deep learning techniques automatically extract important functionality that provides significantly better performance.

But still, the results obtained using deep learning technologies are not sufficient, and they do not always show good classification accuracy, uses of only AlexNet, for example [15]. The research is challenging, especially for selecting the region of interest (ROI), extracting deep features, using clamorous images, etc.

To minimize the aforementioned problems, in this study, we have proposed a novel and effective technique for brain tumor classification automatically employing the AlexNet CNN feature extractor and a variety of prominent machine learning classifiers, such as BayesNet, sequential minimal optimization (SMO), Naïve Bayes (NB), and random forest (RF) in MRIs.

We know that AlexNet CNN is a good tool to determine ROI for extracting deep features and can easily deal with clamorous and raw images, which are very essential for the current research. Moreover, deploying machine learning classification algorithms with AlexNet CNN certainly enhance the performance of the proposed model.

Our contributions to this work are as follows:

(i) We have utilized AlexNet CNN for separating human brain MRI datasets into training and test data and also for extracting the features from those separated datasets

(ii) Those extracted features are then used for classifying the brain tumors as no-tumor, glioma, meningioma, and pituitary tumors by using eminent BayesNet, SMO, NB, and RF classifiers

This paper is organized as follows. Section 1 discusses briefly the problems, some of the existing solutions, research gaps, objectives, contributions, and organizations of the paper. Section 2 presents related works on brain tumor detection and classification. It also discusses the research gaps at the end of this section. Section 3 describes overall working methodology, proposed architecture, and performance measuring metrics. Section 4 describes the results and performance of our model along with a comparison with other contemporary findings. Finally, Section 5 contains the conclusion of the entire study, limitations, and future works.

2. Related Works

Many researchers have published their works to represent different methodologies or approaches for detection and classification purposes over the past few years [15–36]. Various methodologies have been tested using different clinical databases, including magnetic resonance imaging (MRI) of brain tumors. Selvaraj et al. [37] developed a binary classifier for classifying brain MRIs a utilizing first- and second-order statistics and a least square support vector machine (SVM) classifier. Sarkar et al. [38] proposed a computer-aided approach for detecting and classifying brain tumors from MRIs utilizing genetic algorithm as a feature extractor and SVM as classifier, where they obtained 98.3% classification accuracy. John [25] used techniques based on the gray level coocurrence matrix (GLCM) and discrete wavelet transform (DWT) to detect and classify brain tumors. Ullah et al. [39] applied DWT for feature extraction and a feed-forward artificial neural network for brain MRI classification. Kharrat et al. [40] classified brain tumors as normal and abnormal classes through genetic algorithms and SVM classifiers. Papageorgiou et al. [33] obtained 90.26% and 93.22% accuracy for low- and high-grade glioma tumors, respectively, through fuzzy cognitive maps. Diaz-Pernas et al. [26] proposed an architecture for automatic detection and segmentation of brain tumors and acquired 97.3% accuracy. Shree and Kumar [41] used a GLCM feature extractor and probabilistic neural network classifier for classifying MRI datasets and achieved 95% accuracy. Arunachalam and Savarimuthu [42] offered an architecture that covered enhancement of image, transformation of image, extraction of features, and finally, image classification. They employed shift-invariant shearlet transforms for enhancing MRIs and used numerous feature extractors, such as Gabor, GLCM, and DWT to extract the features. Hussain et al. [30] utilized the traditional CNN approach to classify brain tumor images. They utilized fuzzy c-means clustering for segmenting the tumors and achieved an accuracy of 97.87%.

Several deep learning approaches have been extensively utilized by researchers to detect and classify brain tumors in the last decades. Sajid et al. [43] proposed a fully automated hybrid approach for the detection and segmentation of brain tumors in MRIs using deep learning. They utilized the BRATS2013 dataset and obtained 86%, 86%, and 91% dice scores, sensitivity, and specificity, respectively. Saxena et al. [44] utilized ResNet-50, Inception V3, VGG-16, and transfer learning techniques to classify the images; the ResNet-50 outperformed with 95% accuracy. In [27], Çinar and Yildirim utilized numerous convolutional neural network (CNN) models, such as Inception V3, GoogLeNet, ResNet-50, AlexNet, and DenseNet-201, for classifying the MRI datasets; whether they acquired respectable accuracies in every case. Khwaldeh et al. [45] customized the AlexNet architecture and employed it to classify brain MRIs, and obtained 91% accuracy. Preethi and Aishwarya [46] presented an architecture that combined wavelet-based GLCM to develop matrices of features. They utilized a deep neural network classifier to detect and categorize brain MRIs and obtained 92% accuracy. In [28], Hemanth et al. proposed a modified deep convolutional neural network (DCNN) to classify brain tumors, and they obtained excellent model performance in terms of accuracy. Khan et al. [47] introduced a computerized multimodel categorization technique to detect and classify brain tumor images. They utilized two preprocessed CNN models, known as VGG16 and VGG19, to extract features. They deployed the BRATS2018 dataset to validate their model and obtained about 97.80% accuracy.
In [48], Siar and Teshnehlab utilized a central clustering algorithm to extract features, then fed those features directly to the CNN model and achieved 96% accuracy. Kuang et al. [49] introduced a 3D CNN model to classify the BRATS2018 brain tumor image dataset. Preprocessing was performed using the grey-level normalization technique and contrast adjustment. Afterward, those images were transferred to the 3D CNN model for classification. They achieved a dice score of 92%. In another study [29], Zhou et al. considered 3D images as 2D cross-sections to utilize those images as the input dataset. They employed different models, such as DenseNet-RNN, DenseNet-LSTM, and DenseNet-DenseNet, and obtained accuracies of 87%, 91% and 92%, respectively. In [31], Krishnammal and Raja utilized a presupervised AlexNet model to classify the MRI dataset. Features were extracted through curvelet transform and GLCM matrix, and 100% accuracy was obtained. Updated versions of CNN architectures, namely, AlexNet and VGG16, were introduced by Ali et al. [15] with 96% and 98% of accuracy, respectively. Deepak and Ameer [32] proposed a fully computerized technique for classifying brain tumor images into three categories: glioma, meningioma, and pituitary tumors. Their designed architecture was an updated version of the GoogLeNet model and obtained 98% accuracy. In another study, Sultan et al. [34] presented a DCNN model for accurately classifying MRI data into glioma, meningioma, and pituitary tumors, where they achieved 96% accuracy. Abiwinanda et al. [35] designed numerous CNN architectures for classifying brain tumors into three classes: glioma, meningioma, and pituitary tumors. One of the architectures showed the best result with 98.51% of accuracy. In [36], Seetha and Raja classified the brain tumor images using the CNN. They extracted the features using ImageNet and obtained 97.50% accuracy.

Every method has some advantages and disadvantages. Some research gaps are identified in the above-mentioned papers and summarized briefly as follows.

(i) Many researchers have used different techniques in particular to determine the region of interest (ROI) for extracting features

(ii) Many researchers have used statistical features and texture features, but these have not given so much good accuracy with respect to deep features

Table 1: Statistical description of data.

<table>
<thead>
<tr>
<th>Image type</th>
<th>Total images</th>
<th>Training images</th>
<th>Testing images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glioma</td>
<td>900</td>
<td>630</td>
<td>270</td>
</tr>
<tr>
<td>Meningioma</td>
<td>900</td>
<td>630</td>
<td>270</td>
</tr>
<tr>
<td>No-tumor</td>
<td>900</td>
<td>630</td>
<td>270</td>
</tr>
<tr>
<td>Pituitary</td>
<td>900</td>
<td>630</td>
<td>270</td>
</tr>
<tr>
<td>Total</td>
<td>3600</td>
<td>2520</td>
<td>1080</td>
</tr>
</tbody>
</table>

Figure 1: Block diagram of the proposed method for identification and classification of brain tumors.
Most of the studies have used good quality MRI images. Raw images or noisy images have been found well dealt by deep learning methods.

Manual classification of brain tumors from MRIs is time-consuming, nonreproducible, and impractical if the data size is larger. Thus, an excellent model is required for feature extraction and classification of brain tumors accurately from MRIs.

3. Materials and Methods

3.1. Research Implementation Block Diagram. Figure 1 shows the overall working architecture of our proposed model to accurately identify and classify brain tumors from MRIs. Each key component of this architecture, such as the dataset including training and test data, preprocessing, feature extraction, prediction model, classified results, and performance measurement of the diagram, are sequentially narrated in the upcoming subsections. We have used AlexNet CNN in the MATLAB platform [50] for categorizing the dataset as training and test as well as extracting the features from it. Then, the extracted features are employed in the WEKA platform [51] to classify the brain tumors of MRIs as no-tumor, glioma, meningioma, and pituitary tumors utilizing prominent BayesNet, SMO, NB, and RF classifiers and establishing a prediction model. The performance of the model is assessed by the classified results.

3.2. Data Description. The selection of an appropriate dataset is the first and foremost concern for identifying and classifying medical images. Therefore, we collected the dataset from a trustworthy website, the Kaggle database [52, 53], a well-known web-based data source for brain MRIs. From [52], we have utilized the glioma, meningioma, and pituitary tumor data, and from [53], we have utilized the no tumor data. In total, 3600 images have been used in this study, where each class contains the same number of images (900...
Table 2: Confusion matrix for (a) test data itself, (b) CVF 4, (c) CVF 10, and (d) PS 34% of test data using BayesNet, SMO, NB, and RF classifiers.

(a)  
<table>
<thead>
<tr>
<th></th>
<th>BayesNet classifier</th>
<th>SMO classifier</th>
<th></th>
<th>BayesNet classifier</th>
<th>SMO classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Glioma</td>
<td>Meningioma</td>
<td>No tumor</td>
<td>Pituitary</td>
<td>Glioma</td>
</tr>
<tr>
<td>Glioma</td>
<td>203</td>
<td>36</td>
<td>6</td>
<td>25</td>
<td>Glioma</td>
</tr>
<tr>
<td>Meningioma</td>
<td>30</td>
<td>175</td>
<td>27</td>
<td>38</td>
<td>Meningioma</td>
</tr>
<tr>
<td>No-tumor</td>
<td>5</td>
<td>22</td>
<td>243</td>
<td>0</td>
<td>No-tumor</td>
</tr>
<tr>
<td>Pituitary</td>
<td>14</td>
<td>37</td>
<td>3</td>
<td>216</td>
<td>Pituitary</td>
</tr>
</tbody>
</table>

(b)  
<table>
<thead>
<tr>
<th></th>
<th>BayesNet classifier</th>
<th>SMO classifier</th>
<th></th>
<th>BayesNet classifier</th>
<th>SMO classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Glioma</td>
<td>Meningioma</td>
<td>No tumor</td>
<td>Pituitary</td>
<td>Glioma</td>
</tr>
<tr>
<td>Glioma</td>
<td>197</td>
<td>32</td>
<td>6</td>
<td>35</td>
<td>Glioma</td>
</tr>
<tr>
<td>Meningioma</td>
<td>44</td>
<td>144</td>
<td>30</td>
<td>52</td>
<td>Meningioma</td>
</tr>
<tr>
<td>No-tumor</td>
<td>5</td>
<td>29</td>
<td>235</td>
<td>1</td>
<td>No-tumor</td>
</tr>
<tr>
<td>Pituitary</td>
<td>22</td>
<td>37</td>
<td>4</td>
<td>207</td>
<td>Pituitary</td>
</tr>
</tbody>
</table>

(c)  
<table>
<thead>
<tr>
<th></th>
<th>BayesNet classifier</th>
<th>SMO classifier</th>
<th></th>
<th>BayesNet classifier</th>
<th>SMO classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Glioma</td>
<td>Meningioma</td>
<td>No tumor</td>
<td>Pituitary</td>
<td>Glioma</td>
</tr>
<tr>
<td>Glioma</td>
<td>191</td>
<td>35</td>
<td>8</td>
<td>36</td>
<td>Glioma</td>
</tr>
<tr>
<td>Meningioma</td>
<td>43</td>
<td>138</td>
<td>35</td>
<td>54</td>
<td>Meningioma</td>
</tr>
<tr>
<td>No-tumor</td>
<td>5</td>
<td>29</td>
<td>235</td>
<td>1</td>
<td>No-tumor</td>
</tr>
<tr>
<td>Pituitary</td>
<td>20</td>
<td>37</td>
<td>5</td>
<td>208</td>
<td>Pituitary</td>
</tr>
</tbody>
</table>
Table 2: Continued.

<table>
<thead>
<tr>
<th></th>
<th>NB classifier</th>
<th>RF classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Glioma</td>
<td>Meningioma</td>
</tr>
<tr>
<td>Glioma</td>
<td>189</td>
<td>38</td>
</tr>
<tr>
<td>Meningioma</td>
<td>48</td>
<td>134</td>
</tr>
<tr>
<td>No-tumor</td>
<td>5</td>
<td>29</td>
</tr>
<tr>
<td>Pituitary</td>
<td>22</td>
<td>39</td>
</tr>
</tbody>
</table>

(d)

<table>
<thead>
<tr>
<th></th>
<th>BayesNet classifier</th>
<th>SMO classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Glioma</td>
<td>Meningioma</td>
</tr>
<tr>
<td>Glioma</td>
<td>116</td>
<td>37</td>
</tr>
<tr>
<td>Meningioma</td>
<td>25</td>
<td>98</td>
</tr>
<tr>
<td>No-tumor</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td>Pituitary</td>
<td>10</td>
<td>25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>NB classifier</th>
<th>RF classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Glioma</td>
<td>Meningioma</td>
</tr>
<tr>
<td>Glioma</td>
<td>120</td>
<td>36</td>
</tr>
<tr>
<td>Meningioma</td>
<td>31</td>
<td>84</td>
</tr>
<tr>
<td>No-tumor</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>Pituitary</td>
<td>9</td>
<td>21</td>
</tr>
</tbody>
</table>
The elaborate allocation of the data collected and deployed in this study is shown in Table 1. We split the dataset (3600 MRIs) into training and test datasets: 2520 (70% of the total dataset) images are utilized as training images, and the rest 1080 (30% of the total dataset) images as test images, using AlexNet CNN. Training data are deployed for model prediction; on the other hand, test data are applied for testing the model’s performance.

3.3. Preprocessing. Preprocessing is a technique that augments the quality of images and makes them able for further steps. In a certain cases, the separation between normal and abnormal tissue is complicated because of the high-clamor level. As a result, specialists may normally commit errors in diagnosis. Then again, minor contrasts between normal and abnormal tissues can likewise be concealed by the clamor. Therefore, it is important to evacuate the conceivable clamors through preprocessing. In addition, enhancement of the visual nature of images offers tremendous favor to the specialists. In this research, MRIs are preprocessed with an anisotropic filter. The size of the images in the dataset of MRIs is not the same. These different sizes of images illustrate the input layer of the CNN network; therefore, these images need to be preprocessed, standardized, and resized to $227 \times 227$ pixels.

3.4. Feature Extraction. One presupervised CNN model, AlexNet [15, 31], is utilized in this investigation to include extraction, because it is a fantastically incredible model fit for accomplishing high-level accuracy on testing datasets. AlexNet is the most important known deep CNN structure which comprises twenty-five layers, eight of which add to

![Figure 4: ROC curves of (a) glioma, (b) meningioma, (c) no tumor, and (d) pituitary tumors for test data itself using BayesNet, SMO, NB, and RF classifiers.](image-url)
learning by modifying loads. Five of these are known as convolution layers while staying three are known as fully connected (FC) layers. In the AlexNet design, max-pooling layers are connected to convolution layers in series with each other. The first convolution layer and then max-pooling layers utilize fluctuating kernel sizes [54]. The layer of max-pooling ensues the layer of convolution [54, 55]. FC6 is the first FC layer, and FC7 is the second FC layer in activation are utilized to remove the features of vectors. There are altogether 4096 features of FC6 and FC7 vectors in the AlexNet CNN design [56].

3.5. Prediction Model. Consequently, we have developed predictive models based on features extracted from the trained data. A prediction model is a procedure through which future results or conducts are predicted dependent on the former and present data. It is a measurable examination procedure that empowers the assessment and count of the likelihood of specific outcomes identified with programming or frameworks. The prediction model works by gathering information, making a factual model, and applying probabilistic methods to anticipate the possible results.

3.6. Proposed Network Architecture

3.6.1. Convolutional Neural Network. The convolutional neural network (CNN) is a sort of multilayered feedforward neural network and one of the significant ideas of the deep learning method. Its basic structure consists of several convolution layers, a rectified linear unit (ReLU) layer,
pooling layers, and FC layers, as shown in Figure 2. The basic function of the convolution layer is to extract the local features of the existing layers. The ReLU layer performs element-by-element activation. The pooling layers are introduced for downsampling. Max-pooling is commonly utilized to reduce the accuracy of the feature maps by incorporating linguistically similar features. To overcome the fitting issues, some neurons have been removed from the CNN architecture, known as dropouts. At the end, the FC layer yields the class score value ranging from 0 to 1 which is deployed to attain the classification decisions. Essentially, SoftMax layers are introduced by default in this motive.

3.6.2. AlexNet. Figure 3 shows the architecture of an AlexNet which is mainly comprised of five convolutional layers \((11 \times 11, 5 \times 5, 3 \times 3, 3 \times 3, 3 \times 3)\), three max-pooling layers \((3 \times 3)\), and three FC layers. The first two max-pooling layers are deployed after the first two convolutional layers sequentially. The third, fourth, and fifth convolutional layers are connected directly. The third max-pooling layer is inserted after the fifth convolutional layer, and its output is fed into a series of three FC layers. The success of AlexNet is usually credited to ReLU, stochastic gradient descent (SGD), dropout, etc. ReLU is introduced for accelerating the speed of training processes. The value of the convolutional kernel is extracted by optimizing the total cost function and applying the SGD algorithm. To eliminate the overfitting problem, a dropout layer is applied in the first two FC layers. The third FC layer, also known as the softmax layer, is used to classify various objects in a CNN [57].

Figure 6: ROC curves of (a) glioma, (b) meningioma, (c) no tumor, and (d) pituitary tumors for the CVF 10 of test data using BayesNet, SMO, NB and RF classifiers.
3.7. Classifiers

3.7.1. BayesNet Classifier. The BayesNet classifier is a kind of directed acyclic graph which enciphers a joint probability distribution over a set of random variables [58, 59]. It is a twosome of \( B = (G, \theta) \), where \( G \) represents the graph whose vertices correspond to the random variables \( X_1, \ldots, X_n \) as well as whose edges denote direct dependencies between the variables. \( \theta \) denotes the set of parameters which quantifies the network. For the BayesNet classification model, in the present study, a simple estimator is used, and the batch size is chosen as 100 to get better classification accuracy.

3.7.2. Sequential Minimal Optimization Classifier. Sequential minimal optimization (SMO) is an important algorithm that can solve the quadratic programming problem as well as be materialized by training the SVMs. At this issue, the least plausible optimization problem includes the Lagrange multipliers which have two numbers, because the Lagrange multipliers must comply with a direct equity limitation. First, it discovers two multipliers, and then it attempts to enhance them. It will be repeated until it meets the equilibrium state [60]. Penalty parameter (C) is chosen as 1, and a polynomial kernel is selected based on the trial-and-error method for the SMO classification model, because they provide better model performance.

3.7.3. Naïve Bayes Classifier. The Naïve Bayes (NB) classifier is a probabilistic machine learning model based on Bayes theorem; that is, very useful for classifying a large amount of data. It anticipates based on the probability of an object. Thus, the model is developed, and the common
input to the model is batch size (number of experiments carried out) [61], whose value is selected as 100 for better accuracy of the model.

3.7.4. Random Forest Classifier. The random forest (RF) classifier is a supervised machine learning algorithm that creates a forest with a bunch of trees. Its working procedure consists of building simple decision trees and finally making the decision based on the votes. It looks for the best feature in every step (splitting nodes and growing trees), which predominantly results in a superior model. Batch size, number of features, and seed value are the common inputs in the RF model [61]. For better performance, we have considered batch size as 100 and seed value as 1.

**Table 3**: Calculated values of ROC and PRC areas for test data itself, CVF 4, CVF 10, and PS 34% of test data using BayesNet, SMO, NB, and RF classifiers.

<table>
<thead>
<tr>
<th>Testing options</th>
<th>BayesNet</th>
<th>SMO</th>
<th>NB</th>
<th>RF</th>
<th>BayesNet</th>
<th>SMO</th>
<th>NB</th>
<th>RF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test data itself</td>
<td>0.933</td>
<td>0.984</td>
<td>0.900</td>
<td>1.000</td>
<td>0.798</td>
<td>0.944</td>
<td>0.716</td>
<td>1.000</td>
</tr>
<tr>
<td>CVF 4</td>
<td>0.899</td>
<td>0.925</td>
<td>0.890</td>
<td>0.956</td>
<td>0.723</td>
<td>0.795</td>
<td>0.697</td>
<td>0.894</td>
</tr>
<tr>
<td>CVF 10</td>
<td>0.901</td>
<td>0.924</td>
<td>0.891</td>
<td>0.956</td>
<td>0.727</td>
<td>0.791</td>
<td>0.697</td>
<td>0.897</td>
</tr>
<tr>
<td>PS 34%</td>
<td>0.895</td>
<td>0.889</td>
<td>0.892</td>
<td>0.932</td>
<td>0.718</td>
<td>0.713</td>
<td>0.703</td>
<td>0.848</td>
</tr>
</tbody>
</table>

**Figure 8**: Cost curves of (a) glioma, (b) meningioma, (c) no tumor, and (d) pituitary tumors for the test data itself using BayesNet, SMO, NB, and RF classifiers.
3.8. Classified Result. In this research, it is expected that the final classified results will be obtained as glioma, meningioma, no-tumor, and pituitary tumors from MRI datasets. Then, we wish that the qualitative and quantitative values of the performance measurement of the proposed model’s will be attained to evaluate the model performance.

3.9. Performance Measurement. A confusion matrix has been utilized to calculate the performance of a classifier. The performances are described by four components of the confusion matrix, such as true positive (TP), true negative (TN), false positive (FP), and false negative (FN), which are whole in number and not in rate. The values of the performance parameters, such as accuracy, sensitivity, specificity, kappa statistics, false positive rate, precision, f-measure, and Matthew’s correlation coefficient (MCC), have been calculated for four testing options: the test data itself, the cross-validation fold (CVF) 4 [16, 34], the CVF 10 [19], and percentage split (PS) 34% [51] of test data, by employing the Equations (1)–(11) to measure the performance of our proposed model.

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \quad (1)
\]

\[
\text{Sensitivity or Recall or True Positive Rate} = \frac{TP}{TP + FN}, \quad (2)
\]

\[
\text{Specificity or True Negative Rate} = \frac{TN}{TN + FP}, \quad (3)
\]

\[
\text{Kappa statistics} = \frac{\text{Observed Agreement} - \text{Expected Agreement}}{1 - \text{Expected Agreement}}. \quad (4)
\]
Here,

\[
\text{Observed Agreement} = \frac{\text{Total Number of Correctly Classified}}{n},
\]

(5)

and

\[
\text{Expected Agreement} = \frac{(c^1 \times r^1_m) + (c^2 \times r^2_m) + (c^3 \times r^3_m) + (c^4 \times r^4_m)}{n},
\]

(6)

where \(c^1_m\) is the column 1 marginal, \(c^2_m\) is the column 2 marginal, \(c^3_m\) is the column 3 marginal, \(c^4_m\) is the column 4 marginal, \(r^1_m\) is the row 1 marginal, \(r^2_m\) is the row 2 marginal, \(r^3_m\) is the row 3 marginal, \(r^4_m\) is the row 4 marginal, and \(n\) is the total number of observations.

\[
\text{False Positive Rate (FPR)} = \frac{FP}{FP + TN},
\]

(7)

\[
\text{False Negative Rate (FNR)} = \frac{FN}{FN + TP},
\]

(8)

\[
\text{Precision} = \frac{TP}{TP + FP},
\]

(9)

\[
F - \text{measure} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}},
\]

(10)

Figure 10: Cost curves of (a) glioma, (b) meningioma, (c) no tumor, and (d) pituitary tumors for CVF 10 of the test data using BayesNet, SMO, NB, and RF classifiers.
Matthew’s correlation coefficient (MCC)
\[
\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}
\]

Different types of curves, such as the receiver operating characteristics (ROC) curve, precision-recall curve (PRC), and cost curve, have been plotted to understand the performance of every classifier. The ROC curve is utilized to quantify how well the zone could be recognized from noise by plotting the true positive rate against the false positive rate [28, 32]. The PRC is a well-known model execution measurement to assess the binary classification model [62]. A cost curve can be defined as a visual curve that represents the performance of the classifier based on the misclassification cost [63].

4. Results and Discussion

4.1. Confusion Matrix. Table 2 represents the confusion matrix for the test data itself: CVF 4, CVF 10, and PS 34% of the test data using BayesNet, SMO, NB, and RF classifiers, respectively. For the test data itself, the table shows that our models with the BayesNet classifier can classify 203 gliomas, 175 meningiomas, 243 no-tumors, and 216 pituitary tumor images accurately. The SMO classifier can classify 252 gliomas, 256 meningiomas, 269 no-tumors, and 263 pituitary images accurately. The NB classifier can classify 197 glioma cases, 144 meningioma cases, 235 no-tumor cases, and 207
pituitary cases accurately. The RF classifier, our model, can classify all the classes accurately.

For CVF 4 of the test data, in the table, we can see that our model with the BayesNet classifier can classify 181 glioma images, 158 meningioma images, 233 no-tumors, and 205 pituitary images accurately; with the SMO classifier, it can classify 224 glioma images, 200 meningioma images, 251 no-tumors, and 243 pituitary images accurately; with the NB classifier, it can classify 191 glioma images, 138 meningioma images, 235 no-tumors, and 208 pituitary images accurately; and with the RF classifier can classify 207 glioma images, 187 meningioma images, 256 no-tumors, and 245 pituitary images accurately.

As shown in the table, for CVF 10 of the test data, while we utilizing BayesNet classifier, our model can classify 183 glioma images, 153 meningioma images, 232 no-tumors, and 208 pituitary images accurately; with the SMO classifier, we can classify 211 glioma images, 208 meningioma images, 254 no-tumors, and 244 pituitary images accurately; again, while utilizing NB classifier, our model can classify 191 glioma images, 138 meningioma images, 235 no-tumors, and 208 pituitary images accurately; while utilizing RF classifier, our model can classify 214 glioma images, 182 meningioma images, 256 no-tumors, and 242 pituitary images accurately.

In the case of PS 34% of the test data, the table demonstrates that while we have utilized BayesNet classifier, our model can classify 116 glioma images, 98 meningioma images, 152 no-tumors, and 140 pituitary images accurately; while we have utilized SMO classifier, our model can classify 132 glioma images, 116 meningioma images, 161 no-tumors, and 150 pituitary images accurately; again, while we have utilized NB classifier, our model can classify 120 glioma images, 84 meningioma images, 154 no-tumors, and 144 pituitary images accurately; while we have utilized RF classifier, our model can classify 134 glioma images, 107 meningioma images, 163 no-tumors, and 156 pituitary images accurately.

4.2. ROC and PRC Curves of Test Data. We have plotted ROC (Figures 4–7) and PRC (not shown in this paper) curves of glioma, meningioma, no-tumor, and pituitary tumors for the test data itself, CVF 4, CVF 10, and PS 34% of test data using BayesNet, SMO, NB, and RF classifiers, respectively. And then, we have calculated the values of the area under each ROC and PRC curves. The average values are summarized in Table 3. As shown in the table, the average values of the ROC area (i) for the test data itself are 0.933, 0.984, 0.900, and 1.00; (ii) for CVF 4 are 0.899, 0.925, 0.890, and 0.956; (iii) for CVF 10 are 0.901, 0.924, 0.891, and 0.956; and (iv) for PS 34% are 0.895, 0.889, 0.892, and 0.932 of the test data using BayesNet, SMO, NB, and RF classifiers, respectively. These values indicate that the RF classifier performs better than the other classifiers in every case. The table also similarly reveals that the RF classifier gives the highest PRC area values of 1.00, 0.894, 0.897, and 0.848 for the test data itself, CVF 4, CVF 10, and PS 34%, respectively.
4.3. Cost Curve Analysis of Test Data. We have also analyzed cost curves which yields model performance based on the minimum misclassification cost. The minimum misclassification cost can also be defined as the minimum area under the bottom of the envelope. Therefore, the minimum region below the lower end of the range lowers the cost of misclassification and thus provides better classification performance. Figures 8–11 represent the cost curves for different testing options, such as the test data itself, CVF 4, CVF 10, and PS 34% of the test data employing BayesNet, SMO, NB, and RF classifiers for glioma, meningioma, no-tumor, and pituitary classes, respectively. As seen in Figure 8, the RF classifier very clearly offers minimal misclassification cost; that is, better classification performance. However, the SMO classifier shows a slightly lower misclassification cost in Figures 9–11.

4.4. Performance Evaluation. Finally, we have evaluated the performance of our models based on their respective confusion matrices. We have calculated the values of accuracy, sensitivity, specificity, FPR, FNR, precision, f-measure, kappa and MCC and summarized them in Table 4, which represents the values for the test data itself: CVF 4, CVF 10, and PS 34%, respectively.

For the test data itself, Table 4 shows that the values of accuracy, sensitivity, specificity, FPR, FNR, precision, f-measure, kappa, and MCC are higher for the RF classifier compared to those values of the other classifiers, exhibiting that the RF classifier gives excellent performance with 100% accuracy.

For the case of CVF 4 of test data, Table 4 shows that the values of accuracy, sensitivity, specificity, precision, f-measure, kappa, and MCC are higher and the values of FPR and FNR are lower for the SMO classifier than those values for the other classifiers, indicating that the SMO classifier gives better performance with 92.50% accuracy. Similar results are seen in the case of CVF 10 of test data, where the SMO classifier gives 92.45% accuracy.

However, PS 34% of the test data in the table shows that the values of accuracy, sensitivity, specificity, FPR, FNR, kappa, and MCC are higher for the RF classifier than the other classifiers. The values of precision are the same for both SMO and RF classifiers, and the values of f-measure are very slightly lower for the RF classifier compared to the SMO classifier. From these statistical values, we can assume that the RF classifier reveals a better performance with 89.27% accuracy.

\[
\begin{array}{|c|c|c|}
\hline
\text{Ref.} & \text{Classifiers} & \text{Database} & \text{Accuracy (\%)} \\
\hline
[15] & AlexNet & BRATS2015 & 60.20 \\
& VGG16 & & 98.20 \\
\hline
[26] & CNN & Nanfang Hospital, General Hospital, China & 97.3 \\
\hline
[27] & CNN & Kaggle & 97.2 \\
\hline
[28] & DCNN & M/s. Devaki Scan Centre, India & 94.5 \\
& MDCNN & & 96.4 \\
\hline
[29] & DenseNet+RNN & Nanfang Hospital, General Hospital, Tianjin Medical University, China & 84.61 \\
& DenseNet+LSTM & & 92.13 \\
& DenseNet+DenseNet & & 86.68 \\
\hline
[30] & CNN & BRATS & 97.87 \\
\hline
[32] & GoogLeNET & FIGSHARE & 98 \\
\hline
[34] & CNN & Nanfang Hospital and General Hospital, Tianjin Medical University, China, REMBRANDT & 96.13 \\
& CNN & FIGSHARE & 98.7 \\
\hline
[35] & CNN & FIGSHARE & 98.51 \\
\hline
[36] & CNN & BRATS2015, radiopedia & 97.5 \\
\hline
[64] & CNN & FIGSHARE & 98.71 \\
\hline
\hline
\hline
[67] & GoogLeNET+AlexNet & TJU Hospital, China. & 99.51 \\
\hline
\text{[proposed]} & AlexNet CNN+BayesNet & Kaggle & 88.75 \\
& AlexNet CNN+SMO & & 98.15 \\
& AlexNet CNN+NB & & 86.25 \\
& AlexNet CNN+RF & & 100 \\
\hline
\end{array}
\]
4.5. Comparison with Other Works. In this study, we have obtained the highest accuracy of 88.75%, 98.15%, 86.25%, and 100% for AlexNet CNN with (i) BayesNet, (ii) NB, (iii) SMO, and (iv) RF classifiers using the test data itself, respectively. Table 5 represents the comparison of our proposed model with several contemporary findings in terms of accuracy. As shown in the table, our model is highly appreciable, effective, and novel.

5. Conclusion
In this study, a novel and effective technique for feature extraction and classification of brain MRIs is presented for classifying brain tumors into glioma, meningioma, nontumor, and pituitary classes. An AlexNet CNN with BayesNet/SMO/NB/RF models is approached. Performances of the model are evaluated using the test dataset. We have calculated the values of performance parameters from respective confusion matrix, namely, accuracy, sensitivity, specificity, FPP, FNR, precision, f-measure, kappa statistics, MCC, ROC area, and PR area for four testing options: the test data itself, CVF 4, CVF 10, and PS 34% of test data. We have also explained ROC, PRC, and cost curves to measure the classification performance. We have acquired 100%, 98.15%, 88.75%, and 86.25% of accuracy using the AlexNet CNN+RF, AlexNet CNN+SMO, AlexNet CNN+BayesNet, and AlexNet CNN+NB models, respectively. This study presents a paramount technique for the classification of brain tumors from MRIs.

However, the proposed model was evaluated using a moderately sized dataset, which is one of the limitations of the current work. So, it is essential to evaluate the model with big data size in the future to see how well it performs. Another limitation is that the proposed approach has not been tested employing the MRIs of real-patient in Bangladesh. In the future, we will try to deploy this work on a real-time medical diagnostic system by acquiring the MRIs from different hospitals and diagnostic centers in Bangladesh. We believe that our approach is very much applicable in the health sector, offering innovative solutions with high precision for medical imaging, especially for diagnosing brain tumors more accurately.

Data Availability
Experimental codes and the data of this work are available from the corresponding author upon request at any time.

Conflicts of Interest
The authors declare no conflicts of interest.

Acknowledgments
We wish to express our gratitude to Khulna University, Khulna, Bangladesh-9208.

References


