

Research Article

BrainNet: Optimal Deep Learning Feature Fusion for Brain Tumor Classification

Usman Zahid ¹, Imran Ashraf ¹, Muhammad Attique Khan ², Majed Alhaisoni ³,
Khawaja M. Yahya ⁴, Hany S. Hussein ^{5,6} and Hammam Alshazly ⁷

¹Department of Computer Engineering, HITEC University, Taxila 47080, Pakistan

²Department of Computer Science, HITEC University, Taxila, Pakistan

³Computer Sciences Department, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh 11671, Saudi Arabia

⁴Department of Electrical Engineering, Umm Al-Qura University, Makkah, Saudi Arabia

⁵Electrical Engineering Department, College of Engineering, King Khalid University, Abha 62529, Saudi Arabia

⁶Electrical Engineering Department, Faculty of Engineering, Aswan University, Aswan 81528, Egypt

⁷Faculty of Computers and Information, South Valley University, Qena 83523, Egypt

Correspondence should be addressed to Imran Ashraf; imran.ashraf@hitecuni.edu.pk

Received 16 May 2022; Accepted 5 July 2022; Published 4 August 2022

Academic Editor: Abdul Rehman Javed

Copyright © 2022 Usman Zahid 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.

Early detection of brain tumors can save precious human life. This work presents a fully automated design to classify brain tumors. The proposed scheme employs optimal deep learning features for the classification of FLAIR, T1, T2, and T1CE tumors. Initially, we normalized the dataset to pass them to the ResNet101 pretrained model to perform transfer learning for our dataset. This approach results in fine-tuning the ResNet101 model for brain tumor classification. The problem with this approach is the generation of redundant features. These redundant features degrade accuracy and cause computational overhead. To tackle this problem, we find optimal features by utilizing differential evaluation and particle swarm optimization algorithms. The obtained optimal feature vectors are then serially fused to get a single-fused feature vector. PCA is applied to this fused vector to get the final optimized feature vector. This optimized feature vector is fed as input to various classifiers to classify tumors. Performance is analyzed at various stages. Performance results show that the proposed technique achieved a speedup of 25.5x in prediction time on the medium neural network with an accuracy of 94.4%. These results show significant improvement over the state-of-the-art techniques in terms of computational overhead by maintaining approximately the same accuracy.

1. Introduction

With the arrival of deep learning (DL), a remarkable change has occurred in the fields related to medical imaging, i.e., magnetic resonance imaging (MRI), computer vision (CV), and many more [1]. A brain tumor is one of the leading causes of death in both males and females. The survival rate of people with a brain tumor is very low, but it can be significantly increased if tumors are detected at an early stage [2]. According to the WHO standard, around 700,000 humans are affected by brain tumors, and since 2019, around 86,000 patients are being diagnosed with this. There have

been 16,830 fatalities due to brain tumors since 2019, and the average survival rate is only 35% [3]. In the USA, during 2021, the estimated cases of brain tumors are 83,570 which include 24,530 malignant and 59,040 nonmalignant tumors. The numbers of deaths occurred during 2020 are 18,600. In 2022, the estimated case of brain tumors will be 700,000 in the United States [4].

Tumors are basically malignant cells that are made by the uncontrolled development of the cancerous cells in any part of the body, whereas if this occurs in the brain it results in a brain tumor [5]. There are a lot of medical imaging techniques (MITs) used for the detection of diseases, i.e.,

computed tomography (CT) [6], magnetic resonance imaging (MRI) [7], ultrasonography [8], and many more [9]. Among all these techniques, magnetic resonance imaging (MRI) is one of the best techniques for the detection of brain tumors [10]. This is because it gives detailed information about the size, type, and position of the cell and is also very sensitive to the local changes in the tissue density [11, 12]. The images gathered from magnetic resonance imaging (MRI) must be analyzed by the experienced neuroradiologist to check for abnormalities within the brain, which requires a lot of time with manual effort. This ends up having a drawback of high cost, due to the need of a highly skilled neuroradiologist, and also because of the time-consuming process [13], so automated procedures are proposed by the researchers.

Automatic detection of brain tumors based on CV has been proposed by most of the researchers [14–16]. These techniques sometimes start with the preprocessing step which is generally used to enhance the image to achieve higher accuracy [17]. However, this is not the obvious case as it depends on the situation whether you need to do the preprocessing or not. As many of the researchers skip this part [18], the images are then used for feature extraction. As we explained earlier in the introduction, DL has shown remarkable results in so many fields, i.e., medical and computer vision (CV). The main problem with the deep learning approach is that it required a large amount of data and very high computational power to train the machine. However, this problem has been solved by the arrival of transfer learning (TL) [19]. In the case of transfer learning, the layers of the pretrained model are generally modified in a way that they can be used for specific problems. Typically, it can be performed by the modification of input and output layers to tune them according to your problem. Most of the researchers used several pretrained deep convolutional neural network (DCNN) models for computer vision and medical imaging, i.e., ResNet [20], Inception-V3 [21], VGG [22], and GoogLeNet [23].

With all this work, there is still a need for a lot more work in the phase of detection and classification. To address these problems, in this article, we proposed an optimal deep learning feature fusion for the classification of brain tumors. Our work is carried out in many steps. However, the main focus of this article is the optimization of deep learning features and after that the fusion of them into one matrix. To summarize, the contribution of our work is as follows:

- (1) Preprocessing of the dataset (normalization and conversion)
- (2) Selection of the optimal deep features using two algorithms differential evaluation (DE) and particle swarm intelligence (PSO)
- (3) Fusion of optimal features by serial fusion to obtain the fused optimal feature vector, which is passed to classifiers for the classification

This article is organized in the following sequence: Section 2 discusses the previous relevant techniques. The proposed work, which includes database creation, selection

of an optimal solution, and fusion, is presented in Section 3. Section 4 presents detailed classification results, and finally, the conclusion of this paper is discussed in Section 5.

2. Related Work

Brain tumor classification is an important and hot research topic nowadays. Several techniques have been introduced such as deep learning based, best features selection based, and many more [24–26]. In the literature, in [14], the authors presented an automated system for MRI-based brain tumor images with the help of machine learning techniques. The dataset they used for this is BraTS2017. They carried out this whole process in four steps that are preprocessing, segmentation, feature extraction, and then classification. Initially, in the preprocessing phase, they manually removed the skull and reduced the noise with the help of the median filter. In the segmentation, they used Chan–Vese (CV), and then, the features were extracted with the help of a gray-level co-occurrence matrix (GLCM). They used two classifiers SVM and KNN that are used for the performance evaluation, which outperformed the existing methods. They achieved an accuracy of about 98.13% for the SVM and 92.30% for the KNN.

A computer-aided diagnosis (CAD) system has been proposed in [26]. To achieve promising experimental evaluation for two different types of datasets, a recognition scheme named multi-level attention network (MANet) that is both cross-channel and spatial attention was proposed. The evaluated datasets are BraTS2018 and Figshare. The experimental result shows that this CAD technique achieved an accuracy of 94.91% on BraTS2018 and 96.51% on Figshare. Another novel computer-aided diagnosis (CAD) technique is presented in [27]. This technique extracts the features with the help of discrete wavelet transform (DWT). Later, these extracted features are then passed to the CNN to classify the input MRI images. In the experimental process, they achieved an accuracy of 99.3%.

A deep learning approach is presented in [18] to classify brain tumor disease. The datasets they used for this were BraTS2018 and BraTS2019. To extract the features, they utilized transfer learning to fine-tune the Densenet201 model. Later, they applied the entropy-kurtosis-based high feature value (EKbHFV) and the modified genetic algorithm (MGA) to select the optimal features. Fusion is performed with the help of a nonredundant serial-based approach and then classified by the cubic SVM. They achieved an accuracy of more than 95%. The authors of [17] examine the performance of multiple deep learning models, i.e., VGG16, AlexNet, GoogleNet, and ResNet50, in terms of their ability to examine the brain tumor. For evaluation, they used the criteria of accuracy and processing time. The result shows that ResNet50 gave the highest accuracy of about 95.8%, and AlexNet has the fast processing of about 1.2 sec which then decreases to 8.3 msec using GPU.

Brain tumor classification by the combination of both machine learning and deep learning approaches is presented in [28]. They used three different brain tumor classes named glioma, meningioma, and pituitary for classification. To

extract the deep features, they utilized transfer learning to fine-tune the GoogLeNet model. The extracted features are then classified with the help of the support vector machine (SVM), K-nearest neighbor (KNN), and softmax.

A computer-aided approach is presented in [16] to classify the brain MRI images. The author considered two classes that are the normal class and tumor class. The proposed technique was named 2D convolution neural network. The evaluation results are compared based on the recall value, F1-score value, and precision value. Their proposed method gave an accuracy of 97%.

In [29], H. A. Khan et al. presented a convolutional neural network (CNN) model for the classification of the brain tumor along with augmentation and image processing. Initially, in the image processing phase, they used canny edge detection to crop the black portion from an image. After this, they performed the data augmentation to increase the number of images in the dataset. They performed the augmentation by making minor changes in images like rotation, brightness, and flipping. Then, they compared their proposed convolution neural network (CNN) model with the pretrained VGG-16, ResNet-50, and Inception-V3 models. The experimental results were evaluated on very small data which gave the accuracy of about 100%, 96%, 89%, and 75% for their proposed CNN, VGG-16, ResNet-50, and Inception-V3, respectively.

In conclusion, the strategies discussed above primarily aimed to strengthen the extracted features in order to improve the outcome of the presented techniques. They also demonstrated the significance of classifiers in improving classification accuracy. In the classification phase, the failure to extract the best features and the problem of overfitting are the fundamental shortcomings of these strategies. We focused on the fusion of features for brain tumor categorization in this paper. We concentrated on extracting useful deep learning characteristics. Furthermore, we focused on the problem related to overfitting and then feature selection, with the goal of reducing prediction time without sacrificing too much on the accuracy. We proposed an end-to-end automated technique to address these gaps.

3. Proposed Methodology

We present a fully automated technique for brain tumor classification in this paper. This study looks at four different types of brain tumors. The following are the steps involved in this implementation:

- (1) Preprocessing is applied to normalize the dataset and convert images from single to multichannel (requirement of the deep learning model).
- (2) The ResNet101 pretrained model is fine-tuned using the transfer learning technique in order to use it for disease classification.
- (3) The DE and PSO algorithms are used to calculate two optimized feature vectors.
- (4) Using the serial fusion technique, we combine the two feature vectors to get the fused feature vector.

- (5) We apply PCA to the fused vector in order to select the top high variance features.
- (6) Different classifiers are used to classify these features. Figure 1 shows the block diagram of the proposed approach.

3.1. Database Preparation. In this paper, BraTS2018 is used for evaluation purpose. This dataset is clinically acquired preoperative multimodal MRI scans of glioblastoma (GBM/HGG) [30]. This dataset is divided into four categories: (a) native (**T1**), (b) postcontrast T1-weighted (**T1CE**), (c) T2-weighted (**T2**), and (d) T2 fluid-attenuated inversion recovery (**FLAIR**). T1 is composed of 28,446 images, T1CE is composed of 28,969 images, T2 is composed of 28,759 images, and FLAIR is composed of 28,413 images. These images are all in grayscale format with a resolution of 240 × 240 pixels. Figure 2 shows a few examples of images. Table 1 provides a summary of the overall images.

3.2. Resnet101 Deep Model. Deep neural networks have made significant progress in the field of image classification in recent years. A deep model, by definition, is the combination of low-level, mid-level, and high-level features, as well as a classifier. We used ResNet101 [31] to extract deep features in this paper. The VGG19 pretrained network, which is one of the deepest convolutional neural networks (CNNs), inspired this architecture. A CNN model, as previously stated, is made up of many different layers that are connected to each other. These layers are used for a variety of tasks, such as natural language processing and medical image classification. The convolutional filter size in ResNet101 is 33, and the stride value is 2. Downsampling is performed in the convolutional layers based on the stride value. This network has 347 layers and 379 connections. The input to the network has a dimension of 224 × 224 × 3. The filter size is [7, 7], the number of channels is 3, and the number of filters is 64 in the first convolution layer. The filter size in the max pooling layers is 3 × 3, and the stride value is 2. The number of channels and filters in the second convolutional layer is 64. The number of filters in the final convolution layer is 2048, with 512 channels [32]. We get an output vector of dimension $N \times 2048$, where N denotes the number of features, by extracting features from the pool5 layer. Figure 3 depicts the ResNet101 architecture in its entirety.

3.3. Transfer Learning-Based Network Training. In deep learning, data reliance is a severe issue. In comparison to typical machine learning techniques, a large amount of data is needed to train a deep model. The fundamental reason for this enormous amount of training data is that it is necessary to learn hidden patterns. However, a large amount of data is not often accessible for training a deep learning model in a few study domains, particularly in medical imaging. The concept of transfer learning (TL) [33] is to train a model with less data. It is not necessary to train the target model from

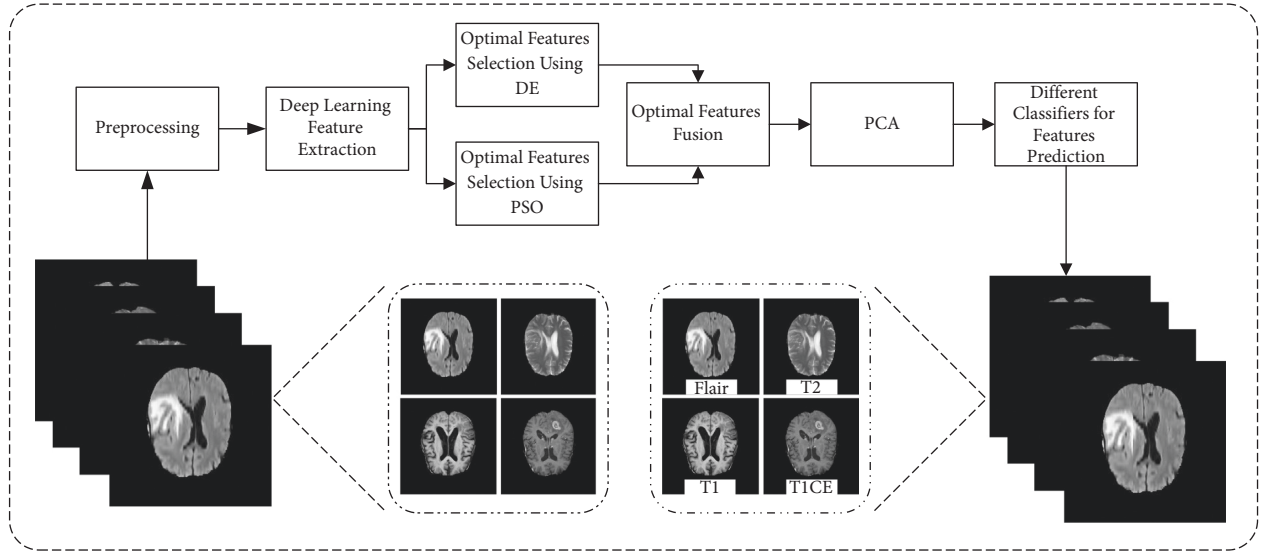


FIGURE 1: The proposed approach for the brain tumor classification.

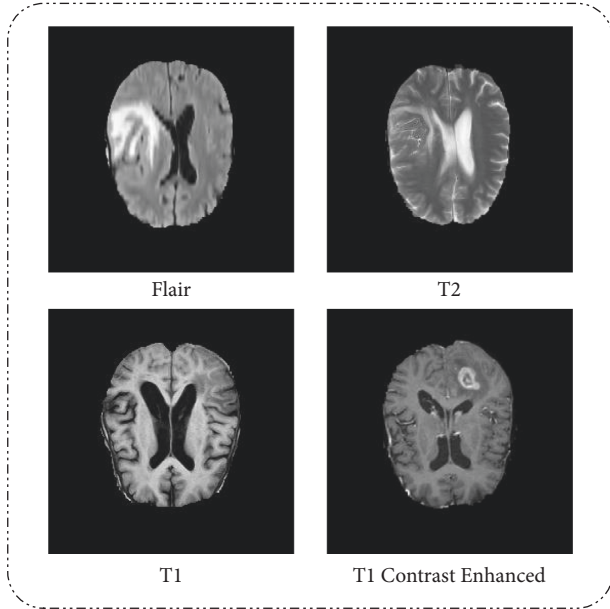


FIGURE 2: Sample database images.

TABLE 1: Summary of the dataset.

Class	Database name	Total images
T1	BraTS2018	28,446
T1CE		28,969
T2		28,759
FLAIR		28,413

scratch in TL. Deep transfer learning is defined mathematically as follows.

Given a transfer learning task with the following parameters: $\langle D_s, T_s, D_t, T_t, F_t(\cdot) \rangle$, D_s represents the source domain, T_s and T_t represent the learning task from the source and the target, D_t represents the target domain, and $F_t(\cdot)$ represents the nonlinear function that represents a deep

Layer Name	output Size	Architecture
conv 1	112x112	7x7, 64, stride 2
conv 2_x	56x56	3x1 maxpool, stride 2
		$\begin{bmatrix} 1 \times 164 \\ 3 \times 364 \\ 1 \times 1256 \end{bmatrix} \times 3$
conv 3_x	28x28	$\begin{bmatrix} 1 \times 1128 \\ 3 \times 3128 \\ 1 \times 1512 \end{bmatrix} \times 4$
conv 4_x	14x14	$\begin{bmatrix} 1 \times 1256 \\ 3 \times 3256 \\ 1 \times 11024 \end{bmatrix} \times 23$
conv 5_x	7x7	$\begin{bmatrix} 1 \times 1512 \\ 3 \times 3512 \\ 1 \times 12048 \end{bmatrix} \times 3$
	1x1	average pool, 1000-d fc, softmax
FLOPs		7.6×10^9

FIGURE 3: ResNet101 architecture.

neural network. Figure 4 depicts the model learning process using transfer learning graphically. In this figure, it is shown that the original ResNet101 model is trained on the ImageNet dataset [34], and then, knowledge is transferred using deep transfer learning for retraining this model on the target database. As a target database, the brain tumor database is used. We extract features from the pool5 layer and output a vector of dimension $N \times 2048$. We started with a learning rate of 0.0001 and a minimum batch size of 32 in the learning.

3.3.1. Training Process. Figure 5 depicts the deep learning model's complete training procedure. The following is a description of the figure.

- (1) We divide each class such as T1, T1CE, T2, and FLAIR into 50% training and testing images. A

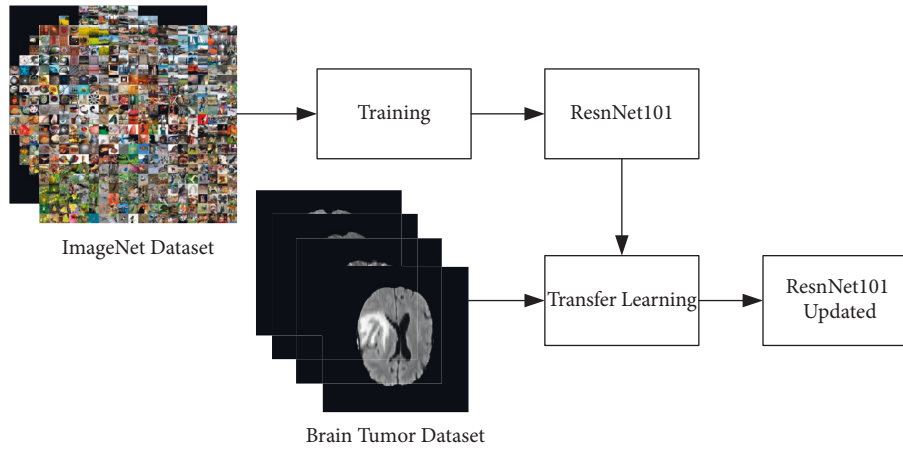


FIGURE 4: Deep transfer learning process.

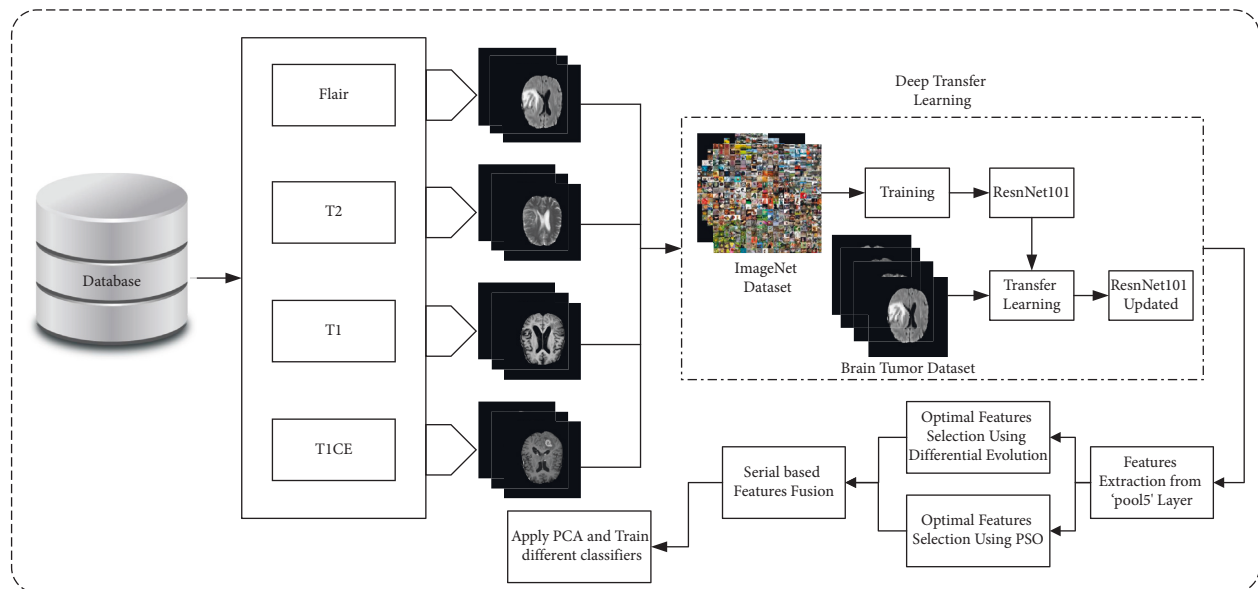


FIGURE 5: The proposed training process of the deep learning model for brain tumor classification.

- randomized procedure is used to separate these images.
- (2) The ResNet101 model was trained using deep transfer learning for brain tumor classification.
 - (3) Deep learning features are extracted from the pool5 layer.
 - (4) DE and PSO are two optimization strategies that are used to improve the extracted features.
 - (5) As the output, two optimal feature vectors are returned.
 - (6) We perform fusion of both optimal vectors using serial fusion.
 - (7) We apply PCA to the fused vector.
 - (8) We train the classifier and save the model for testing.

We train and save our model for brain tumor categorization based on the aforementioned methods. The following sections go over the optimization, fusion, PCA, and classification stages in detail.

3.4. Feature Optimization. The classification accuracy is improved by selecting the most optimal set of features from the initial set of features. These features were chosen for learning from original features with the least amount of loss. The main advantages are that they improve accuracy, take less time, and eliminate the issue of overfitting. In feature selection, the optimization process entails determining the optimum feasible values depending on the established objective function. Many evolutionary strategies for identifying the closest optimal solution are provided for this aim. We

implemented two algorithms in this article: differential evolution (DE) and particle swarm intelligence (PSO).

3.4.1. Differential Evolution. The DE is a global search optimization problem-solving evolutionary approach [35]. Because it uses fewer control factors than the genetic approach (GA), this technique is easier to use. It is significantly more effective in the realm of medical imaging because it has fewer control factors. It starts with a set of randomly generated starting values in the search space. The input data are then subjected to mutation and crossover, followed by a selection procedure to establish a new population. The steps involved in this project are listed as follows:

Input: original $N \times 2048$ dimensional deep feature vector.

Output: optimal feature vector of dimension $N \times 1119$.

(1) Step 1: We initialize the following parameters:

- (1) Population = 50
- (2) Minimum bound and maximum bound
- (3) Use the following expression to find these bounds:

$$\lambda_i^j = \lambda_{\min}^j + \text{rand}(0, 1) \cdot (\lambda_{\max}^j - \lambda_{\min}^j). \quad (1)$$

(2) Step 2: We calculate the fitness function, where fine KNN is opted as the fitness function, and the mean square error rate (MSER) is used for the performance evaluation.

(3) Step 3: We perform mutation as shown in the following equation:

$$M_r = 0.2. \quad (2)$$

(4) The following equation is used to define mutation:

$$v_i^j = \lambda_{r_1}^j + F(\lambda_{r_2}^j - \lambda_{r_3}^j). \quad (3)$$

(5) Step 4: We perform crossover as shown in the following equation:

$$C_r = 0.7,$$

$$C_i^j = \begin{cases} v_i^j, & \text{if } \text{rand}(0, 1) \leq C_r \text{ or } j = j_{\text{rand}} \lambda_i^j, \\ \text{Otherwise.} \end{cases} \quad (4)$$

(6) Step 5: We find fitness evaluation and selection. We repeat steps 2, 3, and 4 until the required optimal feature vector is obtained. An optimal feature vector of dimension $N \times 1119$ is obtained as a result.

3.4.2. Particle Swarm Optimization. Particle Swarm Optimization (PSO) is inspired by swarm behavior such as bird flocks and schooling fish. PSO is basically a population-based metaheuristic technique [36]. It is an efficient evolutionary algorithm, that is why it is extensively used to solve single or multiple-objective problems [37]. Furthermore, PSO is also a powerful computing tool in terms of speed and memory usage [38].

Particle Swarm Optimization (PSO) works on the basis of 5 steps that are mentioned as follows:

Input: original $N \times 2048$ dimensional deep feature vector.

Output: optimal feature vector of dimension $N \times 1125$.

(1) Step 1: We perform generation of population as shown in the following equation:

$$\text{Population} = p_{\text{gen}}(\text{gen} = 0, 1, \dots, \text{Max}_{\text{gen}})$$

$$p = \begin{bmatrix} b_{11}^L & b_{12}^L & b_{13}^L & \dots & b_{1M}^L \\ b_{21}^L & b_{22}^L & b_{23}^L & \dots & b_{2M}^L \\ \dots & \dots & \dots & \dots & \dots \\ b_{N1}^L & b_{N2}^L & b_{N3}^L & \dots & b_{NM}^L \end{bmatrix}, \quad (5)$$

where $b_{11}^L, b_{12}^L, b_{13}^L, \dots, b_{1M}^L$ denotes the particle/candidate solution. Single individual, i.e., b_{11}^L called as an agent. "P" basically denotes the population.

(2) Step 2: We calculate the fitness function, where fine KNN is opted as the fitness function, and the mean square error rate (MSER) is used for the performance evaluation.

(3) Step 3 (a): We find the local best. We find the local best from the first candidate solution

$$C = b_{11}^L b_{12}^L b_{13}^L \dots b_{1M}^L, \quad (6)$$

where C is the first candidate solution,

(1) $lb = C_1$

(2) $\forall i = 2$ to L

(3) if $\text{fitness}(C_i) > \text{fitness}(lb)$

(4) then

(5) $lb^i = C_i$

(6) else

(7) continue

(8) end

(4) Step 3 (b): We find the global best with the help of the following steps:

(1) To find the global best

(2) $gb = lb_1$

(3) $\forall k = 1$ to N

(4) if $\text{fitness}(lb_k) > \text{fitness}(gb)$

(5) then

(6) $gb = lb_k$

(7) else

(8) continue

(9) end

(5) Step 4: We update the speed and position. The equation to update the speed is

$$V_{n+1} = \omega V_3 + C_1 \text{rand}(lb - b_i) + C_2 \text{rand}(gb - b_i), \quad (7)$$

where V_3 is the old speed, C_1 and C_2 are the controlling parameters, ω is the inertia, $(lb - b_i)$ is the local updation, $(gb - b_i)$ is the global updation,

$C_1 \text{rand}(lb - b_i)$ is the local intelligence, and $C_2 \text{rand}(gb - b_i)$ is the global intelligence. All these parameters are combined to generate the updated speed which is V_{n+1} .

(6) The equation to update the position is

$$X_{n+1} = X_n + V_{n+1}, \quad (8)$$

where X_n is the old position and V_{n+1} is the updated speed. By the combination of these parameters, we achieved the updated position.

(7) Step 5: We find fitness evaluation and selection and repeat steps 2, 3, and 4 until the required optimal feature vector is obtained. In the output, an optimal feature vector of dimension $N \times 1125$ is obtained.

3.5. Feature Fusion. Feature fusion is a process in which two feature vectors are combined to get one feature vector, which is more appealing and discriminating than the two input feature vectors.

One of the biggest advantages of feature fusion is that it improves the image information in terms of features. In this paper, we implement the serial-based extended (SBE) approach for feature fusion.

Following are the two optimal feature vectors denoted by $FV_{(DE)}^{k_1}$ and $FV_{(PSO)}^{k_2}$ with the length of $N \times 1119$ and $N \times 1125$, respectively. Let $FV_{(fus)}^k$ be the fused feature vector with the dimension of $N \times K$, where k is basically the length of the feature vector after the fusion. Let 'Y' be the number of features of DE and 'Z' be the number of features of the PSO, then serial fused features 'K' have $(Y+Z)$ dimensions.

The fusion process includes the following steps:

- (1) Input both optimal feature vectors
- (2) The size of these vectors is $N \times 1119_{DE}$ and $N \times 1125_{PSO}$
- (3) A final resulting fused feature vector of dimension $(DE + PSO) N \times 2244$ is obtained in the output

At last, the fused vector is fed into the different classifiers, which produces two outputs: labeled prediction results and numerical results. Figure 6 depicts the labeled results, whereas Section 4 has the numerical results.

3.6. Principal Component Analysis. PCA (principal component analysis) is a dimensionality reduction technique [39]. This technique is normally carried out to reduce the dimensionality of huge data/feature vectors. This is performed because smaller data are easier to understand and analyze, and machine learning and deep learning algorithms can interpret them much more efficiently and rapidly [40]. PCA keeps only those features that carry a massive amount of information. This is accomplished by preserving just those components that have high variance [41].

Following is the optimal feature vector obtained by the optimal feature fusion denoted by $FV_{(FUS)}$ with the dimension of $N \times 2244$. We passed this feature vector to the

PCA. The high variance features which we have selected from this are of dimension $N \times 1000$. The detailed discussion about the accuracy achieved and the prediction time speedup is discussed below in the result section.

4. Results and Discussion

This section covers the detailed discussion about the numerical results we obtained for this work.

The dataset that we used for our experiments is 'BraTS2018'. This dataset is clinically acquired preoperative multimodal MRI scans of glioblastoma (GBM/HGG) [30]. Originally, this dataset consists of 130,200 images with 4 classes that are (a) native (T1), (b) postcontrast T1-weighted (T1CE), (c) T2-weighted (T2), and (d) T2 fluid-attenuated inversion recovery (FLAIR). Most of the images in the given dataset are totally black without any information in them.

In the preprocessing phase, we cleaned the dataset by removing the blank images. After the preprocessing, we ended up having 114,587 images, which we used for the training purpose. We computed the results in multiple steps: (a) with the help of a pretrained deep learning model, we extracted the deep features, and then, we obtained the accuracy by passing them to the different classifiers; (b) to select the optimal feature selection, we first used the differential evolution (DE); (c) then, we used the particle swarm optimization (PSO) and evaluated the results; (d) we fused both of these optimal feature vectors; (e) we compared the results. In the training testing phase, 10-fold cross-validation is used.

Multiple classifiers are utilized in order to compare the accuracies. These classifiers are the fine tree, linear discriminant, cubic SVM, boosted tree, bagged tree, subspace discriminant, narrow neural network, medium neural network, and wide neural network. Various performance metrics are used to report the results, such as the accuracy (%), prediction time (sec), sensitivity, precision, FPR, FNR, and area under curve.

The hyper parameters that we used for our work were as follows:

- (1) Epochs = 100
- (2) Learning rate = 0.05
- (3) Optimizer = stochastic gradient descent
- (4) Loss function = cross entropy
- (5) Momentum = 0.7

In order to conduct our experiments, we used Intel Core-i7 6th generation with 16 GB RAM and Nvidia GeForce GTX1070 GPU with 8 GB RAM. We used Matlab 2021a for our simulations.

This section contains the numerical results that we obtained from our experiments. Prediction accuracy of brain tumor disease with the help of original ResNet deep features is presented in Table 2. Among all the classifiers, the cubic SVM gives the highest accuracy of about 96.7%. The negative rate of cubic SVM is 3.3% with a prediction time of 4461.8 seconds. The wide neural network gives the second highest accuracy of about 96.1%. The negative rate of the

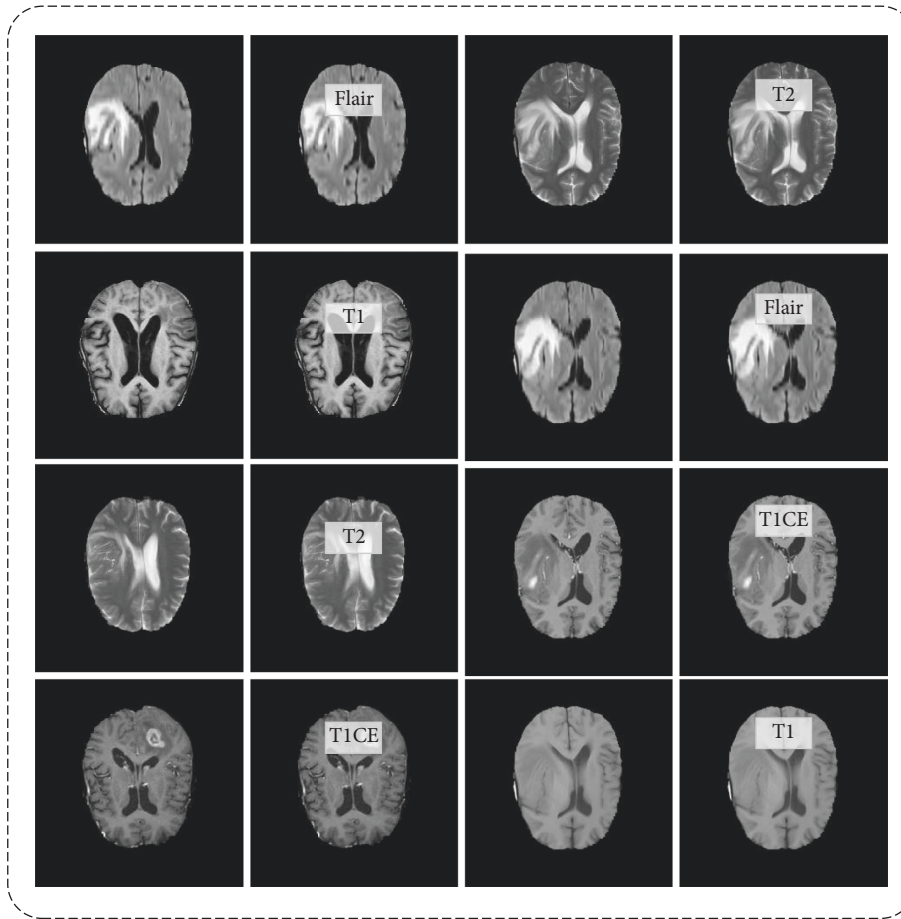


FIGURE 6: Prediction results in the form of labeled Images (a) numerical results.

TABLE 2: Comparison of prediction accuracy of brain tumors.

Classifiers	Org. ResNet101	Accuracies %		Feature fusion	%Age difference (org VS FF)
		Optimized			
		PSO	DE		
Fine tree	89.6	88.8	88	92.6	3.3%
Linear discriminant	96	95.7	95.6	95.7	-0.3%
Cubic SVM	96.7	96.7	96.6	96.7	0%
Boosted trees	92.8	92.1	92.3	92.5	-0.3%
Bagged trees	95.5	95.4	95.3	94.5	-1%
Subspace discriminant	95.8	95.5	95.4	95.4	-0.4%
Narrow neural network	95.7	95.5	95.4	93.9	-1.9%
Medium neural network	96	95.8	95.8	94.4	-1.7%
Wide neural network	96.1	95.9	96.1	95.4	-0.7%

WNN is 3.9% with a prediction time of 1310.4 seconds. The medium neural network gives the third highest accuracy of about 96%. The negative rate of the MNN is 4% with a prediction time of 4906.5 seconds. The prediction time of brain tumor disease with the help of original ResNet deep features is presented in Table 3. Based on the prediction time, the fine tree is executed in the minimum time of 286.34 seconds. The linear discriminant is the second and the bagged tree is the third best classifier with a minimum prediction time of 336.99 and 774.26 seconds, respectively. The accuracies of these classifiers are 7.1%, 0.7%, and 1.2%

with a less difference of 4175.46, 4124.81, and 3687.54 seconds, respectively, in the prediction time as compared to the cubic SVM. The confusion matrix of the cubic SVM on original features is also shown in Figure 7.

Results gathered after the optimal feature selection by PSO are shown in Table 2. Of all the classifiers, the cubic SVM gives the highest accuracy of about 96.7%. The negative rate of the cubic SVM is 3.3% with a prediction time of 954.87 seconds. The negative rate of the cubic SVM by PSO is the same as that of the original one with a decrease of 3506.93 seconds in the prediction time. The wide

TABLE 3: Comparison of the prediction time (sec) of brain tumors.

Classifiers	Prediction time (sec)				
	Original ResNet101	Optimized		Feature fusion	Speedup (org VS FF)
		PSO	DE		
Fine tree	286.34	139.28	159.25	156.73	1.8x
Linear discriminant	336.99	62.48	117.11	52.093	6.5x
Cubic SVM	4461.8	954.87	2088.8	3901.4	1.1x
Boosted trees	5615.4	2691.4	3013.2	2688.8	2.1x
Bagged trees	774.26	326.44	460.75	403.45	1.9x
Subspace discriminant	3063.2	688.63	1287.2	552.78	5.5x
Narrow neural network	5999.1	1332.7	2754.2	1662.1	3.6x
Medium neural network	4906.5	1457.7	3636.3	192.56	25.5x
Wide neural network	1310.4	404.63	701.48	100.38	13.1x

Brain Tumor Classes	Brain Tumor Classes			
	T1CE	Flair	T1	T2
T1CE	13975	155	214	140
Flair	159	13737	192	118
T1	225	125	13733	140
T2	205	81	116	13977

FIGURE 7: The confusion matrix of the cubic SVM after the original feature classification.

neural network gives the second highest accuracy of about 95.9%. The negative rate of this classifier is 4.1% with a prediction time of 404.63 seconds. In the case of the wide neural network by PSO, the negative rate is increased by only 0.2% than that of the original one with a decrease of 905.77 seconds in the prediction time. The medium neural network gives the third highest accuracy of about 95.8%. The negative rate of this classifier is 4.2% with a prediction time of 1457.7 seconds. In the case of the medium neural network by PSO, the negative rate is increased by only 0.2% than that of the original one with a decrease of 3448.8 seconds in the prediction time. The prediction time of brain tumor disease with the help of PSO is presented in Table 3. Comparison based on time shows that the linear discriminant will be executed in the minimum time of 62.48 seconds. The fine tree is the second and bagged tree is the third best classifier with a minimum prediction time of 139.28 and 326.44 seconds, respectively. The accuracies of these classifiers are 1%, 7.9%, and 1.3% with a less difference of 892.39, 815.59, and 628.43 seconds, respectively in the prediction time as compared to the cubic SVM. The confusion matrix of the cubic SVM after PSO is also shown in Figure 8.

Results gathered after the optimal feature selection by DE are shown in Table 2. Of all the classifiers, the cubic SVM gives the highest accuracy of about 96.6%. The negative rate of the cubic SVM is 3.4% with a prediction time of 2088.8 seconds. The negative rate of the cubic SVM by DE is increased by only 0.1% than that of the original one with a decrease of 2373 seconds in the prediction time. The wide neural network gives the second highest accuracy of about 96.1%. The negative rate of this classifier is 3.9% with a

Brain Tumor Classes	Brain Tumor Classes			
	T1CE	Flair	T1	T2
T1CE	13969	158	220	137
Flair	165	13730	199	112
T1	230	130	13716	147
T2	204	85	114	13976

FIGURE 8: The confusion matrix of the cubic SVM after applying the particle swarm optimization (PSO)-based feature selection.

prediction time of 701.48 seconds. The negative rate of the wide neural network by DE is the same as that of the original one with a decrease of 608.92 seconds in the prediction time. The medium neural network gives the third highest accuracy of about 95.8%. The negative rate of this classifier is 4.2% with a prediction time of 3636.3 seconds. In the case of the medium neural network by DE, the negative rate is increased by only 0.2% than that of the original one with a decrease of 1270.2 seconds in the prediction time. The prediction time of brain tumor disease with the help of DE is presented in Table 3. Now, if we compare this in terms of time, then the linear discriminant will be executed in the minimum time of 117.11 seconds. The fine tree is the second and the bagged tree is the third best classifier with a minimum prediction time of 159.25 and 460.75 seconds, respectively. The accuracies of these classifiers are 1%, 8.6%, and 1.3% with a less difference of 1971.69, 1929.55, and 1628.05 seconds, respectively, in the prediction time as compared to the cubic SVM. The confusion matrix of the cubic SVM after DE is also shown in Figure 9.

Results gathered after the PSO and DE feature fusion are shown in Table 2. Of all the classifiers, the cubic SVM gives the highest accuracy of about 96.7%. The negative rate of the cubic SVM is 3.3% with a prediction time of 3901.4 seconds. The negative rate of the cubic SVM by fusion is the same as that of the original one with a decrease of 560.4 seconds in the prediction time. The linear discriminant gives the second highest accuracy of about 95.7%. The negative rate of this classifier is 4.3% with a prediction time of 52.093 seconds. In the case of the linear discriminant by fusion, the negative rate is only increased by only 0.3% than that of the original

Brain Tumor Classes	Brain Tumor Classes			
	T1CE	Flair	T1	T2
T1CE	13954	156	231	143
Flair	165	13738	188	115
T1	237	122	13711	153
T2	215	84	118	13962

FIGURE 9: The confusion matrix of the cubic SVM after applying the differential evolution (DE)-based feature selection.

Brain Tumor Classes	Brain Tumor Classes			
	T1CE	Flair	T1	T2
T1CE	14006	133	212	133
Flair	176	13726	192	112
T1	229	126	13733	135
T2	223	75	117	13964

FIGURE 10: The confusion matrix of the cubic SVM after applying the optimal feature fusion.

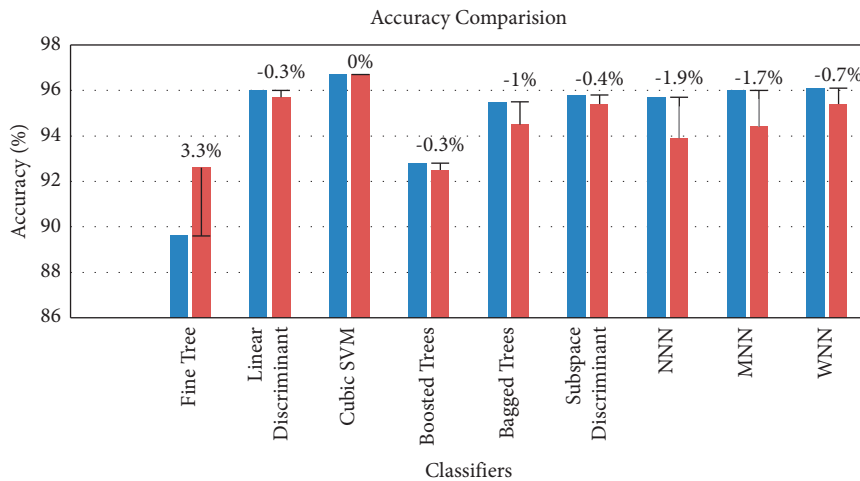


FIGURE 11: Comparison of accuracy results with state of the art.

one with a decrease of 284.897 seconds in the prediction time. The wide neural network and the subspace discriminant give the third highest accuracy of about 95.4%. The negative rates of these classifiers are 4.6% with a prediction time of 100.38 and 552.78 seconds, respectively. In the case of the wide neural network and the subspace discriminant by fusion, the negative rate is increased by only 0.7% and 0.4%, respectively, than that of the original one with a decrease of 1210.02 and 2510.42 seconds, respectively, in the prediction time. The prediction time of brain tumor disease with the help of feature fusion is presented in Table 3. Now, if we compare this in terms of time, then the linear discriminant will be executed in the minimum time of 52.093 seconds. The wide neural network is the second and the fine tree is the third best classifier with a minimum prediction time of 100.38 and 156.73 seconds, respectively. The accuracies of these classifiers are 1%, 1.3%, and 4.1% with a less difference

of 3849.307, 3801.02, and 3744.67 seconds, respectively, in the prediction time as compared to the cubic SVM. The confusion matrix of the cubic SVM after feature fusion is also shown in Figure 10.

Table 2 also shows the percentage difference between the original and fused features. This is also shown visually in Figure 11. The cubic SVM shows the same accuracy after fusion as that of the original one. The accuracy of the fine tree is increased by around 3.3%.

In case of the linear discriminant and the boosted tree, the decrease in the accuracy is only about 0.3%.

The accuracy of the subspace discriminant, wide neural network, bagged tree, medium neural network, and narrow neural network is decreased by around 0.4%, 0.7%, 1%, 1.7%, and 1.9%, respectively. From the results, it is clear that the accuracy remains the same. However, we come up with a decrease in the prediction time as compared to that of the

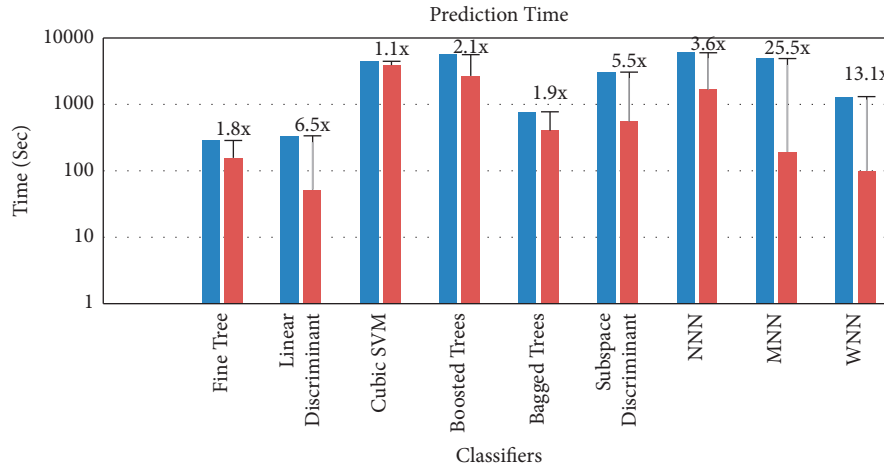


FIGURE 12: Comparison of the prediction time (logarithmic scale) with state of the art.

TABLE 4: Detailed result after feature fusion.

Classifiers	Fusion detail results				
	Sensitivity	FNR	Precision	FPR	AUC
Fine tree	92.575	7.425	92.575	0.0275	0.925
Linear discriminant	95.7	4.3	95.775	0.0125	0.9575
Cubic SVM	96.75	3.25	96.75	0.01	0.97
Boosted trees	92.5	7.5	92.8	0.0225	0.925
Bagged tree	94.475	5.525	94.55	0.02	0.945
Subspace discriminant	95.45	4.55	95.45	0.015	0.9525
Narrow neural network	93.9	6.1	93.9	0.02	0.94
Medium neural network	94.425	5.575	94.425	0.02	0.9425
Wide neural network	95.375	4.625	95.375	0.015	0.955

TABLE 5: A comparative study of the proposed methodologies on the BraTS2018 dataset.

Research papers	Maximum achieved accuracy (%)	Maximum achieved execution-time speedup
PSO features + softmax [42]	92.50	NA
PLS features + ELM [43]	93.40	NA
Two-channel DNN [44]	93.69	NA
MANet [26]	94.91	NA
Proposed	96.7	25.5x

Bold represents the best values.

original one, which we will explain later in Table 3 explanation.

Table 3 also shows the speedup between the original and fused features. This is also shown visually in Figure 12. The highest speedup which we achieved is about 25.5 times in the case of the medium neural network. In the case of the wide neural network, the speedup is 13.1 times. The speedup of 6.5, 5.5, and 3.6 times is achieved in the case of the linear discriminant, subspace discriminant, and narrow neural network. In the case of the boosted tree, bagged tree, fine tree, and cubic SVM, we achieved a speedup of 2.1x, 1.9x, 1.8x, and 1.1x, respectively.

Apart from the accuracy and prediction time results presented in Tables 2 and 3, respectively, we also report the

detailed results after feature fusion. These results are provided in Table 4 where sensitivity, FNR, precision, FPR, and AUC results are provided for various classifiers. All the classifiers have a good performance on all these metrics.

Finally, we also compare our results with the state-of-the-art techniques performing brain tumor classification on the BraTS2018 dataset. This comparison is provided in Table 5. Column 2 shows that the proposed technique has the highest accuracy as compared to the other techniques. Though this improvement in accuracy is not very high, if we bring execution time into the picture (Column 3), we can see that most of the work did not focus on this aspect and did not report it. Our work achieved an improvement of about 25.5x, which is a significant reduction in the prediction time.

5. Conclusion

The manual procedure of brain tumor detection and classification is not a good choice as it is tedious, time consuming, and expensive. A fully automated optimal deep feature fusion-based architecture is proposed in this work for brain tumor classification. A database of MRI images is prepared which consists of four different categories of tumors to perform evaluation. The proposed method achieved an accuracy of about 96.7%, which is the highest compared to the existing techniques. Based on the results presented in this work, it is observed that a few redundant and irrelevant features were still perceived. Therefore, it is essential to select the optimal features. It is also shown that the fusion of optimal features improved the accuracy, but reduction in the prediction time is quite significant, obtaining the main goal of this work. The major dark side of this work is the fusion process that increases the computational time during the testing process. In the future, lightweight deep learning frameworks will be opted, and we will utilize an optimized feature fusion approach for classification and detection of tumors [45–49].

Data Availability

The BraTs2018 dataset has been used in this work for the experimental process (<http://www.med.upenn.edu/sbia/brats2018/data.html>).

Conflicts of Interest

All authors declare that they have no conflicts of interest.

Acknowledgments

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for supporting this research through the Research Groups Program under Grant (RGP.2/16/43).

References

- [1] A. Rehman, I. Razzak, and G. Xu, “Federated learning for privacy preservation of healthcare data from smartphone-based side-channel attacks,” *IEEE Journal of Biomedical and Health Informatics*, vol. 1, p. 1, 2022.
- [2] V. Kumar and K. P. Kumar, “ConvNet based detection and segmentation of brain tumor from MR images,” in *Proceedings of the 2021 Second International Conference on Smart Technologies in Computing, Electrical and Electronics (ICST-CEE)*, pp. 1–4, IEEE, Bengaluru, India, December. 2021.
- [3] A. Rehman, M. A. Khan, T. Saba, Z. Mehmood, U. Tariq, and N. Ayesha, “Microscopic brain tumor detection and classification using 3D CNN and feature selection architecture,” *Microscopy Research and Technique*, vol. 84, no. 1, pp. 133–149, 2021.
- [4] Quick Brain, “Quick brain tumor facts | national brain tumor society,” 2022, <https://braintumor.org/brain-tumor-information/brain-tumor-facts/>.
- [5] N. B. Bahadure, A. K. Ray, and H. P. Thethi, “Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM,” *International Journal of Biomedical Imaging*, vol. 2017, Article ID 9749108, 12 pages, 2017.
- [6] S. Rosmini, A. Aggarwal, D. H. Chen et al., “Cardiac computed tomography in cardio-oncology: an update on recent clinical applications,” *European Heart Journal - Cardiovascular Imaging*, vol. 22, no. 4, pp. 397–405, 2021.
- [7] N. Tayal and S. Tayal, “Advance computer analysis of magnetic resonance imaging (MRI) for early brain tumor detection,” *International Journal of Neuroscience*, vol. 131, no. 6, pp. 555–570, 2020.
- [8] E. Pivetta, A. Goffi, M. Tizzani et al., “Lung Ultrasonography for the Diagnosis of SARS-CoV-2 pneumonia in the emergency department,” *Annals of Emergency Medicine*, vol. 77, no. 4, pp. 385–394, 2021.
- [9] S. Whitehead, S. Li, F. O. Ademuyiwa, R. L. Wahl, F. Dehdashti, and K. I. Shoghi, “Co-clinical FDG-PET radiomic signature in predicting response to neoadjuvant chemotherapy in triple-negative breast cancer,” *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 49, no. 2, pp. 550–562, 2022.
- [10] S. Ali, J. Li, Y. Pei, R. Khurram, K. Rehman, and T. Mahmood, “A comprehensive survey on brain tumor Diagnosis using deep learning and emerging hybrid techniques with multimodal MR image,” *Archives of Computational Methods in Engineering*, vol. 1, p. 3, 2022.
- [11] T. Habib, S. Boulila, W. Koubaa et al., “A survey on COVID-19 impact in the healthcare domain: worldwide market implementation, applications, security and privacy issues, challenges and future prospects,” *Complex & intelligent systems*, pp. 1–32, 2022.
- [12] A. M. Hasan, F. Meziame, R. Aspin, and H. A. Jalab, “Segmentation of brain tumors in MRI images using three-dimensional active contour without edge,” *Symmetry*, vol. 8, no. 11, p. 132, 2016.
- [13] R. Ranjbarzadeh, A. Bagherian Kasgari, S. J. Ghouschi, S. Anari, M. Naseri, and M. Bendecheche, “Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images,” *Scientific Reports*, vol. 11, no. 1, Article ID 10930, 2021.
- [14] A. K. Katta and R. B. Katta, “An automated brain tumor detection and classification from MRI images using machine learning techniques with IoT,” *Environment, Development and Sustainability*, vol. 2021, pp. 1–15, 2021.
- [15] T. Chithambaram and K. Perumal, “Automatic detection of brain tumor from magnetic resonance images (MRI) using ANN based feature extraction,” *International Journal of Advanced Research in Engineering & Technology*, vol. 12, no. 1, pp. 109–118, 2021.
- [16] M. M. Thongam and K. Thongam, “Retracted article: computer-aided detection of brain tumor from magnetic resonance images using deep learning network,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 7, pp. 6911–6922, 2020.
- [17] A. A. Abbood, Q. M. Shallal, M. A. Fadhel, and Q. M. Shallal, “Automated brain tumor classification using various deep learning models: a comparative study,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 22, no. 1, p. 252, 2021.
- [18] M. I. Sharif, M. A. Khan, M. Alhussein, K. Aurangzeb, and M. Raza, “A decision support system for multimodal brain tumor classification using deep learning,” *Complex & Intelligent Systems*, vol. 2021, pp. 1–14, 2021.

- [19] E. S. Bayoumi, M. K. Abd-Ellah, A. A. M. Khalaf, and R. R. Ghariieb, "Brain tumor automatic detection from MRI images using transfer learning model with deepe," *Convolutional Neural Network*, vol. 41, no. 2, 2022, <http://jaet.journals.ekb.eg>.
- [20] R. L. Kumar, J. Kakarla, B. V. Isunuri, and M. Singh, "Multi-class brain tumor classification using residual network and global average pooling," *Multimedia Tools and Applications*, vol. 80, no. 9, Article ID 13438, 2021.
- [21] N. Noreen, A. Palaniappan, M. O. Qayyum, S. Ahmad, and I. O. Allassafi, "Brain tumor classification based on fine-tuned models and the ensemble method," *Computers, Materials & Continua*, vol. 67, no. 3, pp. 3967–3982, 2021.
- [22] P. Kora, S. Mohammed, M. J. Surya Teja, C. Usha Kumari, K. Swaraja, and K. Meenakshi, "Brain tumor detection with transfer learning," in *Proceedings of the 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, pp. 443–446, IEEE, Palladam, India, December. 2021.
- [23] S. Gull, S. Akbar, and H. U. Khan, "Automated detection of brain tumor through magnetic resonance images using convolutional neural network," *BioMed Research International*, vol. 2021, Article ID 3365043, 14 pages, 2021.
- [24] S. Rinesh, K. Maheswari, B. Arthi et al., "Investigations on brain tumor classification using hybrid machine learning algorithms," *Journal of Healthcare Engineering*, vol. 2022, Article ID 2761847, 9 pages, 2022.
- [25] S. Safdar, M. Rizwan, T. R. Gadekallu et al., "Bio-imaging-based machine learning algorithm for breast cancer detection," *Diagnostics*, vol. 12, no. 5, p. 1134, 2022.
- [26] N. S. Shaik and T. K. Cherukuri, "Multi-level attention network: application to brain tumor classification," *Signal, Image and Video Processing*, vol. 16, no. 3, pp. 817–824, 2022.
- [27] A. M. Sarhan, "Brain tumor classification in magnetic resonance images using deep learning and wavelet Transform," *Journal of Biomedical Science and Engineering*, vol. 13, no. 06, pp. 102–112, 2020.
- [28] A. Sekhar, S. Biswas, R. Hazra, A. K. Sunaniya, A. Mukherjee, and L. Yang, "Brain tumor classification using fine-tuned GoogLeNet features and machine learning algorithms: IoMT enabled CAD system," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 3, pp. 983–991, 2022.
- [29] H. A. Ali Khan, W. Jue, M. Jue, and M. U. Mushtaq, "Brain tumor classification in MRI image using convolutional neural network," *Mathematical Biosciences and Engineering*, vol. 17, no. 5, pp. 6203–6216, 2020.
- [30] M. I. C. C. A. I. BraTS 2018, "Data | section for biomedical image analysis (SBIA) | perelman school of medicine at the university of Pennsylvania," 2022, <https://www.med.upenn.edu/sbia/brats2018/data.html>.
- [31] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conf. Comput. Vis. Pattern Recognit*, pp. 770–778, IEEE, Las Vegas, NV, USA, June 2016.
- [32] M. A. Khan, M. S. Sarfraz, M. Alhaisoni, A. A. Albeshier, S. Wang, and I. Ashraf, "StomachNet: optimal deep learning features fusion for stomach abnormalities classification," *IEEE Access*, vol. 8, Article ID 197981, 2020.
- [33] L. Shavlik and J. Shavlik, "Transfer learning," *Handbook of Research on Machine Learning Applications and Trends*, pp. 242–264, 2010.
- [34] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: a large-scale hierarchical image database," in *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, IEEE, Miami, FL, USA, March 2010.
- [35] K. Price, R. M. Storn, and J. A. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*, Springer Science & Business Media, Berlin, Germany, 2006.
- [36] A. P. Piotrowski, J. J. Napiorkowski, and A. E. Piotrowska, "Population size in particle swarm optimization," *Swarm and Evolutionary Computation*, vol. 58, Article ID 100718, 2020.
- [37] X. W. Zhang, H. Liu, and L. P. Tu, "A modified particle swarm optimization for multimodal multi-objective optimization," *Engineering Applications of Artificial Intelligence*, vol. 95, Article ID 103905, 2020.
- [38] A. Beghi, L. Cecchinato, G. Cosi, and M. Rampazzo, "A PSO-based algorithm for optimal multiple chiller systems operation," *Applied Thermal Engineering*, vol. 32, no. 1, pp. 31–40, Jan. 2012.
- [39] F. Salo, A. B. Nassif, and A. Essex, "Dimensionality reduction with IG-PCA and ensemble classifier for network intrusion detection," *Computer Networks*, vol. 148, pp. 164–175, 2019.
- [40] S. M. Holland, "Principal components a N alysi S (PCA)," 2019.
- [41] A. Krishn, V. Bhateja, and A. Sahu, "Medical image fusion using combination of PCA and wavelet analysis," in *Proceedings of the 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 986–991, IEEE, Delhi, India, November 2014.
- [42] M. I. Sharif, J. P. Li, M. A. Khan, and M. A. Saleem, "Active deep neural network features selection for segmentation and recognition of brain tumors using MRI images," *Pattern Recognition Letters*, vol. 129, pp. 181–189, 2020.
- [43] M. A. Khan, I. Ashraf, M. Alhaisoni et al., "Multimodal brain tumor classification using deep learning and robust feature selection: a machine learning application for radiologists," *Diagnostics*, vol. 10, no. 8, p. 565, 2020.
- [44] J. D. Bodapati, N. S. Shaik, V. Naralasetti, and N. B. Mundukur, "Joint training of two-channel deep neural network for brain tumor classification," *Signal, Image and Video Processing*, vol. 15, no. 4, pp. 753–760, 2021.
- [45] F. Afza, M. Sharif, M. A. Khan, U. Tariq, H. S. Yong, and J. Cha, "Multiclass skin lesion classification using hybrid deep features selection and extreme learning machine," *Sensors*, vol. 22, no. 3, p. 799, 2022.
- [46] M. A. Azam, K. B. Khan, S. Salahuddin et al., "A review on multimodal medical image fusion: compendious analysis of medical modalities, multimodal databases, fusion techniques and quality metrics," *Computers in Biology and Medicine*, vol. 144, Article ID 105253, 2022.
- [47] M. I. Sharif, J. P. Li, M. A. Khan, S. Kadry, and U. Tariq, "M3BTCNet: multi model brain tumor classification using metaheuristic deep neural network features optimization," *Neural Computing & Applications*, pp. 1–16, 2022.
- [48] M. Nawaz, T. Nazir, M. Masood et al., "Melanoma segmentation: a framework of improved DenseNet77 and UNET convolutional neural network," *International Journal of Imaging Systems and Technology*, 2022.
- [49] M. A. Khan, "Gastrointestinal diseases recognition: a framework of deep neural network and improved moth-crow optimization with dcca fusion," *Human-centric Comput. Inf. Sci*, vol. 12, p. 25, 2022.