IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 26 May 2024, accepted 22 June 2024, date of publication 17 July 2024, date of current version 29 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3430109

RESEARCH ARTICLE

Enhancing Brain Tumor Classification by a Comprehensive Study on Transfer Learning Techniques and Model Efficiency Using MRI Datasets

NADIA SHAMSHAD[1], DANISH SARWR[1], AHMAD ALMOGREN[2], (Senior Member, IEEE), KIRAN SALEEM[1], ALIA MUNAWAR[1], ATEEQ UR REHMAN[3], AND SALIL BHARANY[4]

[1]School of Software, Dalian University of Technology, Dalian 116024, China
[2]Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 11633, Saudi Arabia
[3]School of Computing, Gachon University, Seongnam 13120, Republic of Korea
[4]Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India

Corresponding authors: Salil Bharany (salil.bharany@gmail.com) and Ateeq Ur Rehman (202411144@gachon.ac.kr)

This work was supported by King Saud University, Riyadh, Saudi Arabia, through the Researchers Supporting Project under Grant RSP2024R184.

ABSTRACT Brain tumors, a significant health concern, are a leading cause of mortality globally, with an annual projected increase of 5% by the World Health Organization. This work aims to comprehensively analyze the performance of transfer learning methods in identifying the types of brain tumors, with a particular emphasis on the necessity of prompt identification. The study demonstrates how useful it is to use pre-trained models, including models VGG-16, VGG-19, Inception-v3, ResNet-50, DenseNet, and MobileNet—on MRI datasets and used to obtain a precise classification. Using these methods model accuracy and efficiency have been enhanced. The research aims to contribute to improved treatment planning and patient outcomes by implementing optimal methodologies for precise and automated brain tumor analysis, evaluation framework encompasses vital metrics such as confusion matrices, ROC curves, and the achieved Area Under the Curve (AUC) for each approach. The comprehensive methodology outlined in this paper serves as a systematic guide for the implementation and evaluation of brain tumor classification models utilizing deep learning techniques. The integration of visual representations, code snippets, and performance metrics significantly enhances the clarity and understanding of the proposed approach. Among our proposed algorithms, VGG-16 attains the highest accuracy at 97% and consumes only 22% of time as compared to our previous proposed methodology.

INDEX TERMS Brain tumors, CNNs, machine learning programming, deep learning models, VGG-16, MobileNet, ResNet-50, artificial intelligence.

I. INTRODUCTION

A brain tumor refers to irregular brain tissue growth, causing elevated pressure inside the skull and interfering with regular brain functions. Brain tumors pose a major health issue, and cancer stands as one of the most prevalent and life-threatening illnesses globally. Brain cancer stands as one of the most fatal forms of cancer, with the potential to severely impair brain functions if left untreated or overlooked [2]. The World Health Organization (WHO) forecasts a yearly increase of approximately 5% in brain tumor cases. Brain tumors are also identified as the 8th leading cause of death in the overall population. An estimated minimum of 18,600 individuals have lost their lives to brain or central nervous system (CNS) tumors in recent years [3]. Timely detection of brain tumors significantly enhances survival chances and enables less

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

invasive treatment options. Various imaging techniques allow for detailed insights into tumor specifics such as location, size, shape, and metabolism [4], [5].

In the domain of deep learning and computer vision, categorizing MRI scans of the human brain is a vital task known as brain tumor image classification [6], a field where advancements such as the FUSE-AI system developed by a Hamburg-based startup have shown significant promise. The FUSE-AI system demonstrates the capability to identify and analyze types of tumors in MRI scans using machine learning classifiers, underscoring the potential of such technologies in enhancing diagnostic accuracy and efficiency [7]. This complex procedure involves labeling these scans into distinct classes and encompasses several stages: acquiring image data, preprocessing, detecting tumors, segmenting them, extracting features, and ultimately classifying the tumors. Precise classification of brain tumors holds immense significance in healthcare, greatly assisting in accurate disease detection and diagnosis, and contributing significantly to medical advancements [8]. Brain tumor analysis and diagnosis heavily rely on image segmentation techniques, essential for transforming images into meaningful forms for assessment. Segmentation divides images, aiding in tumor area identification, yet it remains challenging due to image noise and unclear boundaries. Manual methods are time-consuming, prompting the adoption of deep learning techniques like Convolutional Neural Networks (CNNs). These methods offer automatic feature extraction, showing promise in achieving high segmentation performance, and revolutionizing medical image analysis [9].

Traditional methods like Support Vector Machine (SVM) and Neural Networks previously excelled in brain tumor classification [10]. However, recent advancements in deep learning, particularly transfer learning using models like VGG-16, MobileNet, and ResNet-50, have surpassed these approaches. Transfer learning, leveraging pre-trained models, offers efficient feature extraction from MRI scans [4], [11], demonstrating significant improvements in the precision of brain tumor classification methods [12].

A. LIMITATIONS

Analyzing brain tumors from MRI scans shows obstacles affecting model accuracy and practical adoption. Limited availability of diverse datasets hampers training, introducing biases in classifying tumor types. Variability in imaging protocols, including resolution and contrast, poses additional hurdles. Tumor heterogeneity, both intra- and inter-tumor, complicates accurate classification, requiring consideration of diverse features. Interpretability issues arise as deep learning models operate as black boxes, hindering trust in clinical settings. Generalizing models to cases with rare subtypes or different characteristics proves challenging, impacting real-world applicability. Ensuring reproducibility demands robust validation, including cross-validation and testing on multiple datasets. Integrating models into clinical workflows is challenging, requiring seamless compatibility with existing systems. Ethical and legal considerations, especially regarding patient data privacy, are critical. Additionally, the high computational resource requirements, often involving GPUs, may limit adoption. Clinical validation trials are crucial to demonstrate practical performance, utility, and accuracy for widespread adoption in real-world clinical decision-making scenarios.

B. PROBLEM STATEMENT

The motivation behind this research thesis stems from the challenges encountered in identifying and analyzing brain tumors by previous systems and equipment in hospitals. Traditional image processing tools available in scan centers often fall short of accurately identifying tumor types and segmenting specific regions of abnormal tissue. Additionally, inaccessible areas within the brain due to varying edema and tumor levels pose further difficulties. MRI is a non-invasive method for creating three-dimensional tomographic images of the human body that overcomes these limitations, especially for spotting lesions and anomalies in soft tissues like the brain. However, the qualitative analysis of MRI films by radiologists can be time-consuming and requires significant manpower [10]. This study seeks to use modern image processing methods and computer-aided methods to automatically identify and quantify anomalies in MRI brain images. By applying flexible computing techniques, such as transform learning and soft computing methods, this thesis seeks to achieve accurate classification of brain tumors from other soft tissues in the head. The proposed approaches aim to enhance the efficiency of the diagnostic process, reduce the reliance on manual interpretation, and improve the overall quality of brain tumor analysis in MRI scans.

C. SIGNIFICANCE

This research aims to assess the effectiveness of transfer learning methods, using publicly available datasets, to classify brain tumors based on MRI scans. The study aims to develop a reliable, automated brain tumor classification system to assist radiologists with treatment planning, accurate tumor detection, and enhancing patient outcomes. Specifically, this study focuses on assessing the effectiveness of the transfer learning method to achieve this goal. It focuses on categorizing brain tumors from MRI scans, utilizing the architectures of VGG-16, MobileNet, and ResNet-50. Moreover, this study aims to thoroughly assess the efficiency of transfer learning methods, specifically VGG-16, MobileNet, and ResNet-50, for brain tumor classification. It will compare related research outcomes, presenting experimental results from these models. Performance evaluation includes metrics like confusion matrices, ROC curves, and achieved AUC for each approach. This research strives to advance medical image analysis by implementing optimal methodologies for brain tumor classification.

D. CONTRIBUTIONS

1) **Improving Brain Tumor Classification Efficiency:** This research contributes by investigating the precision and efficiency of brain tumor classification based on MRI images enhanced by the implementation of transfer learning methods, specifically utilizing VGG-16, MobileNet, and ResNet-50 architectures.
2) **Comparative Analysis of Transfer Learning Models:** This research contributes by comparing three transfer learning models—ResNet-50, MobileNet, and VGG-16–in terms of their precision and computational requirements. The analysis offers valuable insights into the strengths and efficiency of each model.
3) **Impact of Dataset-Specific Evaluation on Model Performance:** The study contributes by examining how evaluating pre-trained models on a specific brain tumor dataset influences classification accuracy and the models' generalization capacity in transfer learning. This analysis aims to understand the dataset's role in shaping model performance.
4) **Optimizing Hyperparameters for Enhanced Results:** The research contributes by identifying the best hyperparameters and training methods for transfer learning models, aiming to produce optimal results in classifying brain tumors. This optimization effort enhances the practical implementation and robustness of the proposed models.
5) **Application of Transfer Learning Models:** The study contributes by exploring how the created transfer learning models can assist radiologists in identifying, diagnosing, and treating tumors accurately in the clinical setting. This contribution highlights the potential practical implications of the research for improving patient care.
6) **Key Deliverables for Evaluation and Comparison:** The study provides key deliverables, including a comprehensive evaluation of the effectiveness of transfer learning techniques, specifically VGG-16, MobileNet, and ResNet-50 architectures, using performance metrics such as confusion matrix, ROC curve, and AUC.

II. LITERATURE WORK

A. MACHINE LEARNING METHODS OVERVIEW

Making use of pre-trained models, like the VGG-16 created by the University of Oxford's Visual Geometry Group, has greatly facilitated the ability to identify brain tumors. Leveraging MRI images, this model's 16-layer structure effectively learns and accurately categorizes brain tumors. Meanwhile, the VGG-19, with its deeper 19-layer design, provides enhanced performance, albeit with higher complexity. Alternatively, DenseNet, a widely used architecture, fosters feature reuse and propagation through dense connectivity among layers, further refining tumor classification methods. DenseNet-121 and DenseNet-169 are applied in brain tumor classification for their ability to detect crucial tumor patterns. Google's Inception-v3, using inception modules, offers computational efficiency and complex feature learning in this classification. ResNet-50, addressing gradient issues, captures intricate medical image details effectively. MobileNet, tailored for mobile devices, and its variations aid real-time brain tumor classification on resource-limited devices. Researchers fine-tune these pre-trained models on extensive MRI datasets for accurate classification.

B. BRAIN TUMOR CLASSIFICATION TECHNIQUES

Various medical imaging techniques like SPECT and CT scans are commonly employed for brain tumor detection [13]. These scans provide details on tumor position, size, and shape [4]. The deep CNN method classifies brain tumors into four main categories: Glioma, Meningioma, Pituitary, and Healthy. this method improves recovery and data analysis, making MRI brain results more efficient. [30].CNNs are specifically utilized in research for tumor classification neural network characteristics to suggest brain tumor diagnosis by programming. Completing the usage of small holes is the most important aspect of building. With a 97.5 accuracy rate, CNN is less predictable [4], [29]. Labeled brain image datasets train algorithms to distinguish tumor classes using learned patterns. VGG-16, MobileNet, and ResNet-50 are convolutional neural network (CNN) models that provide the precision and efficiency of brain tumor detection. A large dataset of brain tumors is used for testing and developing the model [32]. Transfer learning with CNN models VGG-16, MobileNet, or ResNet-50, fine-tuned for specific tasks, yields promising results. Preprocessing steps, like normalization and denoising, enhance image quality, crucial for accurate tumor classification. Extracting pertinent features from medical images significantly aids in classification, as seen in studies by [14]. Various techniques extract discriminative features like shape, texture, and intensity from images to differentiate tumor types [11]. Traditional machine learning techniques such as SVM, Random Forests, or k-NN, reliant on manual features from labeled data, classify tumors [15], [16]. Data optimization, involving alterations like rotation or noise addition, expands and diversifies training datasets, reducing overfitting and boosting algorithm generalization [17], [18].

C. MACHINE LEARNING VS DEEP LEARNING

Deep learning models, unlike traditional machine learning, grasp data structures and hierarchies to classify brain tumors based on detailed and accurate descriptive information. There's a shift from manual characteristics to data-driven methods in tumor classification, supported by deep learning capabilities [19]. Convolutional neural networks (CNN) serve as a useful approach to identifying brain tumors offering a variety of methods and innovations. Multiple methods with different datasets have been demonstrated through analysis, preprocessing, ROI segmentation, and personalized vs. pre-trained models. For instance, [20] utilized enhanced MRI scans of meningioma, glioma, and pituitary tumors from different perspectives. Preprocessing techniques conducted from the illustrations involved scaling, normalization, and augmentation (90-degree rotation, vertical flip). Using Glorot weights as the initialization, the Adam optimizer was utilized to instruct a CNN classifier using a mini-batch size of 16. The model was evaluated using 10-fold cross-validation. The findings revealed that the overall accuracy was 95.4%, and the sensitivity scores for meningiomas, gliomas, and pituitary tumors were 89.8%, 96.2%, and 98.4%, respectively. The specificity scores were 90.2%, 95.5%, and 97.7%, with an F1-score of 94.94%. The healthcare industry makes use of a variety of deep learning techniques, including CNN, 3-D CNN, and LSTM [4], [17], [19], [21].

Using magnetic resonance imaging (MRI) scans, The study of the literature emphasizes a thorough analysis of several machine learning as well as deep learning methods of classifying tumors in the brain. However, due to inter-observer error, radiologists' manual review of MRI scans may produce different results for different medical professionals, We will develop a system that utilizes convolutional neural networks (CNN) and transfer learning algorithms to identify different types of brain tumors from MRI images [31]. Notably, The use of pre-trained models combined with traditional machine-learning techniques has greatly enhanced the accuracy and efficiency of tumor identification and classification. Despite these advancements, there remains a gap in the comprehensive evaluation of the effectiveness of transfer learning methods across a broader spectrum of MRI datasets, particularly in comparing the performance of different architectures like VGG-16, MobileNet, and ResNet-50 within the same experimental framework. Furthermore, although these models have been used for classification tasks in previous research, there is a discernible difference in the preprocessing, feature extraction, and data augmentation approaches used, which may affect the systems' stability and adaptability in real-time clinical settings.

This study seeks to bridge these gaps by conducting a thorough evaluation of transfer learning techniques to classify brain tumors more effectively. By leveraging publicly available datasets and employing a consistent evaluation framework across multiple pre-trained architectures, this research aims to create an automatic method for categorizing brain tumors. Such a system would not only aid radiologists in treatment planning but also enhance the precision of tumor detection, ultimately improving patient outcomes. Through a detailed comparison of related research outcomes and a comprehensive performance evaluation—including metrics like confusion matrices, ROC curves, and achieved AUC—this research aims to promote MRI results by Identifying the more efficient transfer learning methods for the classification of brain tumors. Aims to push the boundaries of current knowledge and present a foundation for upcoming studies in this crucial field of medical technology.

III. PROPOSED METHODOLOGY

Employing a systematic approach, our proposed methodology ensures precise brain tumor classification from MRI scans. Leveraging deep learning techniques like transfer learning with ResNet-50, MobileNet, and VGG-16 extracts significant features. Further, feature reduction enhances efficiency and prevents overfitting. Investigating SVM, Random forest (RF), K-NN, and Multilayer Perceptron (MLP) algorithms aid in tumor classification using these reduced features. We'll assess model accuracy and efficiency to gauge their performance accurately. Fig. 1 shows the workflow of the methodology in broad strokes. It shows how the various steps—from data collection to the completed classification models—flow in a sequentially manner. The diagram functions as an illustration for the methodology, emphasizing how the different steps are connected to one another and how they all contribute to achieving the ultimate objective of accurately classifying brain tumors.

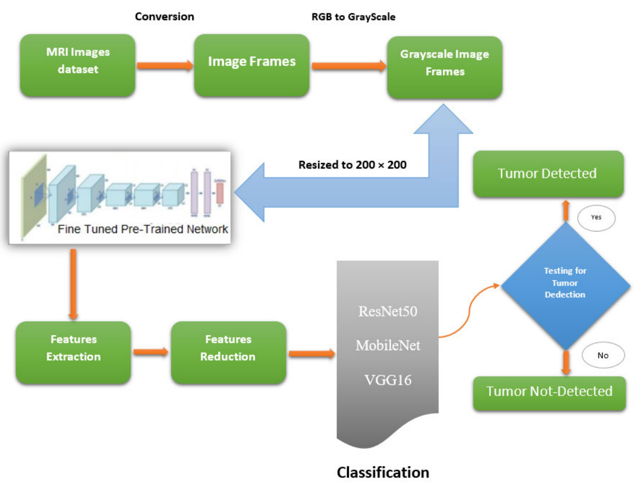


FIGURE 1. Proposed model.

A. DATASETS

This research's dataset consists of results from MRIs of the brain. There are about 256 raw MRI results, each with unique dimensions (measured in pixel ratio). The sample MRI brain results 2 were obtained using the Kaggle dataset [33], and according to the [34] the Joint Photographic Experts Group (JPEG) style is used for these collected results.

B. MODEL FLOW DIAGRAM

This is the flow diagram 3 for our proposed model, which shows the sequence of the multiple process steps.

C. DATA PRE-PROCESSING

To utilize applicable CNN models with the Kaggle brain imaging 4 dataset, it's essential to adhere to the subsequent preprocessing protocols.

1) Import the necessary packages.

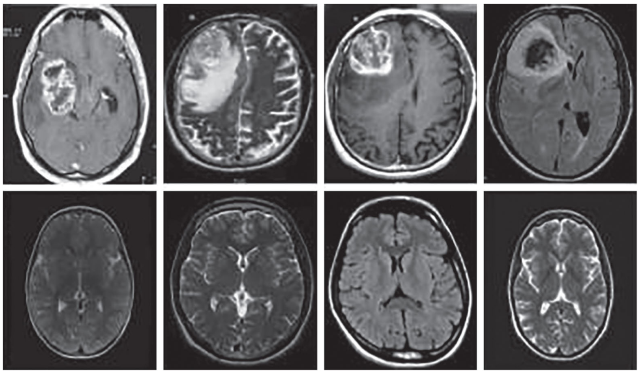


FIGURE 2. Sample dataset of brain MRI results [33].

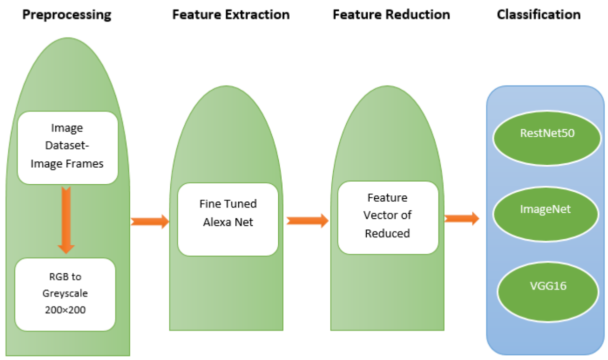


FIGURE 3. Model flow diagram.

2) Access the two data directories containing images categorized as "Positive" and "Negative."
3) Load and convert the results into a tagged format, where "Tumor" corresponds to "Positive" and "No Tumor" translates to "Negative."
4) Store the MRI results via their respective labels in data frames.
5) Reshape the results to a size of 256 × 256 pixels.
6) Employ results cropping and the specified mathematical formula to standardize the results:

$$i = (i - \mu_i)/\sigma_i. \tag{1}$$

7) In Equation (1) symbol 'i' stored resized results.
8) The preprocessing steps for the available MRI results in the dataset were performed as illustrated in Figure 4.

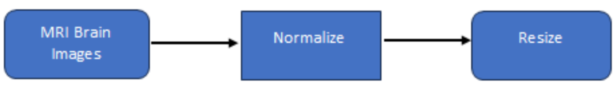


FIGURE 4. The preprocessing steps for the MRI results dataset.

Since different scanners and acquisition times produce different results, normalizing pixel intensity is essential to ensuring a consistent statistical distribution in brain MRI analysis. False positives resulting from MRI image analysis errors, such as poor resolution, deformation, and motion variation, can affect the course of treatment for the patient. Pretrained CNN models need images that are resized; the typical dimensions are 224 × 224 × 3 [22]. Initially, AI techniques are used for cropping to isolate the brain region. Data augmentation techniques such as scaling, cropping, resizing, flipping, rotating, and applying perspective transformations are used to address the issue of limited data for CNN training. When compared to the original dataset, augmented data enhances accuracy and model performance. The dataset of enhanced MRI brain images 5 produced by affine transformations and pixel-level adjustments is presented in the figure.

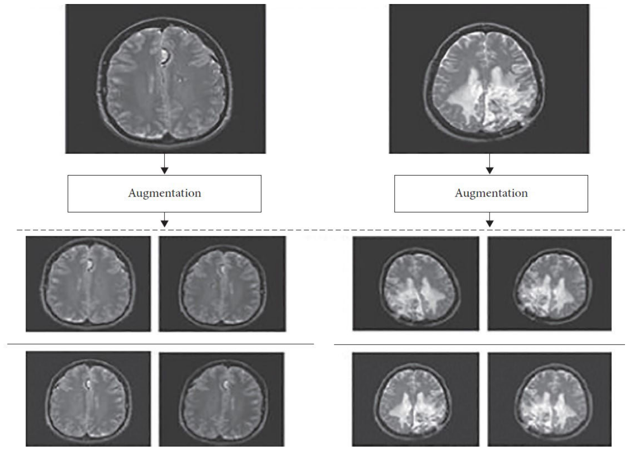


FIGURE 5. The brain MRI image dataset has been expanded through augmentation techniques.

D. CONVOLUTIONAL NEURAL NETWORK (CNN)

Advanced AI techniques, particularly deep transfer learning, excel in predicting medical conditions through image categorization tasks. These models, often built on CNNs with multiple layers like hidden, pooling, and fully connected layers, form a robust foundation. In the realm of MRI brain image categorization, convolution layers using filters are crucial, while pooling layers aid in computational efficiency and mitigating overfitting by reducing spatial representation dimensions. CNNs act as feature extractors, getting rid of the requirement for manually extracting features. and automatically capturing and categorizing relevant features in MRI images. Activation functions like ReLU play a key role in this process, enhancing model performance through mathematical operations.

$$\text{softmax}(z_i) = \frac{exp(z_i)}{\sum_j exp(z_j)} \tag{2}$$

In Equation (2) a vector **z** with n features that represent desired values is given to the SoftMax function. This input vector's z_i elements can all be either positive or negative. e^{z_j} is the result of multiplying each vector element x_i by the exponential function. Valid probability allocations are ensured by the normalization factor e^{z_j}, which is the sum of these exponentials. The suggested model is implemented using Python, which is available on platforms such as Google Colab, Jupyter Notebook, and Anaconda. 120 epochs of training are conducted using different datasets for testing, validation, and training.

E. THE TRANSFER LEARNING METHOD

Transfer learning in advanced deep learning algorithms requires substantial data and high processing power. It leverages pretrained CNN models, like those trained on ImageNet, modifying their parameters for similar tasks, making it efficient and avoiding the need to create CNN models from scratch. This approach minimizes training time and resources by reusing knowledge, making it valuable for healthcare data with small sample sizes. Despite ImageNet's focus on natural images, transfer learning is applied in healthcare, requiring fine-tuning for MRI data. It addresses challenges in computer vision and limited training data, utilizing the ImageNet dataset frequently. Pre-trained models' convolution layers frequently remain unmodified, and their weights are passed for the supreme classification, to advance CNN models, representing an improved learning strategy applying past knowledge to new tasks. See Figure 6 for an overview of common transfer learning procedures.

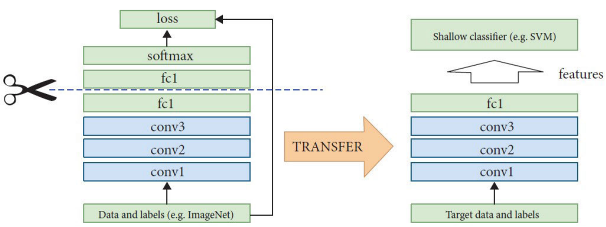


FIGURE 6. Transfer learning model.

F. CONVOLUTIONAL NEURAL NETWORKS (CNNS) BASED DEEP LEARNING METHOD.

The research involves developing a brain tumor prediction model for MRI images. It begins by collecting MRI brain data from a Kaggle dataset [23] and applying preprocessing methods like scaling, trimming, and pixel-level enhancement. Pretrained CNN models are then used to predict brain tumor presence during the stages of testing and training. The research is concentrated on glioma, meningioma, and pituitary tumors, evaluating model performance relying on dataset division. Utilizing pre-trained architectures like VGG-16, MobileNet, and ResNet-50 from datasets such as ImageNet and Kaggle aids in detecting brain tumors. The dense layer in the CNN plays a significant role in image categorization by utilizing results from convolution layers [24].

G. CNN IMPLEMENTATION ON VGG-16

We used the VGG-16 CNN model in our experiments, which was first trained on a limited set of images. Certain Conv layers were frozen for additional model optimization to avoid overfitting. This 2014 model receives 224 × 224 x 3-dimensional brain MRI images and has sixteen convolution layers. Using a lot of ReLU activation functions, it consists of two ultimately connected layers, a SoftMax output layer, max-pooling layers of size 2 × 2, and fixed 3 × 3 filter size Conv layers. The VGG-16 model forms deep neural networks with about 138 million hyperparameters with an emphasis on convolutional layers to learn intricate features. These parameters—activation functions, training %, neuron and batch numbers, and compilation techniques—determine how the model behaves. To maximize the model's performance, hyperparameters such as the number of convolution layers must be tuned. Enhancing the depth of a ConvNet improves its ability to learn hidden features at a more economical expense. In Figure 7, it shows the VGG-16 model structure.

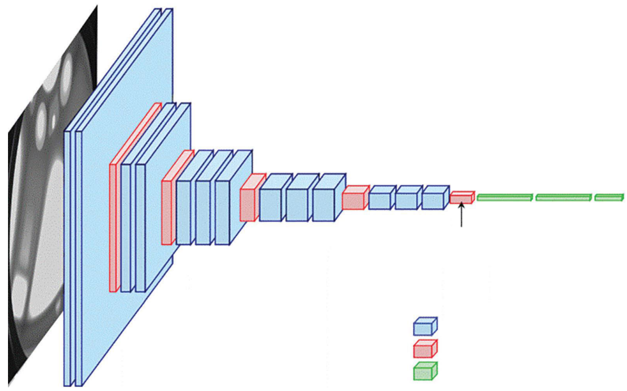


FIGURE 7. VGG-16 model architecture [25].

H. CNN MODEL IMPLEMENTED WITH RESNET-50

Regarding image classification tasks A 50-layer residual network identified as ResNet-50 was developed in 2015 by Kaiming He et al. at Microsoft Research [26]. Unlike conventional deep CNNs, it simplifies training by learning residual features by deducting learned features from each layer's input. ResNet-50, which was trained on ImageNet, establishes direct connections between layers for deeper networks by using skip connections and substantial batch normalization to preserve image classification accuracy. In the attempt, we used a pre-trained ResNet-50 model modified for our image dataset, which had a lower time complexity than VGG-16 or VGG-19.

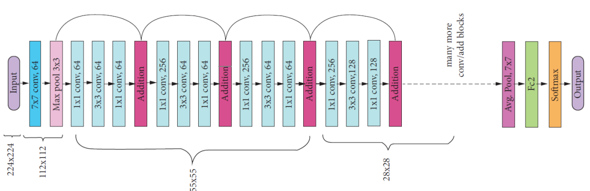


FIGURE 8. RestNet-50 model architecture [27].

I. RESNET-50, MOBILENET AND VGG-16 BASED FEATURES EXTRACTION.

In this research, feature extraction is performed using refined pre-trained networks, specifically ResNet-50, MobileNet, and VGG-16. These networks extract relevant features related to brain tumors from MRI scans by retaining the convolutional layers and removing the classification layers. Pre-processing, which uses augmentation, normalization, and resizing techniques for resilience and generalization, guarantees compatibility with these networks. By capturing a variety of spatial and frequency information, the convolutional layers use MRI images to create feature maps. By utilizing information from extensive datasets, these maps help in the classification of brain tumors by acting as inputs for additional analysis. The information extracted from features is essential for differentiating between different types of tumors and healthy brain tissues.

J. EVALUATION METHOD

Researchers use diverse methods to check the model's efficiency. which involves common metrics like precision, F-measure, recall, specificity, and accuracy. The model's results will be compared with those of recent studies to gauge its performance against current standards.

K. CONFUSION MATRIX

The confusion matrix acts as a grid illustrating system errors. Rows signify human-annotated instances, while columns depict machine-annotated ones. It offers insights into misclassifications, serving as a performance gauge for the proposed model. Here, TP (True Positive) represents accurate positive predictions by the machine, while FP (False Positive) signifies incorrect positive forecasts. Likewise, FN (False Negative) indicates mistaken negative identifications, and TN (True Negative) refers to correct negative classifications by the system.

TABLE 1. Confusion matrix Table.

	Positive	Negative
Positive	TP	FN
Negative	FP	TN

L. PRECISION

Precision assesses the accuracy of positive predictions. Higher precision means fewer negatives are wrongly labeled as positives. Conversely, lower precision indicates more negatives are inaccurately identified as positives. Elevated precision signifies strong detection of true positives, reflecting the accuracy of positive predictions.

$$Precision, P(Positive) = \frac{TP}{TP + FP} \tag{3}$$

M. RECALL

Recall measures how accurately occurrences are identified in the entire text corpus. Lower misidentification of positive sentences corresponds to higher recall. Table 8 displays true positive and false negative values, represented mathematically.

$$Recall, R(Positive) = \frac{TP}{TP + FN} \tag{4}$$

$$Recall, R(Negative) = \frac{TN}{TN + FP} \tag{5}$$

N. F-MEASURE

The harmonic means that computes the average of accuracy and recall, expressed mathematically as follows:

$$F-Measure = \frac{2 * R * P}{R + P} \tag{6}$$

$$F - \text{Measure Positive} = \frac{2TP}{(2TP + FP + FN)} \tag{7}$$

$$F - \text{Measure Negative} = \frac{2TN}{(2TN + FN + FP)} \tag{8}$$

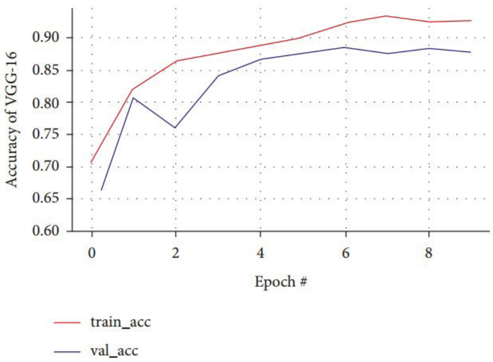
O. ACCURACY

To gauge the closeness of a measurement to the true value and understand digit significance, a standard mathematical rule is applied.

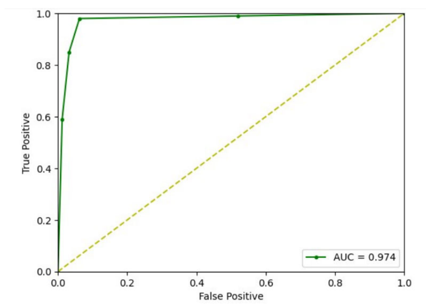
$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

IV. RESULT ANALYSIS AND DISCUSSION

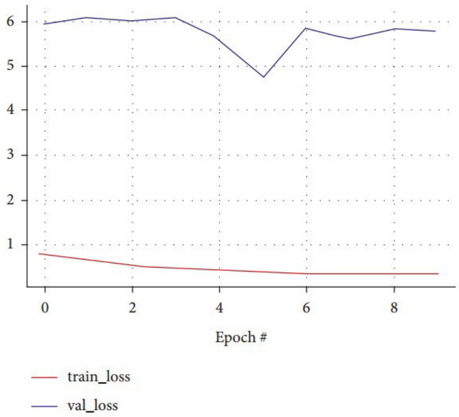
This discussion explores pre-trained CNN models' effectiveness in categorizing brain tumor MRI images, employing VGG-16, MobileNet, and ResNet-50. The dataset comprises 3064 brain tumor MRI images from 233 patients, with different image sizes scaled down to 200 × 200 pixels. VGG-16 utilizes 3 × 3 convolution kernels, contributing to 138 million hyperparameters. MobileNet uses modules to reduce convolution layers, while ResNet-50 accommodates numerous layers without increasing training error significantly. The images are preprocessed and assessed using metrics like accuracy and loss. The analysis uses 2100 scans for training and 900 for validation on a cloud-based GPU virtual machine, displaying predictive graphs based on accuracy and loss over epoch's. Moreover, Recent advancements in medical image processing simplify early disease identification. Medical informatics aids in leveraging extensive medical records. Timely detection of brain tumors is crucial, in guiding treatment decisions. This study proposes an innovative feature ensemble for accurate MR scan-based tumor classification, outperforming existing methods like CNN-based approaches. Our proposed model, especially VGG-16, achieves higher accuracy rates— 97.2% compared to the 96.9% obtained by previous methods. Across VGG-16, MobileNet, and RestNet-50, accuracy scores of 0.97%, 0.87%, and 0.96% were achieved. Additional metrics



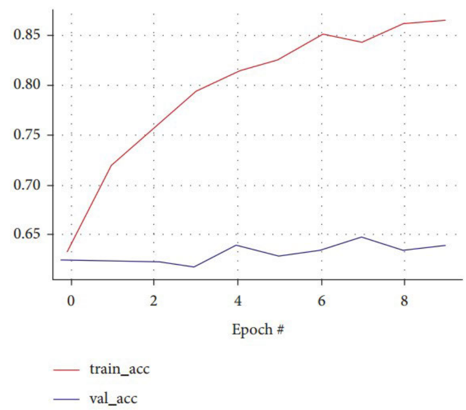
((a)) Accuracy curve of VGG-16



((b)) Classification model of VGG-16



((c)) Training and validation loss curve of MobileNet model



((d)) Training and validation accuracy curves of MobileNet model

FIGURE 9. Performance metrics of the VGG-16 and MobileNet models.

like recall, F1-score, and processing time are provided for each method in Table 8 is presented.

A. VGG-16 MODEL IMPLEMENTATION

a: VGG-16 MODEL IMPLEMENTATION RESULTS

Table 2 presents true label data and predicted label data.

TABLE 2. True label and predicted data.

	Predicted label	
True label	No	Yes
No	128	36
Yes	22	2254

Fig. 11 shows the loss accuracy VGG-16 with multiple epochs in training and validation of curve form.

Fig. 9 illustrates two curves depicting the progression of VGG-16 accuracy across epochs in both training and validation datasets.

In the following Fig. 9(b), the VGG-16 ROC (Receiver Operating Characteristic) curve displays the classification model's performance. It demonstrates, for a range of thresholds, the relationship between true positive rates (sensitivity) and false positive rates (1-specificity). The success of the suggested method is confirmed by the VGG-16 model's remarkable accuracy in recognizing brain tumors in this dataset, which has an AUC value of 0.974.

Metrics for the VGG-16 model, such as accuracy, precision, recall, F1-score, and processing time, are presented in Table 3.

TABLE 3. Matrics values for VGG-16 model.

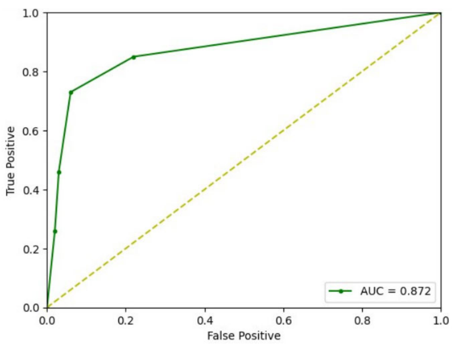
Model	Accuracy	Precision	Recall	F1-score	Time [s]
VGG-16	0.97	0.95	1.0	0.98	4085

B. MOBILENET MODEL IMPLEMENTATION

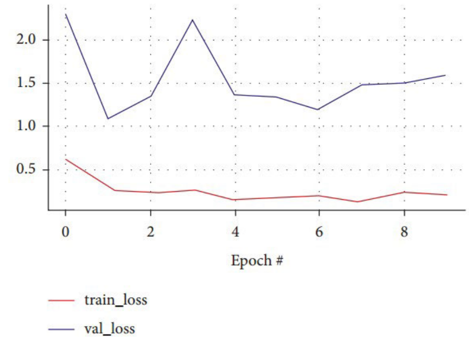
In MobileNet, training accuracy steadily increases, while validation accuracy fluctuates. Validation accuracy notably improves by 0.87%. Graphs in Fig. 9(c) and 9(d) show erratic validation loss but steady training accuracy. The Mobile-net model employs a confusion matrix to anticipate events in table-4.

In Fig. 9(c), the two curves represent the loss accuracy of MobileNet in different epochs in validation and training curves.

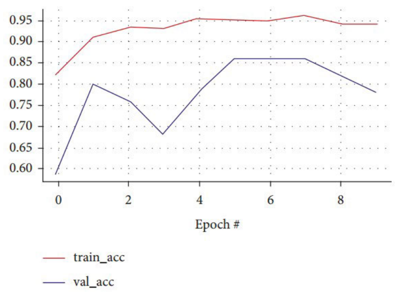
The accuracy of MobileNet across different epochs is depicted in the two curves displayed in Fig. 9(d), illustrating the training and validation curves.



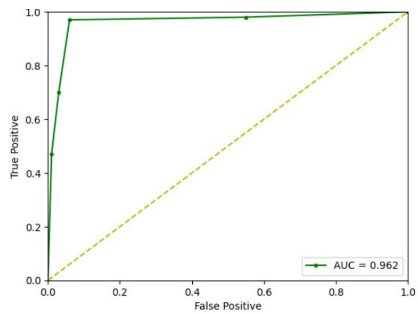
((a)) AUC value of MobileNet



((b)) Training and validation loss of the ResNet-50 model



((c)) Training and validation accuracy of the ResNet-50 model



((d)) Classification model of ResNet-50

FIGURE 10. Performance metrics of the MobileNet and ResNet-50 models.

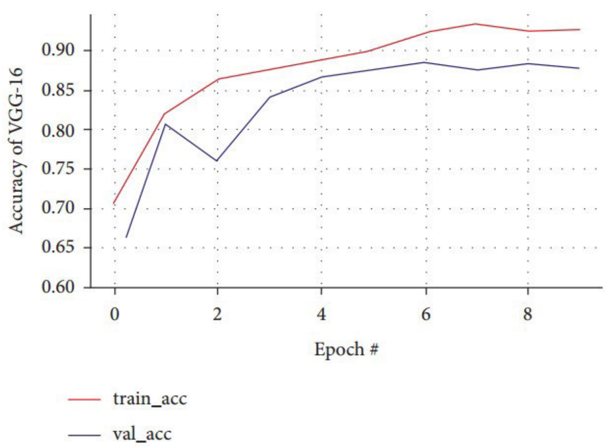


FIGURE 11. Training and validation of VGG-16 model.

TABLE 4. Confusion matrix of mobile net model.

	Predicted label	
True label	No	Yes
No	88	23
Yes	69	1383

In Fig. 10(a), The AUC value of 0.872 indicates MobileNet's strong ability to effectively differentiate tumor and non-tumor samples within the dataset, showcasing its precision and accuracy in classification.

TABLE 5. MobileNet model evaluation.

Model	Accuracy	Precision	Recall	F1-score	Times [s]
MobileNet	0.87	0.84	0.79	0.82	6403

Table 5 shows the MobileNet model's processing time, F1-score, recall, accuracy, and precision.

C. RESNET-50

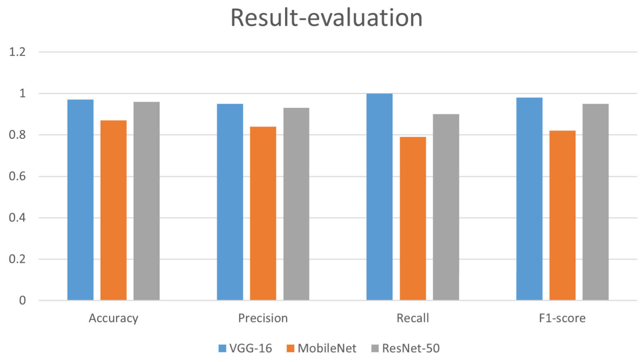
In the ResNet-50 experiment, training accuracy steadily rises, maintaining above 0.96% from the 2nd epoch. Validation accuracy improves initially but declines after the 18th epoch to 0.94%. Consequently, the model struggles to predict new data accurately, as shown in Fig. 10(b) and 10(c) via a confusion matrix for tumor forecasting Table 6.

TABLE 6. Confusion Matrix of ResNet-50 Model.

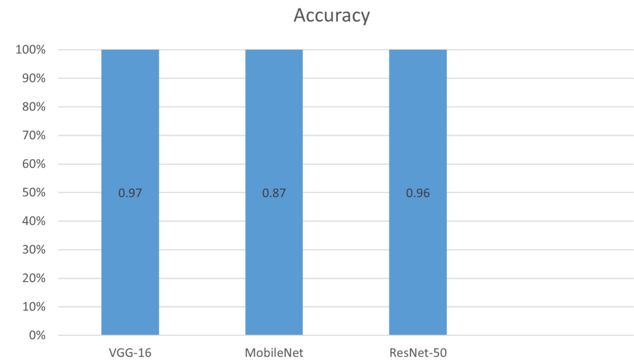
	Predicted label	
True label	No	Yes
No	66	18
Yes	21	2093

Fig. 10(b), displays two curves illustrating the loss and accuracy of ResNet-50 across various epochs, depicted in training and validation curve formats.

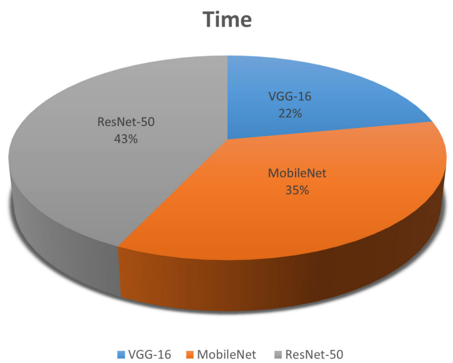
Fig. 10(c) illustrates two curves showcasing the accuracy of ResNet-50 across different epochs, represented in training and validation curve formats.



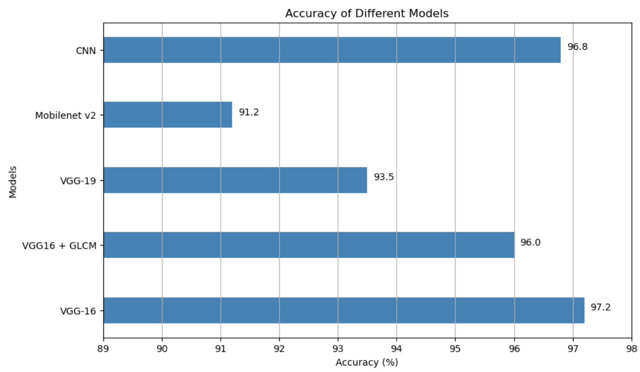
((a)) Result evaluation graph



((b)) Proposed model accuracy



((c)) Proposed model computation time



((d)) Proposed study comparison with existing models

FIGURE 12. Results of the suggested model.

Fig. 10(d)'s AUC value validates the proposed ResNet-50's performance, demonstrating its capability to handle dataset complexities. With a high AUC of 0.962, it showcases strong accuracy in brain tumor classification, promising effective identification and significant implications for medical applications.

Table 7 shows the RestNet-50 model's accuracy, precision, recall, F1 score, and processing time.

TABLE 7. ResNet-50 model evaluation.

Model	Accuracy	Precision	Recall	F1-score	Time [s]
ResNet-50	0.96	0.93	0.90	0.95	7969

Fig. 12(a) compares the evaluation results of three algorithms: VGG-16 with an accuracy of 97%, MobileNet with 87%, and ResNet-50 with 96%. The precision values are 95%, 84%, and 93% for VGG-16, MobileNet, and ResNet-50, respectively. In terms of recall, VGG-16 scores 100%, while MobileNet and ResNet-50 achieve 79% and 90%. Lastly, for F1-score, VGG-16 records 98%, MobileNet 82%, and ResNet-50 95%.

D. DISCUSSION

Fig. 12(b) and 12(c) VGG-16, MobileNet, and ResNet-50 are three deep learning models whose performance examined concerning processing time and accuracy. Fig. 12(b) shows

TABLE 8. Study result of all the algorithms in terms of accuracy.

Model	Accuracy	Precision	Recall	F1-score	Time [s]
VGG-16	0.97	0.95	1.0	0.98	4085
MobileNet	0.87	0.84	0.79	0.82	6403
ResNet-50	0.96	0.93	0.90	0.95	7976

the proportion of total processing time each model requires. VGG-16, which takes 22% of the total time, is the fastest among the three models. MobileNet, occupying 35% of the time, sits in the middle, while ResNet-50 is the slowest, taking 43% of the total time. Fig. 12(c) depicts the accuracy of each model. VGG-16 achieves the highest accuracy at 97%, followed closely by ResNet-50 with 96%, and MobileNet has the lowest accuracy at 87%. This comparison highlights the trade-offs between processing speed and accuracy: VGG-16 offers the best accuracy with the fastest processing time, making it ideal for scenarios where accuracy is critical.

Fig. 12(d) compares the accuracy of five models: CNN, MobileNet v2, VGG-19, VGG16 + GLCM, and VGG-16. VGG-16 achieves the highest accuracy at 97.2%, making it the most accurate model in the comparison. Close behind is CNN with an accuracy of 96.8%, also performing exceptionally well. The combination of VGG16 with GLCM scores 96.0%, slightly less accurate than VGG-16 and CNN. VGG-19 achieves 93.5% accuracy, performing moderately well but not as high as the previous models. MobileNet v2 has the lowest accuracy at 91.2%, reflecting its design for efficiency over precision.

V. COMPARATIVE ANALYSIS WITH PRIOR STUDIES AND BENCHMARKING.

Previous studies by [28] leveraged CNNs to identify brain tumors in MRI results, via bounding boxes initially to locate tumors before classification. In [4] is utilized MobileNet v2, achieving 92% accuracy, while Sharma et al. [27] applied transfer learning with VGG-19, achieving 94% accuracy. Kibriya et al. [25] employed VGG-16 + GLCM on brain MRI scans, reaching 96% accuracy. In contrast, our approach, highlighted in Fig. 12(c) and Table 9, introduces a hybrid feature set, notably with VGG-16 boasting the highest accuracy at 97.2%. This method stands out for its efficiency in brain tumor identification and classification.

TABLE 9. Comparison with existing techniques.

Reference	Algorithm	Accuracy
[30]	CNN	96.7
[4]	MobileNet v2	92
[29]	VGG-19	94.7
[27]	VGG16+GLCM	96
Proposed method	VGG-16	97.2

VI. CONCLUSION

This study compares pre-trained CNN architectures—VGG-16, MobileNet, and ResNet-50 to categorize benign and malignant brain tumors in MRI images. Despite achieving high training accuracy, overfitting issues arise, impacting validation accuracy. VGG-16 demonstrates superior precision and closely aligns authentication correctness with accuracy metrics. Future research could explore additional pre-trained CNN models to further enhance brain tumor predictive analysis using image data.

REFERENCES

[1] D. Lamrani, B. Cherradi, O. E. Gannour, M. A. Bouqentar, and L. Bahatti, "Brain tumor detection using MRI images and convolutional neural network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 7, pp. 452–460, 2023.

[2] D. Saravagi, S. Agrawal, M. Saravagi, S. K. Jain, B. Sharma, A. Mehbodniya, S. Chowdhury, and J. L. Webber, "Predicting lumbar spondylolisthesis: A hybrid deep learning approach," *Intell. Autom. Soft Comput.*, vol. 37, no. 2, pp. 2133–2151, 2023.

[3] R. Anitha and D. S. S. Raja, "Development of computer-aided approach for brain tumor detection using random forest classifier," *Int. J. Imag. Syst. Technol.*, vol. 28, no. 1, pp. 48–53, Mar. 2018.

[4] T. Tazin, S. Sarker, P. Gupta, F. I. Ayaz, S. Islam, M. Monirujjaman Khan, S. Bourouis, S. A. Idris, and H. Alshazly, "A robust and novel approach for brain tumor classification using convolutional neural network," *Comput. Intell. Neurosci.*, vol. 2021, pp. 1–11, Dec. 2021, Art. no. e2392395, doi: 10.1155/2021/2392395.

[5] J. Tidd, *Managing Innovation: Integrating Technological, Market and Organizational Change*. Hoboken, NJ, USA: Wiley, 2020.

[6] M. Halicek, J. D. Dormer, J. V. Little, A. Y. Chen, and B. Fei, "Tumor detection of the thyroid and salivary glands using hyperspectral imaging and deep learning," *Biomed. Opt. Exp.*, vol. 11, no. 3, p. 1383, Mar. 2020.

[7] I. Ahmad, I. Ullah, W. U. Khan, A. Ur Rehman, M. S. Adrees, M. Q. Saleem, O. Cheikhrouhou, H. Hamam, and M. Shafiq, "Efficient algorithms for e-healthcare to solve multiobject fuse detection problem," *J. Healthcare Eng.*, vol. 2021, pp. 1–16, May 2021.

[8] A. Raza, H. Ayub, J. A. Khan, I. Ahmad, A. S. Salama, Y. I. Daradkeh, D. Javeed, A. Ur Rehman, and H. Hamam, "A hybrid deep learning-based approach for brain tumor classification," *Electronics*, vol. 11, no. 7, p. 1146, Apr. 2022.

[9] C. Srinivas, M. Zakariah, Y. A. Alothaibi, K. Shaukat, B. Partibane, and H. Awal, "Deep transfer learning approaches in performance analysis of brain tumor classification using MRI images," *J. Healthcare Eng.*, vol. 2022, pp. 1–17, Mar. 2022.

[10] F. Itoo, Meenakshi, and S. Singh, "Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection," *Int. J. Inf. Technol.*, vol. 13, no. 4, pp. 1503–1511, Aug. 2021.

[11] S. Das, O. F. M. R. R. Aranya, and N. N. Labiba, "Brain tumor classification using convolutional neural network," in *Proc. 1st Int. Conf. Adv. Sci., Eng. Robot. Technol. (ICASERT)*, Singapore, 2019, pp. 1–5.

[12] K. Lamba and S. Rani, "Optimization of MRI- based residual network for prediction of brain tumor disorder using transfer learning," in *Proc. Int. Conf. Comput., Commun., Intell. Syst. (ICCCIS)*, Greater Noida, India, Nov. 2023, pp. 643–648, doi: 10.1109/ICCCIS60361.2023.10425427.

[13] S. Gull, S. Akbar, and K. Safdar, "An interactive deep learning approach for brain tumor detection through 3D-magnetic resonance images," in *Proc. Int. Conf. Frontiers Inf. Technol. (FIT)*, Dec. 2021, pp. 114–119.

[14] D.-E.-M. Nisar, R. Amin, N.-U.-H. Shah, M. A. A. Ghamdi, S. H. Almotiri, and M. Alruily, "Healthcare techniques through deep learning: Issues, challenges and opportunities," *IEEE Access*, vol. 9, pp. 98523–98541, 2021.

[15] E. Irmak, "Multi-classification of brain tumor MRI images using deep convolutional neural network with fully optimized framework," *Iranian J. Sci. Technol., Trans. Electr. Eng.*, vol. 45, no. 3, pp. 1015–1036, Sep. 2021.

[16] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: An overview and application in radiology," *Insights into Imag.*, vol. 9, no. 4, pp. 611–629, Aug. 2018.

[17] H. Yahyaoui, F. Ghazouani, and I. R. Farah, "Deep learning guided by an ontology for medical images classification using a multimodal fusion," in *Proc. Int. Congr. Adv. Technol. Eng. (ICOTEN)*, Jul. 2021, pp. 1–6.

[18] Y. Xie, A. Li, L. Gao, and Z. Liu, "A heterogeneous ensemble learning model based on data distribution for credit card fraud detection," *Wireless Commun. Mobile Comput.*, vol. 2021, pp. 1–13, Jul. 2021.

[19] S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes, "A survey of MRI-based medical image analysis for brain tumor studies," *Phys. Med. Biol.*, vol. 58, no. 13, pp. R97–R129, Jul. 2013.

[20] S. Ali, J. Li, Y. Pei, R. Khurram, K. U. Rehman, and T. Mahmood, "A comprehensive survey on brain tumor diagnosis using deep learning and emerging hybrid techniques with multi-modal MR image," *Arch. Comput. Methods Eng.*, vol. 29, no. 7, pp. 4871–4896, Nov. 2022.

[21] M. A. H. Tuhin, T. Pramanick, H. K. Emon, and W. Rahman, "Detection and 3D visualization of brain tumor using deep learning and polynomial interpolation," Ph.D. thesis, Dept. Comput. Sci. Eng., Brac Univ., Dhaka, Bangladesh, 2019.

[22] M. A. B. Siddique, S. Sakib, M. M. R. Khan, A. K. Tanzeem, M. Chowdhury, and N. Yasmin, "Deep convolutional neural networks model-based brain tumor detection in brain MRI images," in *Proc. 4th Int. Conf. I-SMAC (IoT Social, Mobile, Analytics Cloud) (I-SMAC)*, Oct. 2020, pp. 909–914.

[23] A. Tiwari, S. Srivastava, and M. Pant, "Brain tumor segmentation and classification from magnetic resonance images: Review of selected methods from 2014 to 2019," *Pattern Recognit. Lett.*, vol. 131, pp. 244–260, Mar. 2020.

[24] L. N. Smith, "Cyclical learning rates for training neural networks," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2017, pp. 464–472.

[25] H. Kibriya, R. Amin, J. Kim, M. Nawaz, and R. Gantassi, "A novel approach for brain tumor classification using an ensemble of deep and hand-crafted features," *Sensors*, vol. 23, no. 10, p. 4693, May 2023.

[26] L. Chato and S. Latifi, "Machine learning and deep learning techniques to predict overall survival of brain tumor patients using MRI images," in *Proc. IEEE 17th Int. Conf. Bioinf. Bioengineering (BIBE)*, Oct. 2017, pp. 9–14.

[27] S. Sharma, S. Gupta, D. Gupta, A. Juneja, H. Khatter, S. Malik, and Z. K. Bitsue, "Deep learning model for automatic classification and prediction of brain tumor," *J. Sensors*, vol. 2022, pp. 1–11, Apr. 2022.

[28] S. Irsheidat and R. Duwairi, "Brain tumor detection using artificial convolutional neural networks," in *Proc. 11th Int. Conf. Inf. Commun. Syst. (ICICS)*, Apr. 2020, pp. 197–203.
[29] M. Mruthyunjaya and S. K. Mandala, "A brain tumor identification using a convolution neural network and a fully convolution neural network," in *Proc. MATEC Web Conf.*, vol. 392, 2024, p. 01130.
[30] P. V. Kale and A. B. Gadicha, "Deep CNN-based MRI imaging for brain tumor detection and classification," in *Proc. IEEE Int. Students' Conf. Electr., Electron. Comput. Sci. (SCEECS)*, Bhopal, India, Feb. 2024, pp. 1–6.
[31] I. Salsabila Ardan and R. Indraswari, "Design of brain tumor detection system on MRI image using CNN," in *Proc. ASU Int. Conf. Emerg. Technol. Sustainability Intell. Syst. (ICETSIS)*, Manama, Bahrain, Jan. 2024, pp. 1388–1393.
[32] K. Bhagyalaxmi, B. Dwarakanath, and P. V. Reddy, "Ensemble of deep learning models and machine learning classifiers for the classification of brain tumors," in *Proc. 2nd Int. Conf. Comput., Commun. Control (IC4)*, Indore, India, Feb. 2024, pp. 1–6.
[33] Kaggle. (2023). *Brain Tumors 256x256*. [Online]. Available: https://www.kaggle.com/datasets/thomasdubail/brain-tumors-256x256
[34] A. Mehmood, S. Yang, Z. Feng, M. Wang, A. S. Ahmad, R. Khan, M. Maqsood, and M. Yaqub, "A transfer learning approach for early diagnosis of Alzheimer's disease on MRI images," *Neuroscience*, vol. 460, pp. 43–52, Apr. 2021, doi: 10.1016/j.neuroscience.2021.01.002.

NADIA SHAMSHAD received the bachelor's degree in computer science from the Comsats University of Science Technology, Pakistan, and the master's degree in computer applied technology from Changchun University of Science and Technology, China, in 2019. She is currently pursuing the Ph.D. degree with Dalian University of Technology, China. After that, she was an IT Instructor at the Government Institute, Pakistan. Meanwhile, she contributed to a university research project on traffic flow prediction systems. She joined the University of Education as a Visiting Lecturer. She doing research work on a project "Underwater Water Wireless Communication."

DANISH SARWR received the bachelor's degree in computer science from the Comsats University of Technology, Pakistan, and the master's degree in computer applied technology from Changchun University of Science and Technology, China, in 2020. He is currently pursuing the Ph.D. degree with Dalian University of Technology, China. Afterward, he joined the University of Education as a Visiting Lecturer. Meanwhile, he contributed to a university research project on an E-commerce recommendation system. He is doing research work on a project "Underwater Water Wireless Communication."

AHMAD ALMOGREN (Senior Member, IEEE) received the Ph.D. degree in computer science from Southern Methodist University, Dallas, TX, USA, in 2002. He is a Professor with the Computer Science Department, College of Computer and Information Sciences (CCIS), King Saud University (KSU), Riyadh, Saudi Arabia, where he is currently the Director of the Cyber Security Chair. Previously, he was the Vice Dean of Development and Quality with CCIS. He was the Dean of the College of Computer and Information Sciences and the Head of Academic Accreditation Council, Al-Yamamah University. His research interests include mobile-pervasive computing and cyber security. He served as the General Chair for the IEEE Smart World Symposium; and a Technical Program Committee Member for numerous international conferences/workshops, such as IEEE CCNC, ACM BodyNets, and IEEE HPCC.

KIRAN SALEEM received the B.S. and M.S. degrees in computer science from the University of Lahore, in 2017 and 2020, respectively. She is currently pursuing the Ph.D. degree in software engineering with Dalian University of Technology. She was involved in a software house and developed several web and mobile applications, in 2017. From 2021 to 2022, she was a Lecturer to teach undergraduate students with the University of Lahore. Her main research interests include pervasive computing, ambient intelligence, wireless sensors, body area networks, the IoT, and e-health. She has done several publications in e-health, cognitive intelligence computing, smart computing, and wireless sensor networks.

ALIA MUNAWAR received the bachelor's degree in information technology from the University of Education, Lahore, Pakistan, in 2023. She is currently pursuing the master's degree in software engineering with Dalian University of Technology, China. Subsequently, she worked on several multi-project workload recommendation systems and E-Apps for Android. She also spent time researching the IT industry.

ATEEQ UR REHMAN received the B.S. degree in electrical (telecommunication) engineering from the COMSATS Institute of Information Technology, Lahore, Pakistan, in 2009, the M.S. degree in electrical engineering with a specialization in telecommunications from the Blekinge Institute of Technology (BTH), Karlskrona, Sweden, in 2011, and the Ph.D. degree from the College of Internet of Things (IoT) Engineering, Hohai University (HHU), Changzhou Campus, China, in 2022. He is currently an Assistant Professor with the Faculty of Computing, Gachon University, South Korea. He has contributed to various international IEEE conferences and journals of repute. His research interests include but are not limited to biomedical signal processing, the Internet of Things (IoTs), the social Internet of Things (SIoTs), big data, and renewable energy technologies.

SALIL BHARANY received the B.Tech. degree from Punjab Technical University, Punjab, India, and the M.Tech. and Ph.D. degrees in CSE from Guru Nanak Dev University, India. He joined as an Assistant Professor with more than seven years of experience in teaching undergraduate and post-graduate courses. His research interests include energy, WSN, drones, the IoT, social networks, AI, cybersecurity, and big data analytics. He regularly serves as a TPC member for conferences and reviews for leading journals.