

12-1-2022

Efficient framework for brain tumor detection using different deep learning techniques

Fatma Taher
Zayed University

Mohamed R. Shoaib
Menoufia University

Heba M. Emara
Menoufia University

Khaled M. Abdelwahab
Menoufia University

Fathi E. Abd El-Samie
Menoufia University; Princess Nourah bint Abdulrahman University

See next page for additional authors

Follow this and additional works at: <https://zuscholars.zu.ac.ae/works>



Part of the [Medicine and Health Sciences Commons](#)

Recommended Citation

Taher, Fatma; Shoaib, Mohamed R.; Emara, Heba M.; Abdelwahab, Khaled M.; El-Samie, Fathi E. Abd; and Haweel, Mohammad T., "Efficient framework for brain tumor detection using different deep learning techniques" (2022). *All Works*. 5504.

<https://zuscholars.zu.ac.ae/works/5504>

This Article is brought to you for free and open access by ZU Scholars. It has been accepted for inclusion in All Works by an authorized administrator of ZU Scholars. For more information, please contact scholars@zu.ac.ae.

Author First name, Last name, Institution

Fatma Taher, Mohamed R. Shoaib, Heba M. Emara, Khaled M. Abdelwahab, Fathi E. Abd El-Samie, and
Mohammad T. Haweel



OPEN ACCESS

EDITED BY

Bruno Bonnechère,
Universiteit Hasselt, Belgium

REVIEWED BY

Saneera Hemantha Kulathilake,
Rajarata University of Sri Lanka,
Sri Lanka
Jatinderkumar R. Saini,
Symbiosis Institute of Computer
Studies and Research (SICSR), India

*CORRESPONDENCE

Heba M. Emara
hoba-24@hotmail.com

SPECIALTY SECTION

This article was submitted to
Digital Public Health,
a section of the journal
Frontiers in Public Health

RECEIVED 06 June 2022

ACCEPTED 31 August 2022

PUBLISHED 01 December 2022

CITATION

Taher F, Shoaib MR, Emara HM,
Abdelwahab KM, Abd El-Samie FE and
Haweel MT (2022) Efficient framework
for brain tumor detection using
different deep learning techniques.
Front. Public Health 10:959667.
doi: 10.3389/fpubh.2022.959667

COPYRIGHT

© 2022 Taher, Shoaib, Emara,
Abdelwahab, Abd El-Samie and
Haweel. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Efficient framework for brain tumor detection using different deep learning techniques

Fatma Taher¹, Mohamed R. Shoaib², Heba M. Emara^{2*},
Khaled M. Abdelwahab², Fathi E. Abd El-Samie^{2,3} and
Mohammad T. Haweel⁴

¹College of Technological Innovative, Zayed University, Abu Dhabi, United Arab Emirates, ²Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt, ³Department of Information Technology, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia, ⁴Department of Electrical Engineering, Shaqra University, Shaqraa, Saudi Arabia

The brain tumor is an urgent malignancy caused by unregulated cell division. Tumors are classified using a biopsy, which is normally performed after the final brain surgery. Deep learning technology advancements have assisted the health professionals in medical imaging for the medical diagnosis of several symptoms. In this paper, transfer-learning-based models in addition to a Convolutional Neural Network (CNN) called BRAIN-TUMOR-net trained from scratch are introduced to classify brain magnetic resonance images into tumor or normal cases. A comparison between the pre-trained InceptionResNetv2, Inceptionv3, and ResNet50 models and the proposed BRAIN-TUMOR-net is introduced. The performance of the proposed model is tested on three publicly available Magnetic Resonance Imaging (MRI) datasets. The simulation results show that the BRAIN-TUMOR-net achieves the highest accuracy compared to other models. It achieves 100%, 97%, and 84.78% accuracy levels for three different MRI datasets. In addition, the k -fold cross-validation technique is used to allow robust classification. Moreover, three different unsupervised clustering techniques are utilized for segmentation.

KEYWORDS

MRI, CNN, segmentation, classification, brain tumor classification, deep neural networks, pre-trained models, transfer learning

1. Introduction

The terminology of “brain tumor” involves the growth of abnormal cells in brain tissues. It is a grouping or bulk of abnormal brain cells (1). The skull, which acts as a protective shield for the brain, is extremely rigid. Any growth inside such a confined place might be dangerous. Brain tumors are categorized as being malignant (cancerous) or benign (not cancerous). There are two types of brain cancer: primary and secondary. A primary brain tumor develops within the brain, where many brain tumors in their early stages are not hazardous. A secondary brain tumor, also known as a metastatic brain tumor, occurs when cancer cells move from another organ, such as the lung or breast, to the brain.

A primary brain or spinal cord tumor develops within the brain or spinal cord. Primary malignant tumors of the brain and spinal cord have been detected in 24,530 people in the United States (13,840 males and 10,690 females). The likelihood of developing this type of tumors within one's lifetime is less than 1%. Approximately, 85–90% of all early malignancies are brain tumors. According to Cancer Net (2), brain or central nervous system tumors were detected in about 3,460 children under the age of 15. Pressure or headache around the tumor, loss of balance, and problems with fine motor skills are all symptoms of brain tumors. According to Cancer Net (2), a pineal gland tumor can induce vision alterations such as loss of eyesight, double vision, and inability to gaze upward.

Several researchers compared Computed Tomography (CT) with MRI for brain tumor diagnosis. MRI is more sensitive but less specific (3). There is an ability with MRI to detect abnormalities that are undetected or just faintly visible on CT. When CT scans indicate only a hazy aura, MRI may be used to confirm the tumor exact scope and location. Additionally, MRI with superior contrast discrimination and the ability to record images at several levels can aid in pinpointing the precise site of the lesion with respect to important neuroanatomical structures. As a consequence, we propose a framework for brain tumor detection from MRI datasets with multiple deep learning models.

The main problem considered in this paper is the classification of different brain tumor cases from magnetic resonance images based on segmentation after a first classification stage. This paper is concerned with the utilization of CNN models with different learning strategies for brain tumor detection. First, we consider the transfer-learning-based approach for brain tumor detection from magnetic resonance images. Different pre-trained deep learning models, namely, InceptionResNetv2, Inceptionv3, and ResNet50 are considered and compared for the task of brain tumor detection from magnetic resonance images. The second approach is the training of a proposed CNN model called BRAIN-TUMOR-net from scratch. The classification gives a decision about the case, whether anomalous or not. After that, the suspicious area is segmented. Three different datasets have been considered with different sizes and characteristics (4–6). The main achievement of this approach is the high accuracy of classification with simple implementation.

2. Related work

Several machine and deep learning algorithms have been proposed for detecting brain tumors from magnetic resonance and CT images. Several findings confirm the importance of MRI and image processing tools to identify brain tumors. The MRI scanners are used to create images of organs in the body, for cases such as fractures, bone dislocations, lung infections, pneumonia, and COVID-19.

Sindhumol et al. (7) introduced an approach for improving brain tumor classification from magnetic resonance images using spectrum angle-dependent feature extraction and Spectral Clustering Independent Component Analysis (SCICA). The magnetic resonance images are firstly divided into clusters depending on spectral distance. Then, Independent Component Analysis (ICA) is applied on the clustered data. A Support Vector Machine (SVM) is used for the classification process. Rating was done using $T1$ weighted, $T2$ weighted, and proton density fluid inversion recovery images. To determine the stability and effectiveness of SC-ICA-based classification, a comparison with ICA-based SVM and other conventional classifiers was performed. For a recurrent lesion, ICA-based SVM analysis achieves 98% for accuracy. Hemanth et al. (8) introduced a CNN-based automated segmentation approach. This approach comprises pre-processing, average filtering, segmentation, feature extraction, and a Neural Network (NN) for classification. An accuracy of 91% has been attained. Mallick et al. (9) suggested an image compression strategy based on a Deep Wavelet Auto-encoder (DWA). A Deep Neural Network (DNN) is employed in the classification phase. An accuracy of 96% has been acquired.

Anaraki et al. (10) presented an approach for MRI brain tumor identification based on CNNs and Genetic Algorithms (GAs). In addition, an ensemble approach was used to reduce the variation of prediction error. For classifying three glioma grades, an accuracy of 96% was obtained. Nalepa et al. (11) provided an end-to-end classification approach for Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE-MRI). This strategy attained a 99% accuracy. Amin et al. (12) presented an automated approach for detecting brain tumors from MRI datasets. For the segmentation of potential lesions, several approaches have been used. For the classification procedure, the SVM classifier was used. It achieved an average accuracy of 98%. Gupta et al. (13) developed a non-invasive approach for tumor identification from $T2$ -weighted MRI. Pre-processing improves the magnetic resonance images, which were then segregated using the multilayer customization of the Otsu thresholding technique. From the segmented image, several textural and form features are recovered, and two dominant ones are chosen using an entropy measure.

Sumitra and Saxena (14) introduced an NN technique for identifying magnetic resonance images for brain. It is divided into three steps: feature extraction, dimensionality reduction, and classification. Using Principal Component Analysis (PCA), important characteristics such as mean, median, variance, and correlation values of maximum and minimum intensity are obtained from magnetic resonance images. An NN is built depending on back-propagation. The classifier classifies images as normal, benign, or malignant based on the category to which they belong. The classification accuracy on a brain imaging testing dataset was 73%. Using GAs and an SVM, Jafari and Shafaghi (15) developed a hybrid approach for categorizing

brain tumor tissues in MRI datasets. The introduced system has four stages. Noise reduction and contrast enhancement are done during pre-processing in the first stage. The second stage involves segmentation. Morphological operations are used to remove the skull from the images. The selection and extraction of features is the third stage. The features are classified into four categories: static features, Fourier and wavelet transform histograms, and a mixture of them. The features are chosen using GAs. Finally, the selected features are fed into the SVM classifier, which achieves an accuracy of 83.22% in detecting normal and abnormal activities.

Jayachandran and Dhanasekharan developed a hybrid algorithm for diagnosing brain tumors from magnetic resonance images, based on statistics and SVM classifiers (16). Noise reduction, feature extraction, feature reduction, and classification are the four utilized steps of this algorithm. To reduce noise and prepare the image for feature extraction, the anisotropic filter is used. Using the Gray Level Co-occurrence Matrix (GLCM), the texture features are then extracted. The extracted features are then reduced using PCA. Finally, an SVM classifier is utilized for classification. It yields an accuracy of 95.80%. Selvapandian et al. (17) proposed a Non-Sub-Sampled Contourlet Transform-based (NSCT) method for brain tumor diagnosis. The classification procedure is carried out using the Adaptive Neuro Fuzzy Inference System (ANFIS). After that, morphological functions are used to segment the tumor sections in the glioma brain images.

A learning-based system for robust and automated nucleus segmentation with shape preservation was suggested by Xing et al. (18). Initial deep CNN filtering is followed by iterative region merging segmentation using a selective sparse shape model. It makes use of the benefit of faster computations, making it suitable for real-time applications. This system achieves a sensitivity of 89% and an accuracy of 85%. Narayana and Reddy (19) introduced a median filter GA segmentation technique for the segmentation operation. With an SVM classifier, the GLCM is used including the features. An accuracy of 91.23% has been obtained. Zaw et al. (20) developed an algorithm for detecting tumor locations in distinct brain magnetic resonance images, and predicting whether or not the discovered region is a tumor. Pre-processing, pixel removal, maximum entropy cut-off, statistical feature extraction, and a Naive Bayes classifier are used. The accuracy of this algorithm was 94%. Veeramuthu et al. (21) proposed a Combined Feature and Image-based Classifier (CFIC) for brain tumor classification. This approach was evaluated using the kaggle brain tumor detection 2020 dataset. It has given a sensitivity, a specificity, and an accuracy of 98.86, 97.14, and 98.97%, respectively.

An algorithm for detecting brain tumors was developed by Astina Minz and Chandrakant Mahobiya. It revealed lower error rates and required less training time, but it has a limitation that it can only optimize the margin for features that have previously been described (22). This algorithm achieved an accuracy of

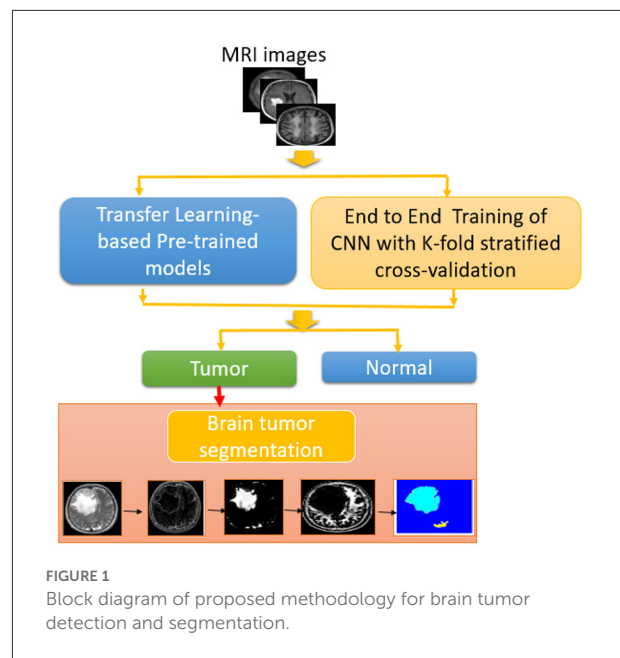


FIGURE 1
Block diagram of proposed methodology for brain tumor detection and segmentation.

89.90% and a precision of 74%. Raju et al. (23) used Bayesian fuzzy clustering segmentation, information-theoretic scatters, and wavelet features for brain tumor diagnosis. An accuracy of 93% has been achieved. Sert et al. (24) presented Single Image Super Resolution (SISR) and a Maximum Fuzzy Entropy Segmentation (MFES) method for brain tumor detection and segmentation. For feature extraction and classification, the ResNet model and the SVM were employed, respectively. An accuracy of 95% has been obtained. Deepak et al. (25) presented a classification technique for extracting features from brain magnetic resonance images based on transfer learning with GoogLeNet. To categorize the extracted features, the SVM classifier was used. The presented method achieved an accuracy of 98%.

3. Materials and methods

3.1. Datasets

Three different datasets of MRI brain tumors are used to evaluate the proposed approach. The datasets are briefly described in this section. There are 155 images for tumor cases and 155 images for normal cases for the first MRI brain tumor dataset (4). The second MRI brain tumor dataset (5) includes 1,500 images for tumor cases and 1,500 images for normal subjects. The third MRI brain tumor dataset (6) comprises 5,504 images for tumor cases and 6,159 images for normal subjects.

The proposed approach is shown in Figure 1. Magnetic resonance images are used as input to the proposed brain tumor

detection approach. Different CNN-based models combined with segmentation techniques, namely, transfer-learning-based models and an end-to-end CNN model have been studied and compared. As previously indicated, the ResNet50, Inceptionv3, and InceptionResNetv2 were used in the transfer-learning-based models. The k -fold stratified cross-validation was used to train the BRAIN-TUMOR-net model from scratch. An input layer, three convolutional layers, three Rectified Linear Unit (ReLU) layers, and three Batch Normalization (BN) layers make up the proposed BRAIN-TUMOR-net model structure. For dimensionality reduction, two pooling layers are employed. A Fully-Connected (FC) layer, a softmax layer, and a classification layer are used at the end of the model.

3.2. Transfer-learning-based approach

Deep learning from scratch is a time-consuming process that requires data classification and division. Transfer learning is ideal for removing the huge strain of this process. According to the input characteristics, transfer learning causes little modifications in deep pre-trained networks. The used dataset is partitioned into two datasets randomly, with a 75/25 training/testing ratio. The pre-trained models were loaded, and the BN, ReLU, and softmax layers were substituted for the last three FC layers.

3.3. State-of-the-art CNNs for transfer learning

Recent CNN models for brain tumor detection are addressed in this section.

- **ResNet.** Deep residual learning network is a new tool for training very deep neural networks. Identity mapping is used for shortcut connections in the deep residual learning network. It is a new way for training very deep neural networks. In a range of computer vision challenges, ResNet exceeded the state-of-the-art networks and won the ImageNet ILSVRC 2015 classification competition.
- **Inceptionv3.** Its architecture is based on Szegedy et al. publication “Rethinking the Inception Architecture for Computer Vision” (2015), which presented an improvement to the inception module to enhance ImageNet classification accuracy, dramatically (26). The authors proposed Inceptionv2 and Inceptionv3 (27). Factorization, which breaks convolutions into smaller convolutions and other minor adjustments to Inceptionv1, were introduced in Inceptionv2. The typical 7×7 convolution has been factored into three 3×3 convolutions. However, Inceptionv3 is a version of

Inceptionv2 that includes a BN-auxiliary. The BN-auxiliary refers to the variant in which the fully-linked layer of the auxiliary classifier, rather than merely convolutions, is normalized. The model [Inceptionv2 + BN-auxiliary] is referred to as Inceptionv3. The Inception module reduces the grid size, which expands the filter banks.

- **InceptionResNetv2.** It is a convolutional neural architecture that uses residual connections from Inception designs. The residual connection takes the place of the filter concatenation stage (28). This network is able to classify 1000 categories.

The dataset is divided into three parts randomly, with the ratio of 75/25 for training or validation/testing. After loading the pre-trained models, the last three fully-connected layers were replaced with BN, ReLU, and softmax layers. The training approaches used in this paper demonstrated their ability to control the degradation problem, while also providing the required convergence in a short time. Due to its high convergence and short running duration, Stochastic Gradient Descent (SGD) is employed for training (29). The ReLU is used to activate all convolutional layers. The main objective of the proposed approach is to combine image batch identification with a fine-tuned classifier to classify many instances as tumor or normal cases (30).

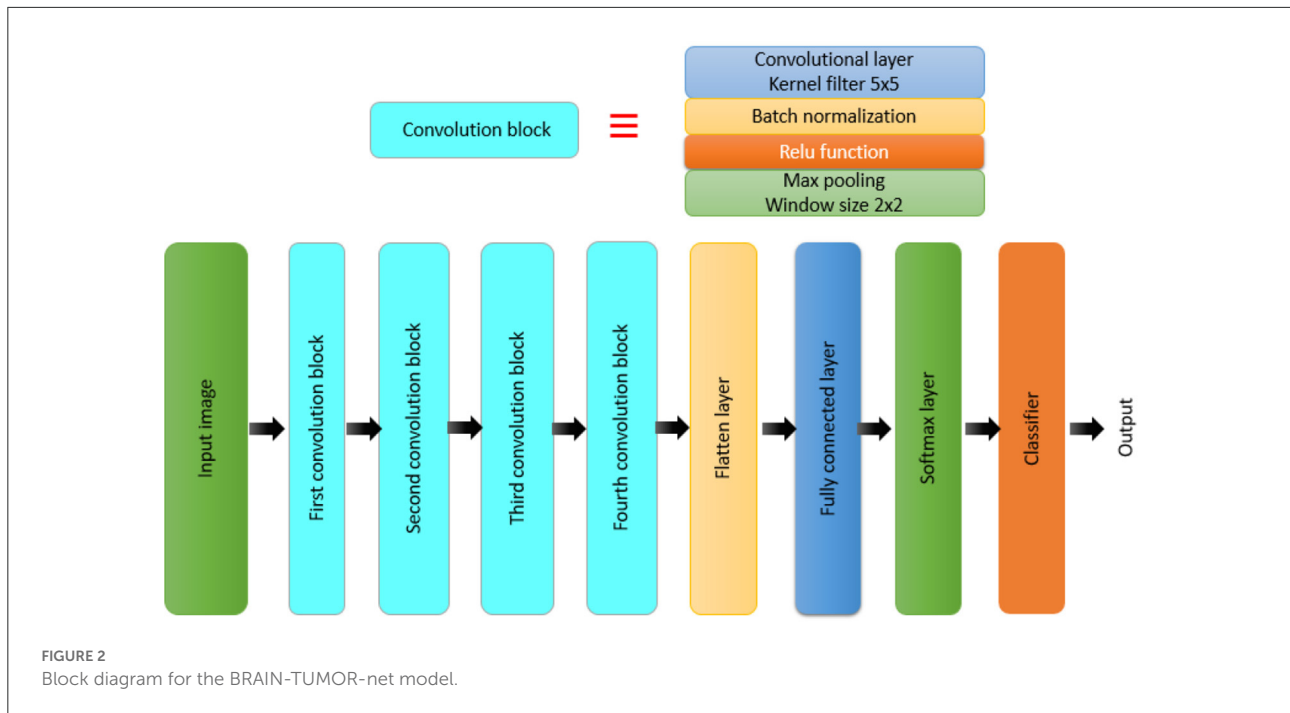
3.4. Convolutional neural network trained from scratch

Deep learning models have been employed in a variety of medical data classification, segmentation, and lesion detection applications. Medical imaging techniques such as MRI, X-ray, and CT are used to generate medical images. Machine learning and deep learning models may be evaluated on MRI, CT, and X-ray datasets. This paper provides a number of CNN-based deep learning models for identifying tumor instances by categorizing magnetic resonance images as normal or tumor cases (31–34).

Furthermore, a model is built from scratch for the classification task. Figure 2 presents the structure of the BRAIN-TUMOR-net.

A CNN model is made up of several layers, including an input layer, convolutional layers, pooling layers, FC layers, and an output layer (32, 33, 35–38). The proposed BRAIN-TUMOR-net is constructed as follows:

- **Input layer.** The inputs are magnetic resonance images with a resolution of 224×224 pixels.
- **COVN layers.** The convolutional layer (Conv), the BN layer, and the ReLU layer make up the COVN layers. The convolutional layer captures and compresses image



features to create feature maps. As a consequence, convolutions were conducted over the input images for the first, second, third, fourth, and fifth Conv layers, with different filters (8, 16, 32, 64, and 128) and a fixed window size of 3. The BN layers are used in optimization to reduce overfitting and improve test accuracy. The activations of the preceding layer are normalized for each batch during training. To incorporate element-wise non-linearity, a ReLU activation function is used.

- Pooling layer. This layer is used to extract the most important features from each feature map. We use the max-pooling method for pooling operations. The max-pooling layer vectors are concatenated to form a fixed-length feature vector. The stride is set to 2 and the max-pooling window is set to 2×2 .
- Fully-Connected (FC) layer. It takes a simple vector as input and returns a single vector as output. The proposed model has four FC levels. The last layer is an FC output layer with softmax activation for classifying the input images into two categories.

The principal structural components of a CNN network are convolution, BN, and pooling layers. The convolution layers extract the local features, and the BN layers normalize them. Pooling layers are used to minimize the number of extracted features. To reflect fluctuations in local activity levels, max-pooling is used. It displays the edges with considerable details. The highest values obtained are mostly associated with edges. Magnetic resonance images are rich

in details. The resulting feature map can be represented as follows:

$$Y_j^l = f\left(\sum_{i \in N_j} Y_i^{l-1} * X_{ij}^l + b_j^l\right) \quad (1)$$

where Y_j^l represents the local features obtained from the previous layer, X_{ij}^l represents the adjustable kernels. In order to prevent the overfitting, the bias is used and denoted by b_j^l . The pooling process is implemented as follows:

$$Y_j^l = \text{down}(Y_j^{l-1}) \quad (2)$$

where $\text{down}(\cdot)$ represents the down-sampling function. The FC layers have full connections to all activations in the previous layer. The FC layer provides discriminative features for the classification of the input image into various classes.

Whenever a classification task, whether binary or multi-class classification, is performed, the data is divided into train and test sets, and the model is trained to improve the accuracy (39). Numerous performance metrics and data splitting mechanisms are becoming increasingly important (40). As a result, stratified k -folds and a variety of performance measures may be employed to help in the development of a reliable deep-learning-based model (41). The model precise accuracy cannot be determined, since the model accuracy is altered by modifying the random state values. It samples the data without consideration of class distributions. In case

of binary classification, and out of a 100% dataset, 80% belong to class 0 and the remainder to class 1. Through the utilization of random sampling to achieve this balance, there is a strong chance to have different class distributions between training and testing. Tearing on such a dataset will result in inaccurate results.

The most popular validation technique is the k -fold technique. The division of the training dataset into k -folds is known as cross-validation. The first $k - 1$ folds are used for training, while the remaining fold is used for testing. This process is repeated for each fold. k folds are fitted and evaluated collectively, and the mean accuracy for all of these folds is returned. This technique produced promising results for balanced classification issues, but it did not work for imbalanced classes. This is because cross-validation randomly divides the data without taking into consideration class imbalance. As a result, rather than splitting the data randomly, the solution is to stratify it. The stratified k -fold cross-validation technique is a variation of the cross-validation commonly used for classification issues. It maintains the same class ratio as in the original dataset throughout the k -fold technique. So, by using a stratified k -fold technique, the same class ratio may be maintained throughout all k folds (42). The essential configuration option for k -fold cross-validation is the number of folds k (43). When the number k is set too high, the bias of the actual error rate estimator becomes minimal, but the estimator variance and time consumption become large. If k is small, the calculation time decreases, and the estimator variance decreases, but the estimator bias increases (44). The most common values are $k = 3$, $k = 5$, and $k = 10$. As a consequence, if a maximum classification accuracy is required, the ideal value of k needs to be chosen. The value of k in this paper is set to 5.

3.5. MRI brain tumor segmentation approach

The pixel value properties of the MRI datasets are used to segment the data. On the MRI datasets, the following steps were used in the segmentation process:

1. Post-processing:

- Image enhancement.
- Utilization of the usual shrink denoising process to remove noise from the images.
- Edge preservation by applying a bilateral filter method on the denoised output.

2. Segmentation:

- Edge-based segmentation using a 3×3 mask and the Kirsch operator. A 3×3 mask is utilized to implement

the Sobel operator. Also, a 5×5 mask is utilized to build the sophisticated Sobel operator.

- Magnetic resonance brain image segmentation based on thresholds. The Otsu threshold algorithm is implemented.
- Clustering-based segmentation. The k -means clustering method with a predetermined number of iterations and a certain value of k was developed. The adaptive k -means clustering technique was used. The number of iterations to convergence has been determined. In addition, the fuzzy c -means clustering technique was used.
- Watershed algorithm with marker control. The segmentation function was accomplished using the Watershed technique with gradient magnitude.

3.6. Performance metrics

The proposed approach performance is assessed using conventional metrics such as sensitivity (SEN), specificity (SPEC), accuracy (ACC), precision (PRECI), Matthews Correlation Coefficient (MCC), $F1_score$, kappa, and false positive rate (F_{pr}). The number of successfully detected anomalous cases (T_p) is the true positive. The number of accurately detected normal instances (T_n) is the true negative. A false positive (F_p) is a collection of normal instances designated as anomaly diagnoses. A false negative (F_n) is a collection of abnormalities seen as normal (45, 46).

Sensitivity is given as:

$$SEN = \frac{T_p}{T_p + F_n} \times 100 \quad (3)$$

Specificity is given as:

$$SPE = \frac{T_n}{T_n + F_p} \times 100 \quad (4)$$

Accuracy is given as:

$$ACC = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \times 100 \quad (5)$$

Precision is given as:

$$PRECI = \frac{T_p}{T_p + F_p} \quad (6)$$

Matthews correlation coefficient (MCC) is defined as:

$$MCC = \frac{T_p \times T_n - F_p \times F_n}{\sqrt{(T_p + F_p) \times (T_p + F_n) \times (T_n + F_p) \times (T_n + F_n)}} \times 100 \quad (7)$$

False positive rate is given as:

$$F_{pr} = \frac{F_p}{T_n + F_p} \quad (8)$$

F1_score is given as:

$$F1_score = \frac{T_p}{T_p + \frac{1}{2}(F_p + F_n)} \times 100 \tag{9}$$

Kappa coefficient is defined as:

$$kappa = \frac{2 \times (T_p \times T_n - F_n \times F_p)}{(T_p + F_p) \times (F_p + T_n) \times (T_p + F_n) \times (F_n + T_n)} \times 100 \tag{10}$$

4. Simulation results

The proposed approach is evaluated on three different publicly-available datasets. The performance of the CNN models differs from one dataset to another according to the size of the dataset.

4.1. Results on the first MRI dataset

Table 1 summarizes the results for the proposed approach in terms of SEN, SPEC, ACC, PRECI, MCC, F_{pr} , $F1_score$,

TABLE 1 Detection performance results for various CNN models with a 75/25 training/testing ratio on the first dataset.

Models	Evaluation metric								
	ACC	SEN	SPE	PRECI	F1_score	MCC	Error	Kappa	FPR
InceptionResNetv2	0.9130	0.9348	0.8913	0.8958	0.9149	0.8269	0.0870	0.8261	0.1087
Inceptionv3	0.8804	0.9130	0.8478	0.8571	0.8842	0.7625	0.1196	0.7609	0.1522
ResNet50	0.9348	0.9348	0.9348	0.9348	0.9348	0.8696	0.0652	0.8696	0.0652
Transfer learning model	0.9130	0.9130	0.9130	0.9130	0.9130	0.8261	0.0870	0.8261	0.0870
BRAIN-TUMOR-net	0.8478	0.8043	0.8913	0.8810	0.8409	0.6983	0.1522	0.6957	0.1087

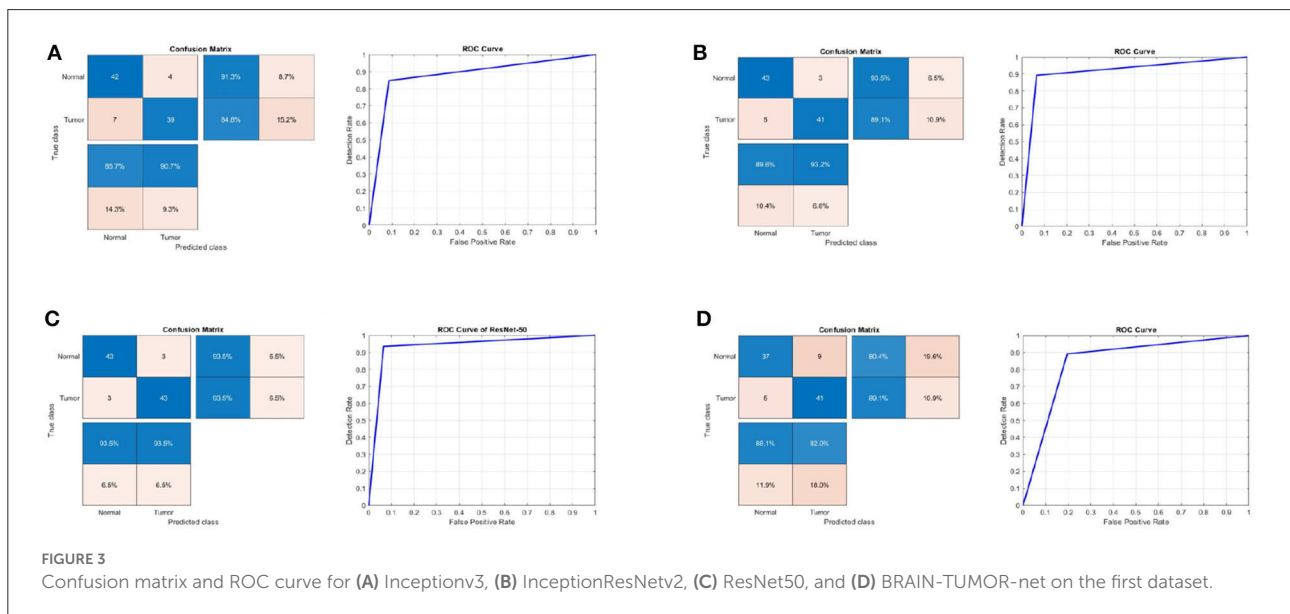


TABLE 2 Detection performance results for the proposed CNN models with a 75/25 training/testing ratio on the second dataset.

Models	Evaluation metrics								
	ACC	SEN	SPE	PRECI	F1_score	MCC	Error	Kappa	FPR
InceptionResNetv2	0.9689	0.9578	0.9800	0.9795	0.9685	0.9380	0.0311	0.9378	0.0200
Inceptionv3	0.9633	0.9622	0.9644	0.9644	0.9633	0.9267	0.0367	0.9267	0.0356
ResNet50	0.9656	0.9822	0.9489	0.9505	0.9661	0.9316	0.0344	0.9311	0.0511
BRAIN-TUMOR-net	0.9700	0.9644	0.9756	0.9753	0.9698	0.9401	0.0300	0.9400	0.0244

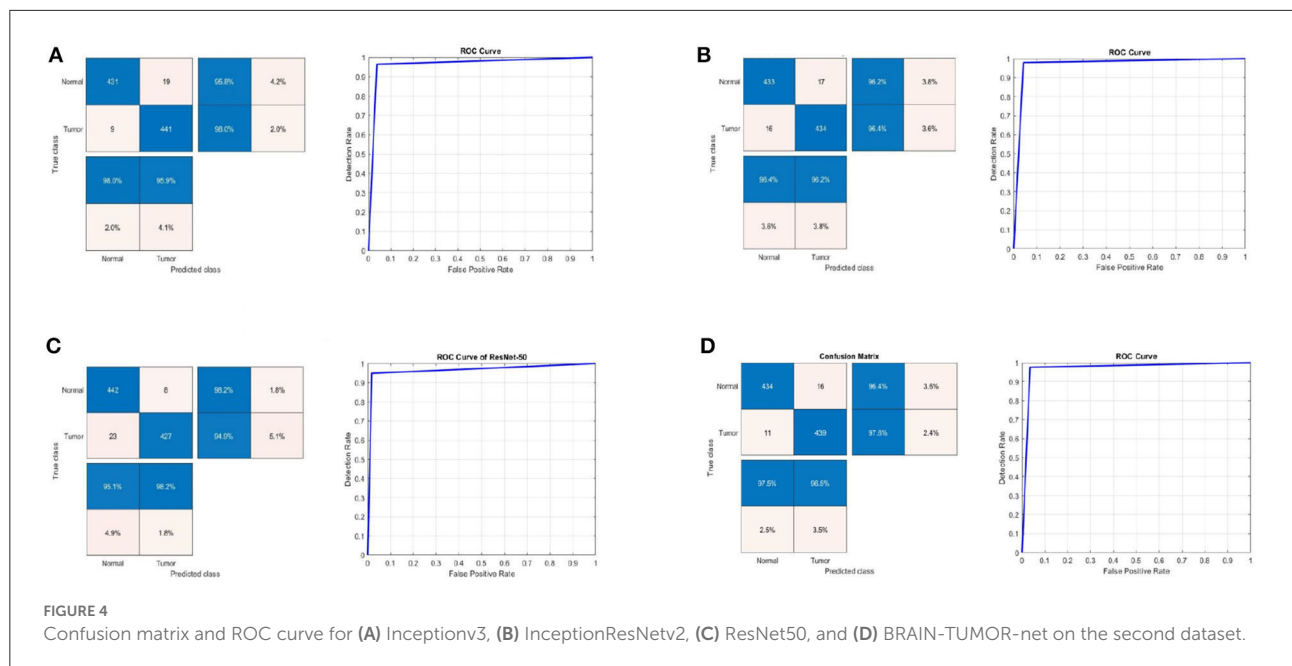


FIGURE 4 Confusion matrix and ROC curve for (A) Inceptionv3, (B) InceptionResNetv2, (C) ResNet50, and (D) BRAIN-TUMOR-net on the second dataset.

TABLE 3 Detection performance results for the proposed CNN models with a 75/25 training/testing ratio on the third dataset.

Models	Evaluation metric								
	ACC	SEN	SPE	PRECI	F1_score	MCC	Error	Kappa	FPR
InceptionResNetv2	0.9446	0.9988	0.8904	0.9011	0.9474	0.8944	0.0554	0.8892	0.1096
Inceptionv3	0.9767	0.9855	0.9679	0.9685	0.9769	0.9535	0.0233	0.9534	0.0321
ResNet50	0.9691	0.9970	0.9412	0.9443	0.9699	0.9397	0.0309	0.9382	0.0588
BRAIN-TUMOR-net	1.00	1.00	1.00	1.00	1.00	1.00	0.0	1.00	0.0

kappa, and error using three transfer-learning-based CNN models and a CNN model trained from scratch on the first MRI dataset. Figure 3 presents the receiver operating characteristic (ROC) curves and confusion matrices for Inceptionv3, InceptionResNetv2, ResNet50 and BRAIN-TUMOR-net models. It is clear that the ResNet50 model achieves the highest performance among other models on the first MRI dataset. It achieves a sensitivity of 93.48%, a specificity of 93.48%, an accuracy of 93.48%, a precision of 93.48%, an MCC of 86.96%, a false positive rate of 0.0652, an *F1_score* of 93.48%, a kappa of 86.96%, and an error of 0.0652.

4.2. Results on the second MRI dataset

Table 2 shows the detection performance results for different CNN models with a 75/25 training/testing ratio on the second dataset. Figure 4 presents the ROC curves and confusion matrices for Inceptionv3, InceptionResNetv2, ResNet50 and

BRAIN-TUMOR-net models, respectively. It is clear from the obtained results that the BRAIN-TUMOR-net model achieves the highest performance among the other proposed ones on the second MRI dataset. It achieves a sensitivity of 96.44%, a specificity of 97.56%, an accuracy 97%, a precision of 97.53%, an MCC of 94.01%, a false positive rate of 0.0244, an *F1_score* of 96.98%, a kappa of 94%, and an error of 0.0300.

4.3. Results on the third MRI dataset

Table 3 shows the detection performance results for different CNN models with a 75/25 training/testing ratio on the third dataset. Figure 5 presents the ROC curves and confusion matrices for Inceptionv3, InceptionResNetv2, ResNet50 and BRAIN-TUMOR-net models. It is clear from the obtained results that the BRAIN-TUMOR-net model achieves the highest performance among the other proposed ones on the third MRI dataset. It achieves a sensitivity

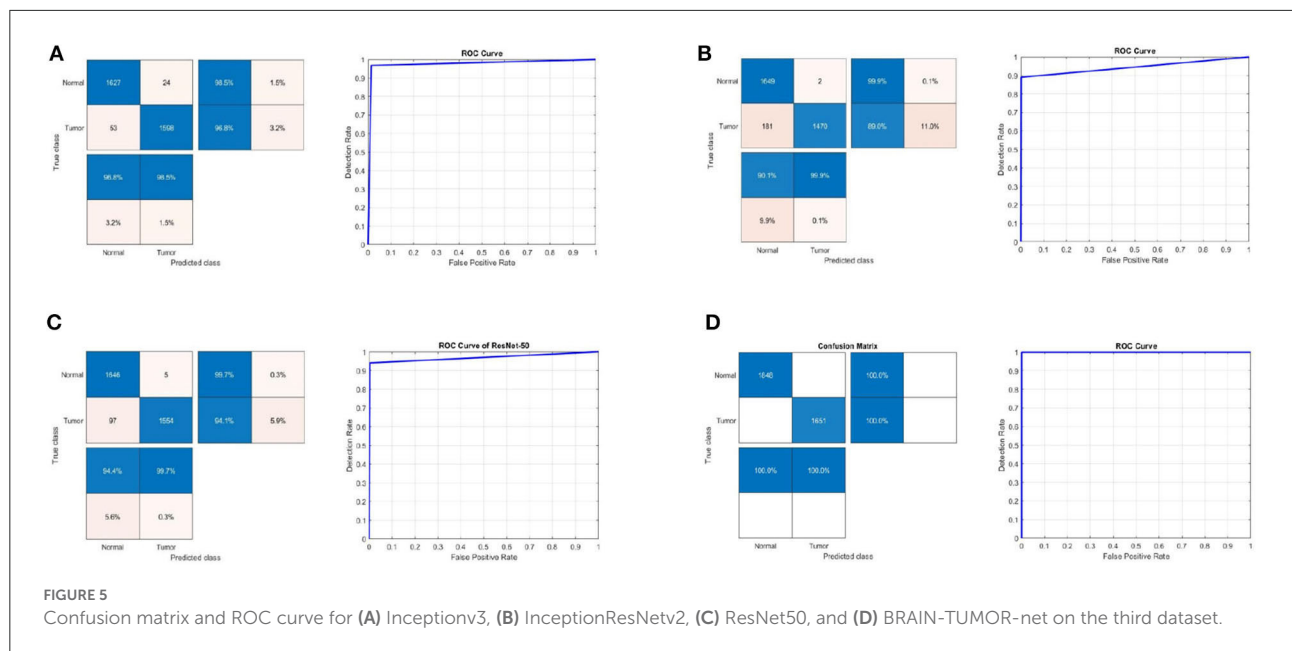


FIGURE 5 Confusion matrix and ROC curve for (A) Inceptionv3, (B) InceptionResNetv2, (C) ResNet50, and (D) BRAIN-TUMOR-net on the third dataset.

TABLE 4 Detection performance results for BRAIN-TUMOR-net based on stratified *k*-fold cross validation with a 75/25 training/testing ratio on the three datasets.

Dataset	Evaluation metrics								
	ACC (%)	SEN (%)	SPE (%)	PRECI (%)	F1_score (%)	MCC (%)	Error	Kappa (%)	FPR
First MRI dataset	89.03	92.26	85.81	86.67	89.38	78.23	0.1097	78.06	0.1419
Second MRI dataset	98.67	98.67	98.67	98.67	98.67	97.33	0.0133	97.33	0.0133
Third MRI dataset	99.49	99.38	99.60	99.60	99.49	98.98	0.0051	98.98	0.0040

of 100%, a specificity of 100%, an accuracy 100%, a precision of 100%, an MCC of 100%, a false positive rate of 0.0, an *F1_score* of 100%, a kappa of 100%, and an error of 0.0.

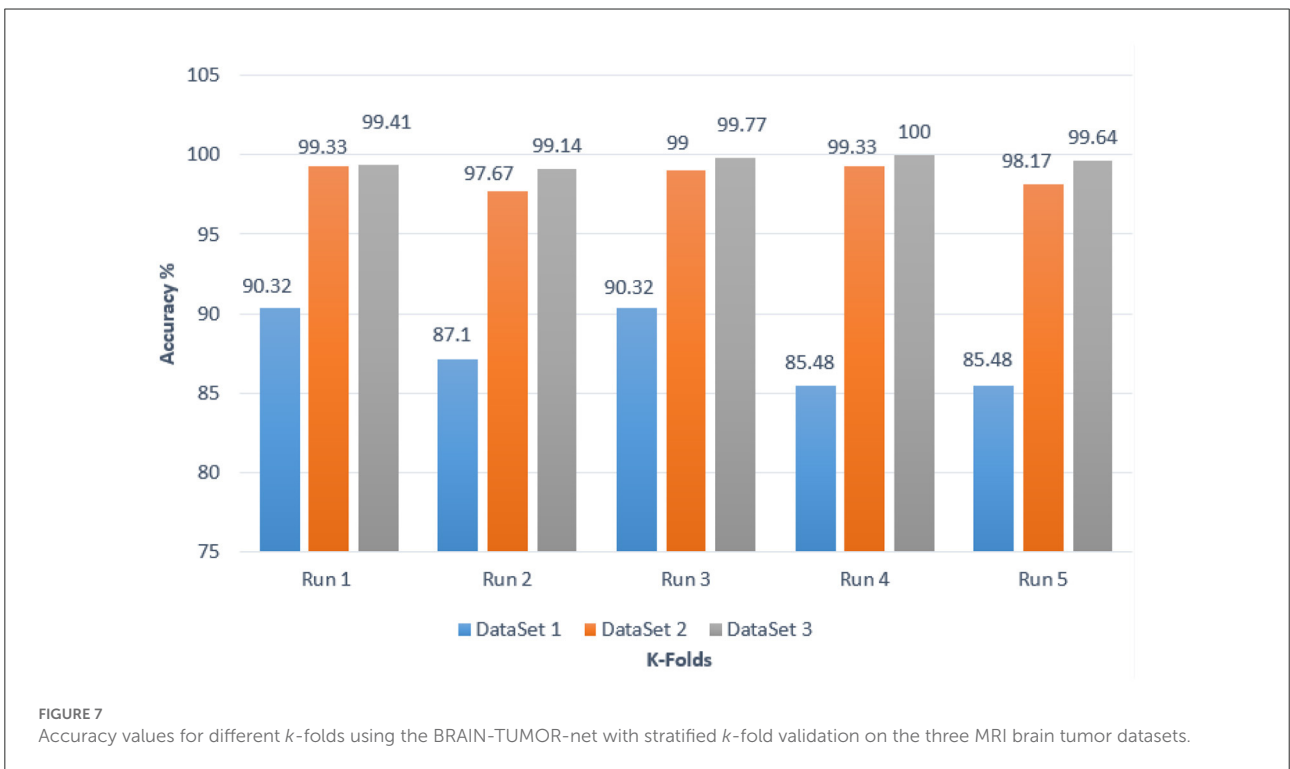
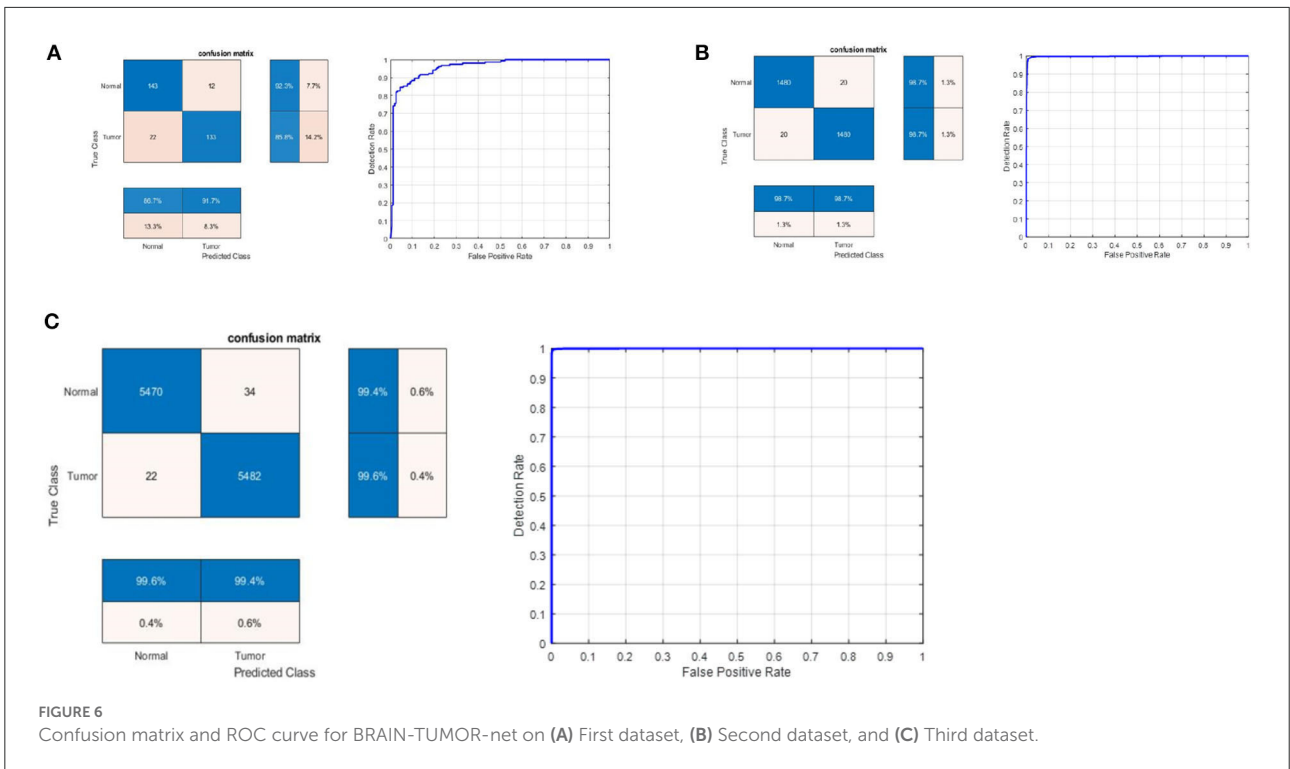
4.4. Results for BRAIN-TUMOR-net based on stratified *k*-fold validation

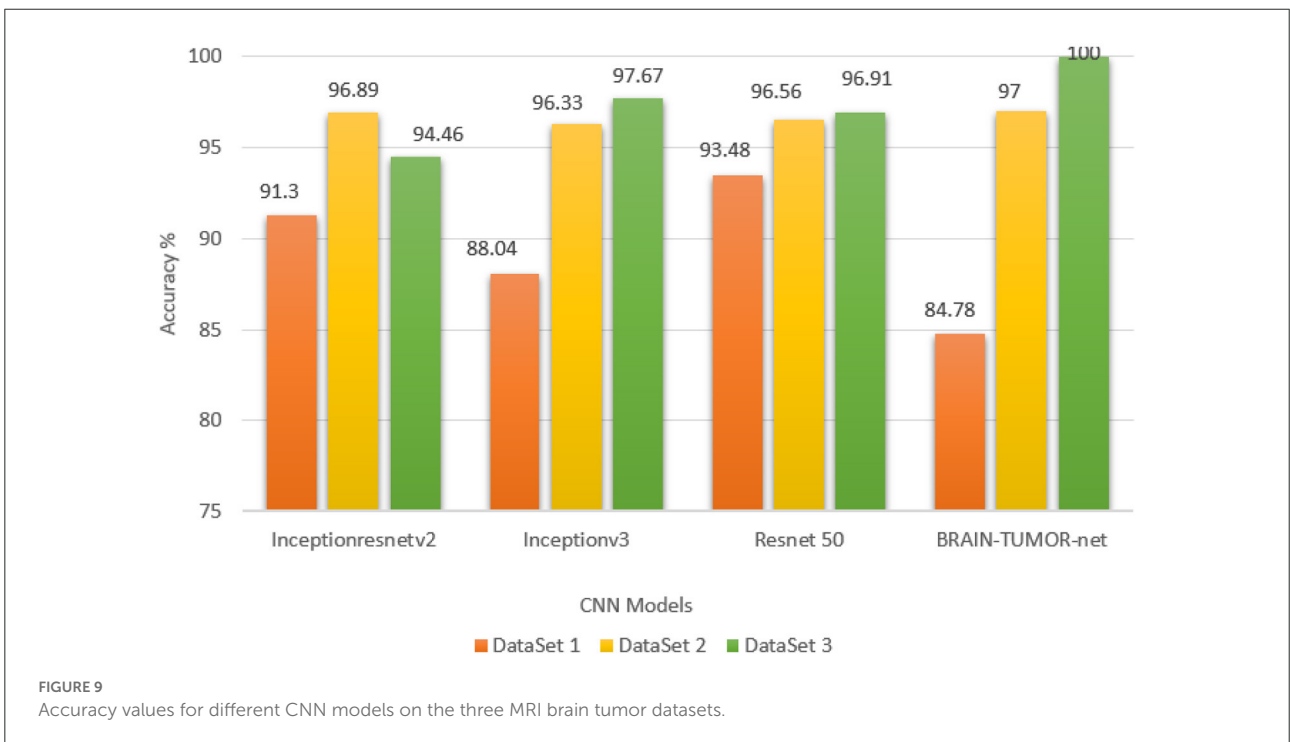
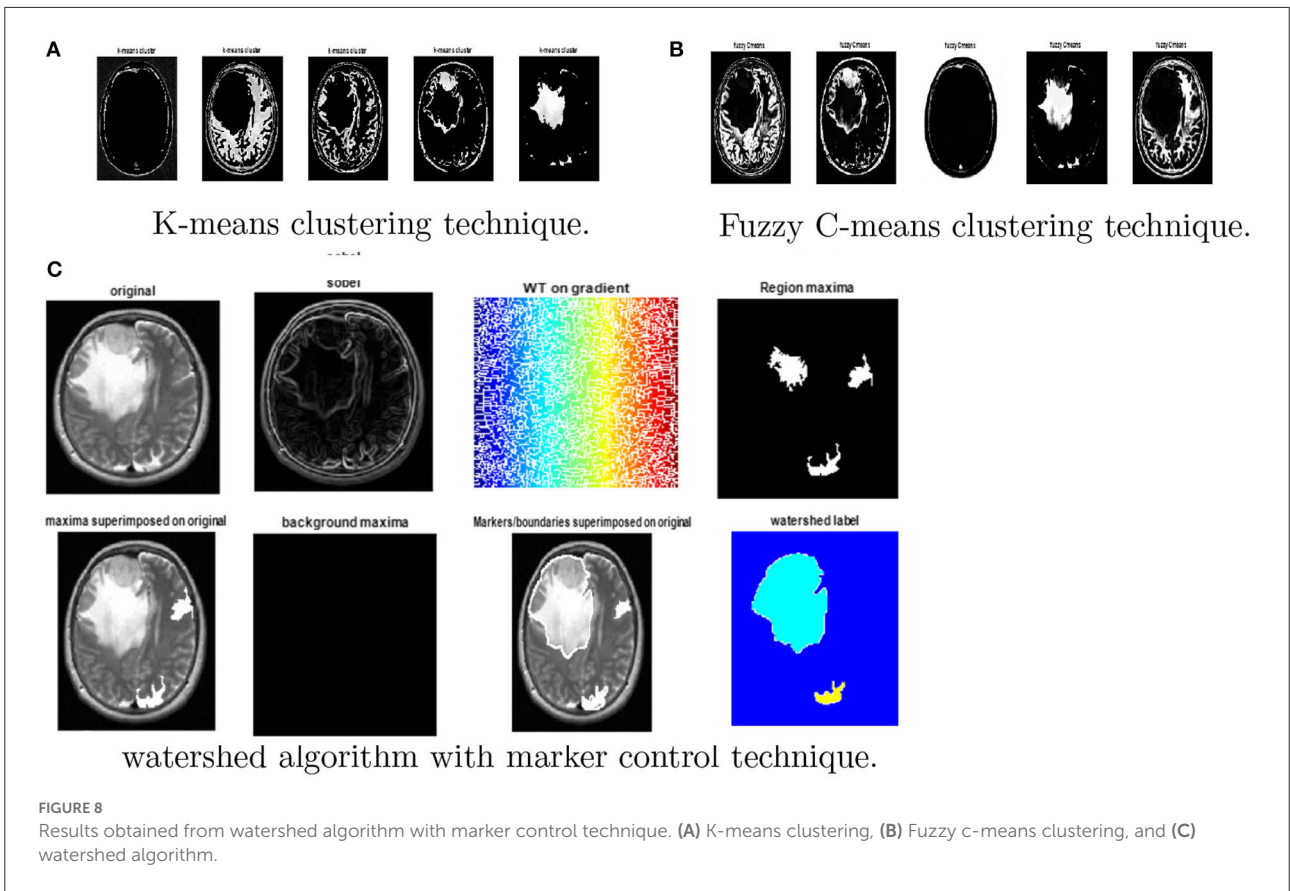
Stratified *k*-fold validation is combined with BRAIN-TUMOR-net model in order to get a more stable error. Table 4 presents the detection performance results from BRAIN-TUMOR-net using a stratified *k*-fold validation model with a 75/25 training/testing ratio on the three MRI datasets. Figure 6 presents the ROC curves and confusion matrices for the first, second and third datasets. Figure 7 provides the value of accuracy for each *k*.

4.5. MRI brain tumor segmentation results

On magnetic resonance brain images, an image segmentation method is used to separate similar sections of the image based on the gray level values of the pixels. The primary goal of segmenting magnetic resonance brain images is to aid in tumor detection. Edge-based segmentation (Krisch and Sobel), threshold-based segmentation (Otsu), and clustering algorithms, namely k-means, adaptive k-means, fuzzy c-means, and marker-controlled watershed, were used as segmentation techniques. Segmentation process consists of seven steps to generate the segmented image. The normal shrink image is generated from the normal shrink denoising algorithm. Edge preservation is achieved by applying a bilateral filter on the denoised output.

Various clustering techniques such as k-means clustering, adaptive clustering, and fuzzy c-means clustering have





been implemented. Figure 8 depicts the outcomes for *k*-means clustering, fuzzy *c*-means clustering, and the results obtained from the watershed algorithm with marker control.

Finally, the magnitude of the gradient is determined and employed as a segmentation function. Then, the watershed transform is applied on the gradient magnitude. The original image is then changed into another image by employing morphological operators to iterate with other images of specified shape and size.

5. Discussion and comparison with the-state-of-the-art methods

As the dataset size grows, the BRAIN-TUMOR-net trained from scratch outperforms the transfer-learning-based techniques. However, the results show that even when using a small dataset, transfer learning produces satisfactory outcomes. State-of-the-art models were trained on 25 million images. Their convolution layer filters were chosen, because they are effective in novel applications like brain tumor detection. Furthermore, the accuracy of the classification application is influenced by the depth of the CNN models.

The stratified *k*-fold validation procedure with 5 folds is employed to get a more steady error rate. This is attributed to the stratified *k*-fold cross-validation capacity to cope with imbalanced data. It keeps the same class ratio as that of the original dataset throughout the *k* folds. The accuracy results for different CNN models on the three MRI brain tumor datasets are shown in Figure 9. The results clearly vary based on the depth of the CNN model, the classification complexity, and the amount of data.

The computation time is the most important metric for comparing various techniques. It is obvious from Table 5 that the ResNet50 model yields run times of 4.68, 9.44, and 24.67 s on the first, second, and third datasets, respectively, which are the shortest times. The Inceptionv3 model was reported to give the second-best runtimes of 7.67, 23.91, and 71.835 s on the first, second, and third datasets, respectively. However, the CNN model trained from scratch using *k*-fold validation has the longest runtimes of 376, 3418, and 16284.818 s on the first, second, and third datasets, respectively. It obtained a level of accuracy of 100%. In the future work, numerous solutions such as image downsizing, adjusting the number of max-pooling layers, and dropout will be investigated to minimize the computation time.

TABLE 5 Computational times of the examined approaches.

MRI dataset	CNN model	Elapsed time (s)
Dataset 1	InceptionResNetv2	31.66
	Inceptionv3	7.67
	ResNet50	4.68
	BRAIN-TUMOR-net	53
	BRAIN-TUMOR-net with K-fold	376
Dataset 2	InceptionResNetv2	68.95
	Inceptionv3	23.91
	ResNet50	9.44
	BRAIN-TUMOR-net	1,069.2
	BRAIN-TUMOR-net with <i>k</i> -fold cross validation	3418
Dataset 3	InceptionResNetv2	210.0371
	Inceptionv3	71.83508
	ResNet50	24.6754
	BRAIN-TUMOR-net	11,702.6
	BRAIN-TUMOR-net with <i>k</i> -fold cross validation	16284.818355

The proposed approach yields an accuracy level of 100%, which is greater than the levels of the traditional approaches shown in Table 6. These findings support the CNN model ability to execute the essential classification task after being trained from scratch.

6. Conclusions

The CNN is regarded as one of the most effective tools for classifying image datasets. It produces the forecast by reducing the image into features without losing the necessary information to make the prediction, correctly. In this paper, three different deep learning models for brain tumor classification have been introduced. Transfer-learning-based models, as well as a CNN model, BRAIN-TUMOR-net, and a model trained from scratch, have been introduced. Three publicly available MRI datasets have been used to test the proposed models. The results show that the BRAIN-TUMOR-net achieves the highest accuracy among the other models as the dataset size increases. It achieves a 100% accuracy on the third dataset, while it achieves 97% and 84.78% accuracy levels on the second and first MRI datasets, respectively. When compared to existing pre-trained models, the proposed model needs extremely less processing power and achieves far higher accuracy outcomes. In future research, optimization techniques can be applied so

TABLE 6 Comparison of the proposed work with state-of-the-art models.

References	Method	Accuracy	
Sindhumol et al. (7)	Spectral angle-based feature extraction method and Spectral Clustering Independent Component Analysis (SC-ICA)	98% for SVM and 96.1% for reproduced lesion	
Sumitra and Saxena (14)	PCA, Back-propagation neural network	73%	
Jafari and Shafaghi (15)	Genetic Algorithm (GA) and Support Vector Machine (SVM),	83.22%	
Jayachandran and Dhanasekaran (16)	Statistical, SVM classifier and PCA	95.80%	
Xing et al. (18)	The HySIME algorithm initial filtering by deep CNN followed by iterative region merging segmentation by selective sparse shape model.	85%	
Zaw et al. (20)	Morphological operation, pixel subtraction, and maximum entropy threshold segmentation with Naive Bayes classifier.	94 %	
Narayana and Reddy (19)	Median filter GA segmentation with SVM classifier.	91.23%	
Minz and Mahobiya (22)	GLCM (Gray Level Co-occurrence Matrix) for classification boosting.	89.90% & 74.00%	
Raju et al. (23)	Bayesian fuzzy clustering segmentation with HSC-based multi SVNN classification method.	93%	
Sert et al. (24)	Single image super-resolution for image enhancement and segmentation with maximum fuzzy entropy (MFE) and SVM classifier.	95%	
Deepak and Ameer (25)	Min-max normalization, with SVM and KNN classifier.	97.8% & 98%	
Hemanth et al. (8)	Average filter, and pixel subtraction with CNN-linkNet classifier.	91%	
Mallick et al. (9)	DICOM image processing, and DWT-DNN features with MLP classifier.	96%	
Selvapandian and Manivannan (17)	NSCT image enhancement, and GLCM texture features with ANFIS classifier.	98.5%	
Nalepa et al. (11)	Sharpening and smoothing filters, threshold segmentation, and SGLD features with ANN classifier.	99%	
Anaraki et al. (10)	Image rescaling, and data augmentation with CNN classifier.	96%	
Amin et al. (12)	Skull stripping-BSE Gaussian filtering, <i>k</i> -Means clustering with SVM classifier.	98%	
Gupta and Khanna (13)	Image enhancement-DSR-AD, and Otsu segmentation with SVM classifier.	98%	
Proposed work			
Dataset 1	InceptionResNetv2	InceptionResNetv2	91.30%
	Inceptionv3	Inceptionv3	88.04%
	ResNet50	ResNet50	93.48%
	BRAIN-TUMOR-net	BRAIN-TUMOR-net	84.78%
	<i>k</i> -fold model	<i>k</i> -fold validation model	89.03%
Dataset 2	InceptionResNetv2	InceptionResNetv2	96.89%
	Inceptionv3	Inceptionv3	96.33%
	ResNet50	ResNet50	96.56%
	BRAIN-TUMOR-net	BRAIN-TUMOR-net	97%
	<i>k</i> -fold Model	<i>k</i> -fold validation model	98.67%

(Continued)

TABLE 6 (Continued)

References	Method	Accuracy
InceptionResNetv2	InceptionResNetv2	94.46%
Inceptionv3	Inceptionv3	97.67%
Dataset 3 ResNet50	ResNet50	96.91%
BRAIN-TUMOR-net	BRAIN-TUMOR-net	100%
k-fold Model	k-fold validation model	99.49%

as to decide the number of layers and filters that can be used in the model.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.kaggle.com/ahmedhamada0/brain-tumor-detection?select=yes>; <https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection>; <https://www.kaggle.com/leaderandpiller/brain-tumor-segmentation>.

Author contributions

FA: study conception and design. MS and HE: data collection. MS, KA, and FT: analysis and interpretation of results. MH, HE, and MS: draft manuscript preparation. All authors reviewed the results and approved the final version of the manuscript.

References

- Healthline. *Brain-tumor*. Available online at: <https://www.healthline.com/health/brain-tumor>
- Cancer Net. *Cancer.net. Cancer-Types.brain-tumor.symptoms-and-signs*. Available online at: <https://www.healthline.com/health/brain-tumor>
- Maravilla KR, Sory WC. Magnetic resonance imaging of brain tumors. In: *Seminars in Neurology*. Vol. 6. copyright 1986 by Thieme Medical Publishers, Inc., Texas (1986). p. 33–42.
- Navoneel K. *Brain MRI Images for Brain Tumor Detection*. Available online at: <https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection>
- Ahmedhamada. *Brain-Tumor-Detection*. Available online at: <https://www.kaggle.com/ahmedhamada0/brain-tumor-detection?select=yes>
- Leaderandpiller. *Brain-Tumor-Segmentation*. Available online at: <https://www.kaggle.com/leaderandpiller/brain-tumor-segmentation>
- Sindhulol S, Kumar A, Balakrishnan K. Spectral clustering independent component analysis for tissue classification from brain MRI. *Biomed Signal Process Control*. (2013) 8:667–74. doi: 10.1016/j.bspc.2013.06.007
- Hemanth G, Janardhan M, Sujihelen L. Design and implementing brain tumor detection using machine learning approach. In: *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*. Tirunelveli: IEEE (2019). p. 1289–94.
- Mallick PK, Ryu SH, Satapathy SK, Mishra S, Nguyen GN, Tiwari P. Brain MRI image classification for cancer detection using deep wavelet autoencoder-based deep neural network. *IEEE Access*. (2019) 7:46278–87. doi: 10.1109/ACCESS.2019.2902252
- Anaraki AK, Ayati M, Kazemi F. Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms. *Biocybern Biomed Eng*. (2019) 39:63–74. doi: 10.1016/j.bbe.2018.10.004
- Nalepa J, Lorenzo PR, Marcinkiewicz M, Bobek-Billewicz B, Wawrzyniak P, Walczak M, et al. Fully-automated deep learning-powered system for DCE-MRI analysis of brain tumors. *Artif Intell Med*. (2020) 102:101769. doi: 10.1016/j.artmed.2019.101769
- Amin J, Sharif M, Yasmin M, Fernandes SL. A distinctive approach in brain tumor detection and classification using MRI. *Pattern Recogn Lett*. (2020) 139:118–27. doi: 10.1016/j.patrec.2017.10.036
- Gupta N, Khanna P. A non-invasive and adaptive CAD system to detect brain tumor from T2-weighted MRIs using customized Otsus thresholding with prominent features and supervised learning. *Signal Process Image Commun*. (2017) 59:18–26. doi: 10.1016/j.image.2017.05.013
- Sumitra N, Saxena RK. Brain tumor classification using back propagation neural network. *Int J Image Graphics Signal Process*. (2013) 5:45. doi: 10.5815/ijgisp.2013.02.07

Acknowledgments

The author would like to thank the Deanship of Scientific Research at Shaqra University for supporting this work.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

15. Jafari M, Shafaghi R. A hybrid approach for automatic tumor detection of brain MRI using support vector machine and genetic algorithm. *Global J Sci Eng Technol.* (2012) 3:1–8. doi: 10.5120/18036-6883
16. Jayachandran A, Dhanasekaran R. Brain tumor detection and classification of MR images using texture features and fuzzy SVM classifier. *Res J Appl Sci Eng Technol.* (2013) 6:2264–9. doi: 10.19026/rjaset.6.3857
17. Selvapandian A, Manivannan K. Fusion based glioma brain tumor detection and segmentation using ANFIS classification. *Comput Methods Programs Biomed.* (2018) 166:33–8. doi: 10.1016/j.cmpb.2018.09.006
18. Xing F, Xie Y, Yang L. An automatic learning-based framework for robust nucleus segmentation. *IEEE Trans Med Imaging.* (2015) 35:550–66. doi: 10.1109/TMI.2015.2481436
19. Narayana TL, Reddy TS. An efficient optimization technique to detect brain tumor from MRI images. In: *2018 International Conference on Smart Systems and Inventive Technology (ICSSIT)*. Tirunelveli: IEEE (2018). p. 168–71.
20. Zaw HT, Maneerat N, Win KY. Brain tumor detection based on Naïve Bayes Classification. In: *2019 5th International Conference on Engineering, Applied Sciences and Technology (ICEAST)*. Luang Prabang: IEEE (2019). p. 1–4.
21. Veeramuthu A, Meenakshi S, Mathivanan G, Kotecha K, Saini JR, Vijayakumar V, et al. MRI brain tumor image classification using a combined feature and image-based classifier. *Front Psychol.* (2022) 13:848784. doi: 10.3389/fpsyg.2022.848784
22. Minz A, Mahobiya C. MR image classification using adaboost for brain tumor type. In: *2017 IEEE 7th International Advance Computing Conference (IACC)*. Hyderabad: IEEE (2017). p. 701–5.
23. Raju AR, Suresh P, Rao RR. Bayesian HCS-based multi-SVNN: A classification approach for brain tumor segmentation and classification using Bayesian fuzzy clustering. *Biocybern Biomed Eng.* (2018) 38:646–60. doi: 10.1016/j.bbe.2018.05.001
24. Sert E, Özyurt F, Doğanekin A. A new approach for brain tumor diagnosis system: single image super resolution based maximum fuzzy entropy segmentation and convolutional neural network. *Med Hypotheses.* (2019) 133:109413. doi: 10.1016/j.mehy.2019.109413
25. Deepak S, Ameer P. Brain tumor classification using deep CNN features via transfer learning. *Comput Biol Med.* (2019) 111:103345. doi: 10.1016/j.compbiomed.2019.10.3345
26. Wang C, Chen D, Hao L, Liu X, Zeng Y, Chen J, et al. Pulmonary image classification based on inception-v3 transfer learning model. *IEEE Access.* (2019) 7:146533–41. doi: 10.1109/ACCESS.2019.2946000
27. Upadhyay M, Rawat J, Maji S. Skin cancer image classification using deep neural network models. In: *Evolution in Computational Intelligence: Proceedings of the 9th International Conference on Frontiers in Intelligent Computing: Theory and Applications (FICTA 2021)*. Vol. 267. Singapore: Springer Nature (2022). p. 451.
28. Szegedy C, Ioffe S, Vanhoucke V, Alemi A. Inception-v4, inception-resnet and the impact of residual connections on learning. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 31. San Francisco, CA (2017).
29. Haji SH, Abdulazeez AM. Comparison of optimization techniques based on gradient descent algorithm: a review. *PalArchs J Archaeol Egypt/Egyptol.* (2021) 18:2715–43.
30. Gazda M, Hireš M, Drotár P. Multiple-fine-tuned convolutional neural networks for Parkinsons disease diagnosis from offline handwriting. *IEEE Trans Syst Man Cybern Syst.* (2021) 52:78–89. doi: 10.1109/TSMC.2020.3048892
31. Yildirim O, Talo M, Ay B, Baloglu UB, Aydin G, Acharya UR. Automated detection of diabetic subject using pre-trained 2D-CNN models with frequency spectrum images extracted from heart rate signals. *Comput Biol Med.* (2019) 113:103387. doi: 10.1016/j.compbiomed.2019.103387
32. Kassani SH, Kassani PH. A comparative study of deep learning architectures on melanoma detection. *Tissue Cell.* (2019) 58:76–83. doi: 10.1016/j.tice.2019.04.009
33. Ribli D, Horváth A, Unger Z, Pollner P, Csabai I. Detecting and classifying lesions in mammograms with deep learning. *Sci Rep.* (2018) 8:1–7. doi: 10.1038/s41598-018-22437-z
34. Shoaib MR, Elshamy MR, Taha TE, El-Fishawy AS, Abd El-Samie FE. Efficient brain tumor detection based on deep learning models. In: *Journal of Physics: Conference Series*. Vol. 2128. Bogota: IOP Publishing (2021). p. 012012.
35. Emara HM, Shoaib MR, Elwekeil M, El-Shafai W, Taha TE, El-Fishawy AS, et al. Deep convolutional neural networks for COVID-19 automatic diagnosis. *Microsc Res Techn.* (2021) 84:2504–16. doi: 10.1002/jemt.23713
36. Khalil H, El-Hag N, Sedik A, El-Shafie W, Mohamed AEN, Khalaf A, et al. Classification of diabetic retinopathy types based on convolution neural network (CNN). *Menoufia J Electron Eng Res.* (2019) 28:126–53. doi: 10.21608/mjeer.2019.76962
37. Kim P. Matlab deep learning. In: *With Machine Learning, Neural Networks and Artificial Intelligence*. California, CA (2017).
38. Shoaib MR, Emara HM, Elwekeil M, El-Shafai W, Taha TE, El-Fishawy AS, et al. Hybrid classification structures for automatic COVID-19 detection. *J Ambient Intell Humaniz Comput.* (2022) 13:4477–92. doi: 10.1007/s12652-021-03686-9
39. Hill E, Binkley D, Lawrie D, Pollock L, Vijay-Shanker K. An empirical study of identifier splitting techniques. *Empir Softw Eng.* (2014) 19:1754–80. doi: 10.1007/s10664-013-9261-0
40. Phan H, He Y, Savvides M, Shen Z. Mobinet: a mobile binary network for image classification. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. Snowmass, CO: IEEE (2020). p. 3453–62.
41. Wong TT, Yeh PY. Reliable accuracy estimates from k-fold cross validation. *IEEE Trans Knowl Data Eng.* (2019) 32:1586–94. doi: 10.1109/TKDE.2019.2912815
42. Tilekar P, Singh P, Aherwadi N, Pande S, Khamparia A. Breast cancer detection using image processing and CNN algorithm with K-Fold cross-validation. In: *Proceedings of Data Analytics and Management*. Singapore: Springer (2022). p. 481–90.
43. Anguita D, Ghelardoni L, Ghio A, Oneto L, Ridella S. The Kin K-fold cross validation. In: *20th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)*. Bruges: i6doc.com Publ (2012). p. 441–6.
44. Jiang G, Wang W. Error estimation based on variance analysis of k-fold cross-validation. *Pattern Recogn.* (2017) 69:94–106. doi: 10.1016/j.patcog.2017.03.025
45. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging.* (2015) 15:29. doi: 10.1186/s12880-015-0068-x
46. Jensen JR. *Introductory Digital Image Processing: A Remote Sensing Perspective*. 2nd ed. Upper Saddle River, NJ: Prentice-Hall Inc. (1996).