Al-Mukhtar Journal of Engineering Research 09 (1): 28-39, 2025
Doi: https://doi.org/10.54172/yzehb352

Research Article

Open Access

Transfer Learning in Convolution Neural Network for Brain Tumor Detection Using a Small Training Dataset

Khamis Faraj Alarabi Aljribi[1*], Mohammed Elhadi Faraj[2]

[2] Department of Computer, Faculty of Electronic Technology, Baniwalid University, Libya.

***Corresponding author:** kf.aljribi@gmail.com , Department of Computer, Faculty of Electronic Technology, Baniwalid University, Libya.

Received:
18 November 2024

Accepted:
23 March 2025

Publish online:
31 December 2025

Abstract

Tumors are a significant risk in today's medical field, requiring fast and reliable automated techniques for detection, particularly for brain tumors. Accurate detection is crucial for effective treatment and saving lives. Various image processing techniques aid doctors in providing appropriate treatment. Manual or human-based identification of brain tumors using MRI images is time-consuming and prone to inaccuracies, especially for an experienced person. Deep learning algorithms have introduced effective solutions for brain tumor detection. The one constraint is that the algorithms need to train on a huge amount of data for reliable performance. This research aims to investigate the effectiveness of transfer learning for brain tumor detection using a small training dataset regardless of the tumor type. Therefore, we have used three models of Convolutional Neural Networks (CNN): traditional model, enhanced model, and transfer learning-based model. The average results have shown that the transfer learning-based model has better performance on the small training dataset than traditional and enhanced models, with a classification accuracy reached 92%.

Keywords: transfer learning, VGG16, CNN, brain tumor, MRI, deep learning

INTRODUCTION

Tumors are a significant risk in today's medical field, which affects a lot of people worldwide. They require fast and reliable automated techniques for detection, particularly for brain tumors. Brain tumors can be detected using imaging techniques such as diffusion tensor imaging (DTI) and magnetic resonance imaging (MRI). Early diagnosis and classification of brain tumors play a vital role in effective treatment and improved outcomes. However, the evaluation and stratification of tumors can be complex and time-consuming for medical professionals, which leads to the use of information technology techniques for implementing automatic detection systems. Artificial intelligence applications, especially machine learning, have been increasingly used for medical applications to develop systems that can learn independently, without human intervention, for a variety of tumor detection tasks.

Machine Learning (ML) is a method that utilizes algorithms and mathematical structures to perform tasks by recognizing and leveraging patterns, without requiring explicit instructions. Deep Learning (DL) is a branch of machine learning that focuses on hierarchical feature learning and data representation. DL algorithms use multiple layers of nonlinear transformations to extract features from

data. Convolutional Neural Network (CNN), as an instance, is a popular type of DL model, especially in visual imaging research. It requires minimal preprocessing and is inspired by the neural activities of the human brain. LeNet, the first deep CNN, was developed in the late 20th century for text recognition. Later, the VGG16 model gained popularity for image classification and achieved remarkable performance compared to other network architectures, leading to significant breakthroughs in the deep learning community focused on CNNs (Deng & Yu, 2014**)**.
Both machine learning and deep learning algorithms need big data for training the model enough to be able to classify the data. In some cases, data are not available in amounts sufficient for training new models. Accordingly, the system accuracy will be significantly low. To solve such a problem, researchers have suggested the investigation of using enhanced deep learning and transfer learning-based algorithms.

Transfer learning

Knowledge gained from solving one type of problem can be applied to similar challenges. Pre-trained networks, which have learned a rich set of features during solving the initial problem, can be utilized to learn and solve new, similar problems. By fine-tuning the pre-trained network through transfer learning, it can be adapted to new applications. In brain tumor classification using CNN, the process involves training and testing stages, with pre-processing, feature extraction, and classification using a loss function. The VGG 16 model is a commonly used pre-trained convolutional neural network for brain tumor classification. Transfer learning is employed to expedite training, with adjustments made to the first and last three layers of the network to adapt it to the classification goal, and the output of the fully connected layer indicates the presence or absence of a tumor.(Bairagi et al., 2023)

In**(Sawant et al., n.d.)** The authors highlight the time-consuming nature of traditional MRI analysis for brain cancer detection by proposing the use of a convolutional neural network (CNN) implemented in TensorFlow. They report using a 5-layer CNN with a dataset of 1800 MRI images, 900 cancerous and 900 non-cancerous. The findings show high training and validation accuracies of 99% and 98.6% respectively, over 35 epochs, demonstrating the effectiveness of the approach. The system is positioned as a potential tool for surgeons and radiologists, offering efficient and accurate brain tumor detection. The authors acknowledge opportunities for further development, such as using GPU-accelerated TensorFlow for faster model creation and integrating more advanced MRI image processing techniques.

In **(Rai & Chatterjee, 2020)** Hari Mohan Rai and Kalyan Chatterjee introduce Lu-Net, a low-complexity and efficient deep neural CNN model for detecting brain abnormalities in MR images. The authors preprocess, resize, crop, and augment the images before training Lu-Net. The model is evaluated using various statistical assessment metrics and compared with other CNN models like LeNet and VGG-16. Lu-Net achieves an impressive overall accuracy of 98%, surpassing the performance of other models. This research highlights the potential of deep learning in biomedical imaging, particularly for early diagnosis and treatment of brain tumors and other abnormalities. The high accuracy and performance metrics of Lu-Net make it a promising tool for enhancing medical diagnostics and patient outcomes.

In **(Badža & Barjaktarović, 2020)** authors introduce a novel CNN architecture for brain tumor classification using MRI images. They utilize an augmented dataset of 3064 MRI images from three tumor types and propose a CNN architecture that takes whole images as input. The network is evaluated using subject-wise 10-fold cross-validation and achieves an impressive accuracy of

97.28%, surpassing state-of-the-art methods. The paper includes a thorough discussion of various network types, pre-trained models, and combinations with neural networks for feature extraction and classification. The results demonstrate the effectiveness of the proposed CNN architecture for accurate brain tumor classification.

In (Abiwinanda et al., 2019) The authors of the study propose the use of Convolutional Neural Networks (CNN) for the automatic classification of the most common types of brain tumors. They train the CNN using a dataset of T-1 weighted CE-MRI brain tumor images, with balanced samples for each tumor class. The CNN achieves a high training accuracy of 98.51% and validation accuracy of 84.19%. The study highlights the potential of CNN as a supportive tool for medical doctors in classifying brain tumors, with comparable accuracy to conventional algorithms. The CNN-based approach shows improved performance compared to region-based pre-processing algorithms typically used in this task.

In(Das et al., 2019) The study proposes a CNN model for classifying brain cancers into three major types: glioma, meningioma, and pituitary. The CNN is trained using an Adam optimizer with a batch size of 256 and 100 epochs after preprocessing the image data. In testing on real-world data, the model achieves an accuracy of 94.39% and a loss of 28.16%. Performance evaluation includes measures such as accuracy, recall, F1-score, and support, with precision values of 88%, 94%, and 98% for glioma, meningioma, and pituitary classes, respectively. This highlights the effectiveness of the proposed CNN model for brain cancer classification.

MATERIALS AND METHODS

This research has been designed to investigate three kinds of CNN models to find out the best algorithm type for classifying data when the dataset is relatively small, especially in the medical field. Therefore, we have chosen a set of MRI images of brain tumors as a dataset to build and train three models of CNN; the first is a traditional CNN model, the second is an enhanced CNN model, and the third is a transfer learning-based model of CNN, as shown by **Figure (1**.

The system consists of five modules: Dataset, Pre-processing, CNN Model Building and Training, using traditional CNN model, enhanced CNN model, and transfer learning-based model (VGG16). To ensure the best classification and avoiding the biasness in the dataset, models have been trained on the dataset following the k-fold cross-validation with k=5. Furthermore, data have been exposed to an augmentation process to increase the dataset instances by changing images orientation or their brightness.

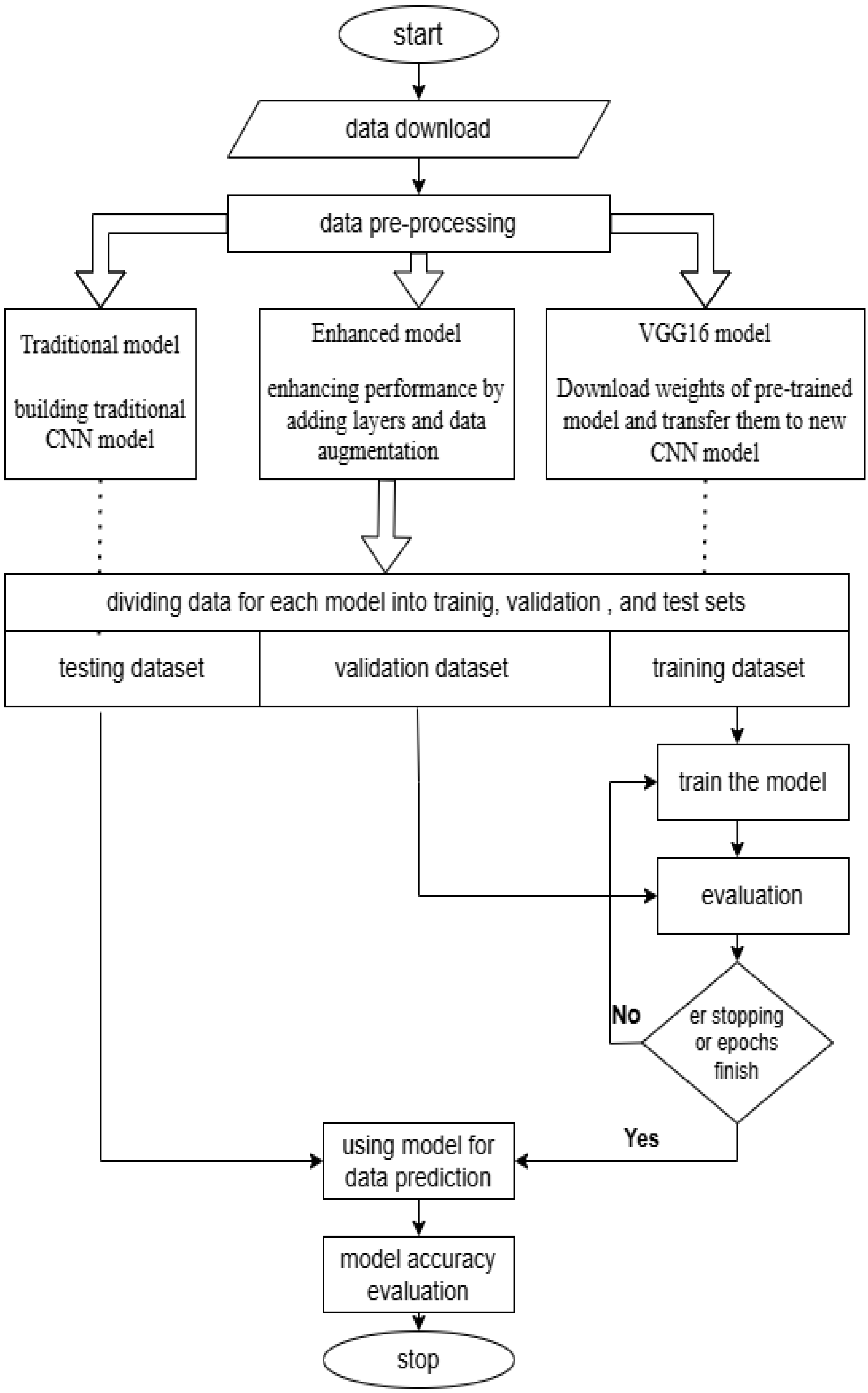

Figure (1): Flowchart of the research processes

DATA SET

Due to the lack of brain tumor MRI images, the dataset used in this study was downloaded from the common "Kaggle" website, which contains brain MRI images categorized as normal (without tumors) or abnormal (with tumors), regardless of the type of tumor. The images are in .jpg format, and there were 253 images in total. Out of these, 98 images show tumor-free brains, and 155 images show brains with tumors. 203 images have been used for models training and 50(50% for normal brain) images for testing. Figure (2 represents a sample of normal brain MRI images without a tumor, while Figure (3 Represents a sample of abnormal brain MRI images with a tumor.

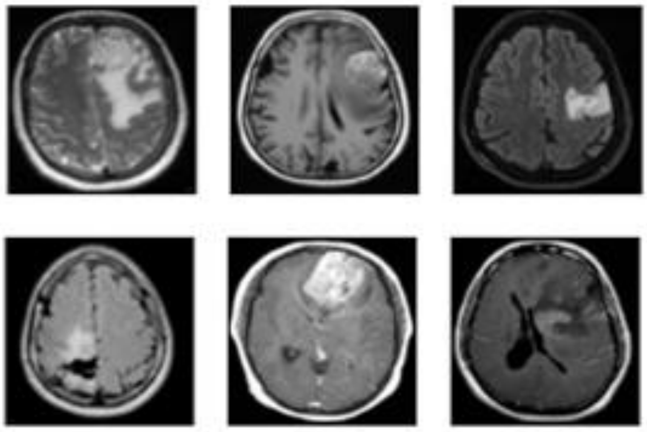

Figure (2): samples of MRI brain images with a tumor

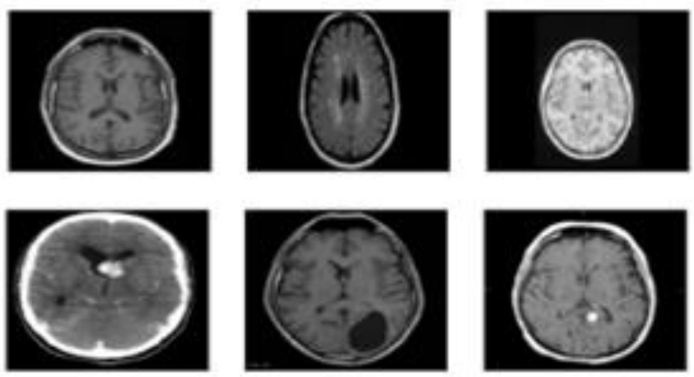

Figure (3): samples of MRI normal brain images

Data pre-processing

The pre-processing stage aims to prepare images to be processed by CNN models for classifying images into two categories: "yes" or "y", when the system decides the image contains a tumor, and "no" or "N", when the system decides the image does not contain a tumor. As a part of the pre-processing step, we have removed extraneous edges; the areas within images but without features, then resized the images to (224x224) for training with VGG16 and to (200x200) for training with CNN and Enhanced CNN. Additionally, as the data sample is small, the data augmentation techniques were utilized to increase the amount of training data, as it is a usual process followed, especially in cases where the available data is limited, as in medical image datasets. Finally, the image dataset is divided into two parts: one represents 70% for training and the other represents 30% for validation.

Convolutional Neural Network (CNN)

The potential of AI to bridge the gap between human and machine capabilities has expanded significantly. Convolutional Neural Networks (CNNs) are deep learning methods that are effectively used to distinguish different aspects within images. CNN requires less pre-processing compared to other classification approaches and can learn complex patterns with sufficient training **(Tamm et al., 2020)**. CNN has been widely used for object detection and has become a common method in computer vision applications.

CNN has a hierarchical architecture consisting of input, hidden units, convolutional features, batch normalization, and convolution layers. CNNs can vary in the number of layers, size, and activation methods used, which are determined through experimental identification and empirical support. Neurons respond to changes in a specific portion of the visual field called the Receptive Field, and if these fields intersect, they cover the entire viewable region. **Figure (4** Explains the general architecture of CNN.

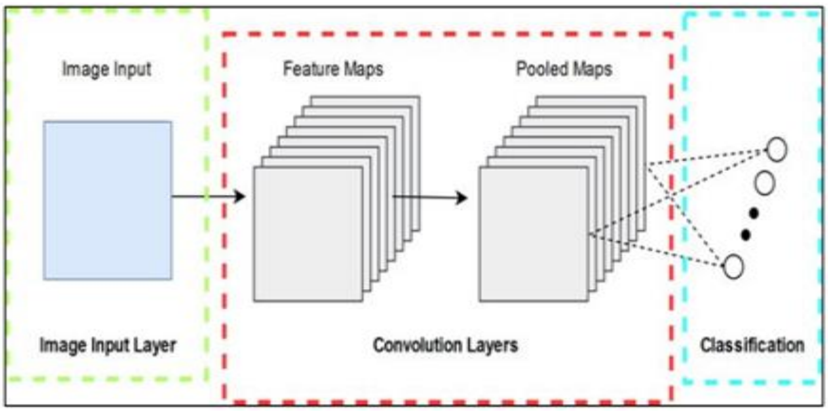

Figure (4): CNN architecture

Enhanced CNN

Enhanced CNN models involve the addition of extra layers to the traditional CNN architecture. These additional layers aim to improve the model's focus on relevant features. However, when the model is trained with limited data, the results are comparable to those obtained using the traditional CNN architecture.

Visual Geometry Group (Vgg16) Model

VGG 16 is a CNN model with 16 layers known for its effectiveness. It focuses on ConvNet layers with a 3x3 kernel size, making it simpler compared to other comprehensive models. The VGG 16 model's values are publicly available and can be downloaded for use in different systems and applications. The minimum expected input size for this model is 224x224 pixels with three channels. The model incorporates nonlinearity through the use of the kernel function and activation function. Max-pooling layers are used for spatial pooling, and three Fully-Connected (FC) layers follow the convolutional layers. The third FC layer performs 1000-way classification. The model ends with a SoftMax layer, and all hidden layers use the rectification (ReLU) nonlinearity. The structure of VGG16 is depicted by Figure (5.

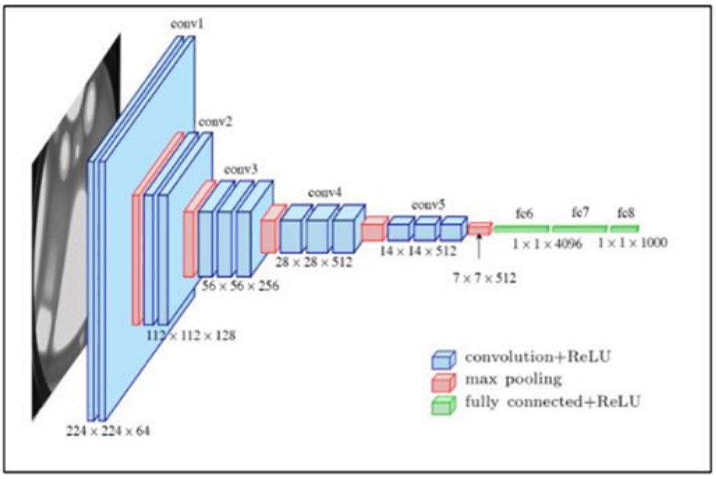

Figure (5): VGG16 architecture

RESULTS

Once all the models have been trained and evaluated using the same dataset, we can examine the results to differentiate between their performances for brain tumor classification from MRI brain images. We conducted experiments using Traditional CNN, enhanced CNN, and pre-trained VGG16 models to obtain the accuracy of these models. All results are presented corresponding to the 5-fold cross-validation for the original size of data and the augmented size of data. First phase, results of models' accuracy when they are applied to original data (i.e., without data augmentation).
50 MRI images from the dataset, 25 for normal brains and 25 for abnormal (with tumor), have been used to test the model. Table (1, Table (2, and Table (3 Present the confusion matrix for each fold for the three models. Table (1 Presents a confusion matrix for the traditional CNN model. It can be seen that the best performance of the model is at k=2, where it has correctly classified 20 images from 25 as abnormal and has also correctly classified 20 images as normal images (without tumor).

Evaluation Criteria

To evaluate models' performance, we have used three common standard statistical metrics associated with average accuracy: precision, recall, and F1-score. To calculate these criteria, we introduce some definitions of variables used in equations (1), (2), (3), and (4):
True positive (Tp): This indicates the number of correctly recognized positive samples "yes" (with tumors) in the dataset.
False positive (Fp): This value shows the number of incorrectly identified positive samples (samples without tumors) in the dataset.
True negative (Tn): This shows the number of correctly identified negative samples "No" (samples without tumors) in the dataset.
False negative (Fn): The sum of incorrectly identified negative samples (with tumors) in the dataset
Accuracy: The ratio of correctly known samples in the dataset, as given in Equation (1).

$$Accuracy = \frac{(Tp + Tn)}{(Tp + Fp + Tn + Fn)} \qquad (1)$$

Precision: The ratio of correctly identified positive images out of all those identified as positive, as given in Equation (2)

$$\text{Precision} = \frac{Tp}{(Tp + Fp)} \qquad (2)$$

Recall: The ratio of accurately identified positive images out of all the actual positive images as given in Equation (3).

$$\text{Recall} = \frac{Tp}{(Tp + Fn)} \qquad (3)$$

F1-Score: The harmonic mean of precision and recall contributes equivalent weight to both measures as a result of Equation (4)

$$\text{F1} - \text{score} = \frac{2 * (\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})} \qquad (4)$$

Table (1): confusion matrices for traditional CNN in K folds as percentages (%)

	K1		K2		K3		K4		K5	
	N	Y	N	Y	N	Y	N	Y	N	Y
N	56	44	80	20	68	32	72	28	60	40
Y	8	92	20	80	20	80	16	84	8	92

From Table (**2**We can infer that the best performance of the enhanced CNN model was in the second fold, with an accuracy of 76% for normal and 92% for abnormal tumor detection.

Table (2): confusion matrices of enhanced CNN in K folds as percentages (%)

	K1		K2		K3		K4		K5	
	N	Y	N	Y	N	Y	N	Y	N	Y
N	56	44	76	24	68	32	64	36	60	44
Y	12	88	8	92	16	84	16	84	20	80

Table (**3** Presents the confusion matrix of the VGG16 model. The best accuracy was achieved at k=2 and k=3, where the model correctly classified 96% images as normal (without tumor) and 84% as abnormal or infected by a brain tumor. At k=3, the model has successfully classified 88% and 92% images for normal and abnormal, respectively.

Table (3): confusion matrices of VGG16 in K folds as percentages (%)

	K1		K2		K3		K4		K5	
	N	Y	N	Y	N	Y	N	Y	N	Y
N	76	24	96	4	88	12	80	20	84	16
Y	4	96	16	84	18	92	12	88	8	92

The second phase is training and testing the models on the dataset after the augmentation process is achieved. The training and validation accuracies with the loss accuracies for the three models: traditional CNN, enhanced CNN, and pre-trained VGG16 are shown by Figure (**6**, Figure (**7**, and Figure (**8** Respectively.

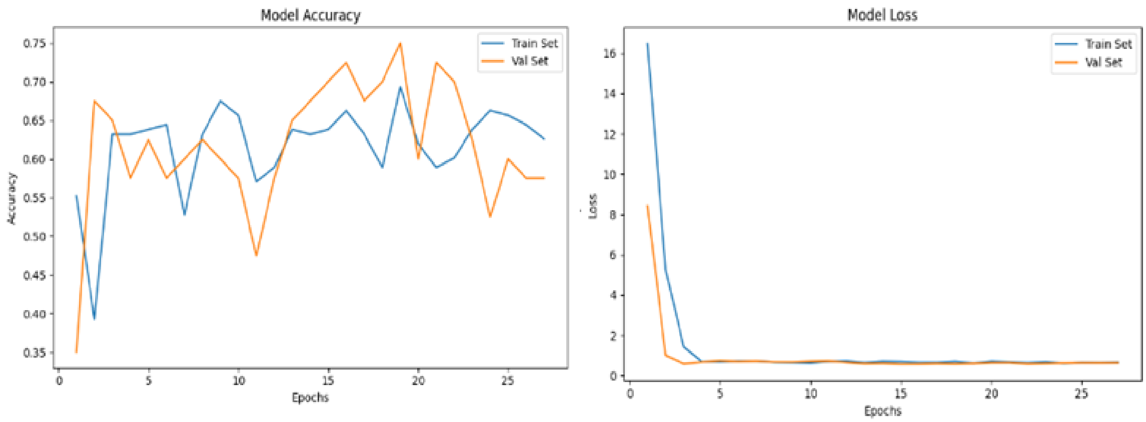

Figure (6): Accuracy and loss graph of the traditional CNN model

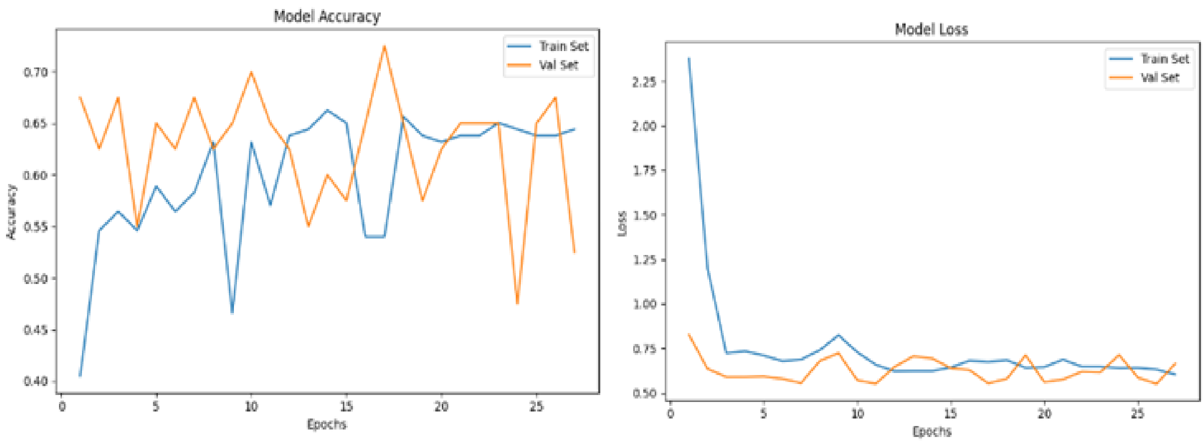

Figure (7): Accuracy and loss graph of the enhanced CNN model

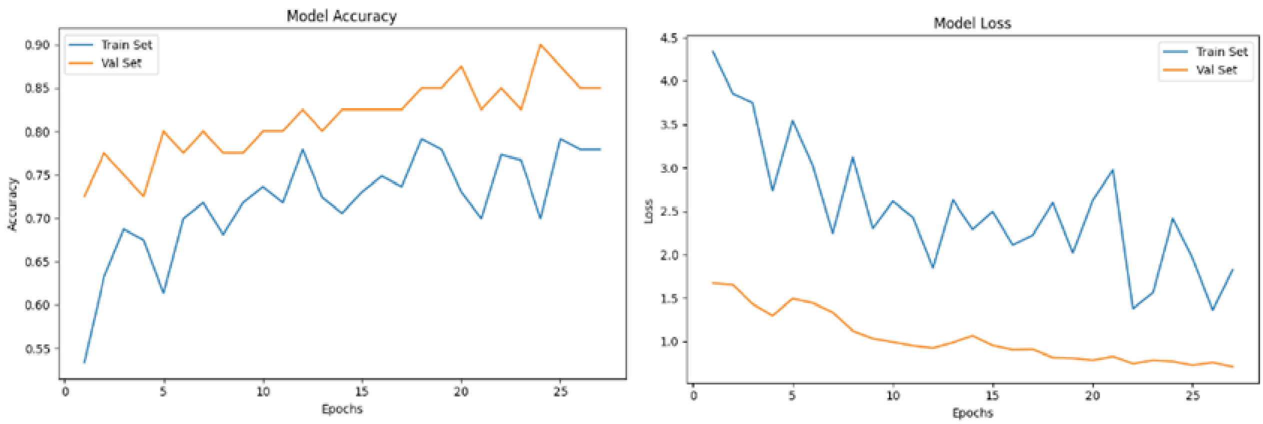

Figure (8): Accuracy and loss graph of the VGG16 CNN model

This section also presents the models' classification accuracies in separate tables based on a 5-fold cross-validation technique. Table (**4** Presents very poor accuracy of classification for the traditional model, and the best performance was 24% at k=5.

Table (4): confusion matrices of traditional CNN in K folds as percentages (%)

	K1		K2		K3		K4		K5	
	N	Y	N	Y	N	Y	N	Y	N	Y
N	0	100	0	100	12	88	20	80	24	76
Y	4	96	0	100	4	96	12	88	8	92

The enhanced CNN model shows a slightly better performance than the traditional model with augmented data, as shown by Table (**5**. At k=1, its accuracy was 84% for normal and 64% for abnormal images.

Table (5): confusion matrices of enhanced CNN in K folds as percentages (%)

	K1		K2		K3		K4		K5	
	N	Y	N	Y	N	Y	N	Y	N	Y
N	84	16	76	24	0	100	72	28	0	100
Y	36	64	28	72	0	100	28	72	0	100

The VGG16 model continuously performs better than both previous models. The best accuracy has been achieved at k=2 and k=4, where it has correctly classified 88% and 96% for normal and abnormal brain images, as shown by Table (**6** Respectively.

Table (6): confusion matrices of the VGG16 model in K folds as percentages (%)

	K1		K2		K3		K4		K5	
	N	Y	N	Y	N	Y	N	Y	N	Y
N	80	20	88	12	84	16	88	12	88	12
Y	12	88	4	96	32	68	4	96	24	76

The accuracy values obtained from this procedure are presented in the Table (**7**. Additionally, calculating the average accuracy of the models provided a clearer understanding of the best-performing model. Among the values of test accuracies all over, we got the best average 88%, 76%, and 75% for VGG16, traditional CNN, and enhanced CNN, respectively.

Table (7). Comparison between the accuracies (%) of the three models

K	With data augmentation			Without data augmentation		
	CNN	Enh. CNN	VGG16	CNN	Enh. CNN	VGG16
1	48%	74%	84%	74%	72%	86%
2	65%	74%	92%	80%	84%	90%
3	54%	50%	76%	74%	76%	90%
4	54%	72%	90%	78%	74%	84%
5	58%	50%	82%	76%	68%	88%
av.	56%	64%	85%	76%	75%	88%

Table (**8** Presents the evaluation standard statistical criteria: precision, recall, and F1 score for the three models, in case data were augmented to increase their size. It can be seen that VGG16 has recorded the highest scores in recall and F1-score, and a little bit lower than traditional CNN in precision.

Table (8). comparison between evaluation metrics for models with (data augmentation)

Mode	Precision	Recall	F1-Score
VGG16	0.96	0.88	0.918
CNN	1.0s	0.50	0.66
Enh CNN	0.72	0.75	0.734

Table (9 Presents the evaluation standard statistical criteria of the three models for the original data (after data augmentation). VGG16 model also has scored good performance over both other models with 95% and .89% for recall and F1-score respectively. Even though, enhanced CNN model has scored a little bit higher than VGG16. On average, a pre-trained model that is a transfer learning-based model shows the best performance for brain tumor detection and classification with a sample of training data.

Table (9). comparison between evaluation metrics for the three models (without data augmentation)

Mode	Precision	Recall	F1-Score
VGG16	0.84	0.95	0.89
CNN	0.80	0.80	0.80
Enhanced CNN	0.92	0.79	0.85

CONCLUSION

Three models: traditional CNN, enhanced CNN, and pre-trained CNN models have been investigated to test their performance when applied to a small dataset of brain tumor images. Two testing phases were included: one with the original dataset size and the second with data size augmentation. The k-fold cross validation also applied to all models. Results show that pre-trained (transfer learning-based VGG16) has far better performance than traditional CNN, enhanced CNN models. On the other hand, data augmentation is not always a solution for a small dataset. Results inform us that the performance of all models was higher when models used the original dataset, although it was smaller than the augmented dataset.

Duality of interest: The authors declare that they have no duality of interest associated with this manuscript.

Author contributions: Contribution is equal between authors.

Funding: No specific funding was received for this work.

REFERENCES

Abiwinanda, N., Hanif, M., Hesaputra, S. T., Handayani, A., & Mengko, T. R. (2019). Brain Tumor Classification Using Convolutional Neural Network. In L. Lhotska, L. Sukupova, I. Lacković, & G. S. Ibbott (Eds.), *World Congress on Medical Physics and Biomedical Engineering*

2018 (Vol. 68/1, pp. 183–189). Springer Nature Singapore. https://doi.org/10.1007/978-981-10-9035-6_33

Badža, M. M., & Barjaktarović, M. Č. (2020). Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network. Applied Sciences, 10(6), 1999. https://doi.org/10.3390/app10061999

Bairagi, V. K., Gumaste, P. P., Rajput, S. H., & Chethan K. S. (2023). Automatic brain tumor detection using CNN transfer learning approach. Medical & Biological Engineering & Computing, 61(7), 1821–1836. https://doi.org/10.1007/s11517-023-02820-3

Das, S., Aranya, O. F. M. R. R., & Labiba, N. N. (2019). Brain Tumor Classification Using Convolutional Neural Network. 2019 1st International Conference on Advances in Science, Engineering *and Robotics Technology (ICASERT)*, 1–5. https://doi.org/10.1109/ICASERT.2019.8934603

Deng, L., & Yu, D. (2014). Deep Learning: Methods and Applications. *Foundations and Trends® in Signal Processing*, *7*(3–4), 197–387. https://doi.org/10.1561/2000000039

Rai, H. M., & Chatterjee, K. (2020). Detection of brain abnormality by a novel Lu-Net deep neural CNN model from MR images. *Machine Learning with Applications*, *2*, 100004. https://doi.org/10.1016/j.mlwa.2020.100004

Sawant, A., Bhandari, M., Yadav, R., Yele, R., & Bendale, M. S. (n.d.). *Brain Cancer Detection from Mri: A Machine Learning Approach (Tensorflow)*. *05*(04).

Tamm, M.-O., Muhammad, Y., & Muhammad, N. (2020). Classification of Vowels from Imagined Speech with Convolutional Neural Networks. *Computers*, *9*(2), 46. https://doi.org/10.3390/computers9020046