Journal of Imaging Informatics in Medicine
https://doi.org/10.1007/s10278-025-01470-1

ORIGINAL PAPER

I-BrainNet: Deep Learning and Internet of Things (DL/IoT)–Based Framework for the Classification of Brain Tumor

Abdullahi Umar Ibrahim[1,2] · Glodie Mpia Engo[3] · Ibrahim Ame[3,4] · Chidi Wilson Nwekwo[1] · Fadi Al-Turjman[3,4]

Received: 22 October 2024 / Revised: 21 February 2025 / Accepted: 24 February 2025

Abstract

Brain tumor is categorized as one of the most fatal form of cancer due to its location and difficulty in terms of diagnostics. Medical expert relies on two key approaches which include biopsy and MRI. However, these techniques have several setbacks which include the need of medical experts, inaccuracy, miss-diagnosis as a result of anxiety or workload which may lead to patient morbidity and mortality. This opens a gap for the need of precise diagnosis and staging to guide appropriate clinical decisions. In this study, we proposed the application of deep learning (DL)–based techniques for the classification of MRI vs non-MRI and tumor vs no tumor. In order to accurately discriminate between classes, we acquired brain tumor multimodal image (CT and MRI) datasets, which comprises of 9616 MRI and CT scans in which 8000 are selected for discrimination between MRI and non-MRI and 4000 for the discrimination between tumor and no tumor cases. The acquired images undergo image pre-processing, data split, data augmentation and model training. The images are trained using 4 DL networks which include MobileNetV2, ResNet, Ineptionv3 and VGG16. Performance evaluation of the DL architectures and comparative analysis has shown that pre-trained MobileNetV2 achieved the best result across all metrics with 99.94% accuracy for the discrimination between MRI and non-MRI and 99.00% for the discrimination between tumor and no tumor. Moreover, I-BrainNet which is a DL/IoT-based framework is developed for the real-time classification of brain tumor.

Keywords Brain tumor · Deep Learning · Internet of Things · MRI · Medical imaging · Diagnosis

Introduction

As part of the central nervous system, the brain plays a vital role in carrying out several functions such as learning, coordination, memory, movement, etc. Therefore, malfunctioning of the brain as a result of diseases, injuries or accident can be detrimental. Diseases associated with the brain such as the brain tumor exert a significant impact on global healthcare system accounting for thousands of cases globally [1]. According to the National Brain Tumor Society (NBTS) and Cancer Research UK, brain cancer is ranked as the 10th leading cause of death globally [2, 3]. Brain tumor or cancer can be described as the uncontrollable growth of brain cells. The symptoms of brain tumor vary based on the region impacted. The most common symptoms include dizziness, seizures, persistent headaches, poor vision and hearing, etc. [4, 5].

Brain tumors can be sub-divided into benign and malignant. Moreover, it can be classified into 4 categories based on molecular features and histopathology which include grade I, II, III and IV. Benign brain tumors are the types that are restricted to a specific region and hardly spread. They can be divided into meningioma, craniopharyngiomas and neuromas. While malignant brain tumor are the types that spread to other neighboring tissues and organs. They can be divided into astrocytoma, oligodendrogliomas, glioma, glioblastoma. The recent edition of the WHO classification of brain tumor also classified glioma into adult and pediatric types [6–8].

✉ Abdullahi Umar Ibrahim
Abdullahi.umaribrahim@neu.edu.tr

1 Department of Biomedical Engineering, Near East University, Mersin 10, Nicosia, Turkey

2 Research Centre for Science, Technology and Engineering (BILTEM), Near East University, TRNC, Mersin 10, 99138 Nicosia, Turkey

3 Artificial Intelligence, Software, Information Systems Engineering Departments, AI and Robotics Institute, Near East University, Mersin10, Nicosia, Turkey

4 Research Centre for AI and Iot, Faculty of Engineering, University of Kyrenia, Mersin10, Kyrenia, Turkey

Published online: 10 March 2025

Diagnosis or screening patients suspected with brain cancer involve physical examination followed by biopsy which is regarded as the standard technique. Other conventional approach includes cerebrospinal fluid (CSF) biomarker evaluations and medical imaging such as Magnetic Resonance Imaging (MRI), Magnetic Resonance Spectroscopy (MRS), Single Photon Emission Computed Tomography (SPECT), Position Emission Tomography (PET), Computed Tomography (CT) scan, etc. [9–11]. Among several approaches, biopsy is regarded as the standard approach for estimation and classification of tumor grades and confirmation of the aggressiveness of the tumor type. The technique enables visualization of the cells, tissues and variations based on color, shape, size and distribution [12, 13].

Accurate and precise diagnosis of patients suspected with brain cancer is highly crucial for timely treatment and cost savings. Despite the dependency of biopsy technique, it is hampered by several limitations which include time-consuming, tediousness and its invasive nature which can be life threatening. In terms of medical imaging, MRI is the most popular technique employed by medical experts. However, the approach suffers from several setbacks which include the requirement of experience radiologists or expert, time consuming, error prone and difficulty in classifying or determining tumor type [14, 15].

As a means to counter these challenges, scientists adopted Artificial Intelligence (AI)-based approach which encompasses Machine Learning (ML) and Deep Learning (DL). AI-based techniques have shown a great promise in diagnosis and treatment of neuro-oncology [1, 16]. AI has shown prospect in management of brain tumors through the detection of abnormalities in medical images or by identifying patterns and features which are not easily discernible to human eyes, enhancing MRI imaging, tumor grading, optimizing workflows, etc. Consequently, numerous stationary models limited to clinical settings have been developed for identification and classification of brain tumors. However, these models cannot be accessible remotely which limit their applications. In order to address these challenges, we developed a remote and on-site AI/IoT-based brain tumor classification framework known as I-BrainNet for accurate and automated discrimination between MRI and non-MRI as well as discrimination between tumor and no tumor. The overall technique revolves around data curation, data preprocessing, implementation of DL models, evaluation of models using several metrics and web development. The main contributions of this study are listed below:

- The deployment of AI-driven models for the accurate discrimination between MRI and non-MRI cases
- The deployment of AI-driven models for the accurate discrimination between malignant and benign brain tumors.
- Performance evaluation and comparative analysis between models and previous studies.
- Development of remote and on-spot framework for the classification of brain tumors.

The remaining part of this study are structured as follows: the "Literature Survey" section presents literature surveys which overview previous studies that implemented DL-based models for the detection or classification of brain tumor. The section outlines the study's methodology which cover data collection, image processing, implementation and training of models. The "Result and Discussion" section covers the result and discussion which include performance evaluation of models, comparative analysis and web development. While the final section entails the conclusion.

Literature Survey

Several DL-based algorithms have shown to aid medical experts in identifying molecular subtypes of tumors and providing quantitative measurements. The field of AI has significantly enhanced medical imaging by providing accurate, precise and detailed image analysis for classification of tumors and grading. Several studies have explored the capabilities of these algorithms for the detection of brain cancer. For example, the study proposed by Lamrani et al. [17] reported the use of Convolutional Neural Network (ANN) CNN model for brain tumor detection and binary classification. The study retrieved image dataset from Kaggle repository which contain 3000 images (1500 tumor and 1500 no tumor). The images are pre-processed based on normalization of sizes and image partitioning (into 80% training, 10% validation and 10% testing). In order to classify the images, a CNN is implemented which is constructed based on 4 convolutional layer, 3 Max pooling layers, flatten and 6 dense layers. The performance evaluation has shown that the model achieved 96.33% accuracy, 97.93% precision, 95% sensitivity, 75.72% specificity and 96.44% F1-Score.

El-Feshawy et al. [18] proposed two different approaches for the detection of tumor from MRI. The first approach revolves around the use of DL-CNN for the direct classification of the acquired images while the second approach revolves around the implementation of IoT-based framework that integrates multiuser detection system by sending the images to the cloud for early detection of brain tumors. In order to train 253 MRI acquired from Kaggle titled Brain MRI Images for Brain Tumor Detection dataset (with 155 tumor and 98 no tumor), a modified version of ResNet-18 known as OMRES. Two vital hyper-parameters are employed to fine-tune the model by testing different optimizers using different learning rates, batch sizes as well as constant number of epochs. The second hyper-parameter revolves around assessing the impact of varying dropout rates. The modified model is compared with traditional pre-trained models and the result has shown that OMRES achieved superior performance with 98.67% highest rated accuracy.

Brindha et al. [19] proposed the implementation of Artificial Neural Network (ANN) and CNN for detecting the presence of brain tumor. The overall methodology revolves around data collection in which MRI is acquired from GitHub. The dataset contains 2065 images (980 noncancerous and 1085 brain tumor). The acquired dataset is processed through image resizing and partition into training, validation and testing. The images are trained using 2 DL techniques. The first technique involves the use of 7-layer ANN compiled with Adam optimization technique and binary cross entropy loss function. The second technique revolves around the implementation of 7-layer CNN with 5 convolution layers designed based on convolution, Max pooling and dropout filled by flatten after the 5th convolutional layer, fully connected layer and output or classification layer. Evaluation of the model has resulted in 80.77% testing accuracy and 71.51% validation accuracy using ANN and 65.21% testing accuracy and 94.00% validation accuracy using CNN.

Sawant et al. [20] implemented a 5-layer CNN architecture known as Le-Net for classification of brain cancer and normal images. The model is designed with 2 convolution layer, 2 pooling layer and fully connected layer. The study employed 1800 MRI (900 tumors and 900 noncancerous) which were curated from two different sources which include BRATS and TCGA. In order to increase training set, two data augmentation techniques which include horizontal flipping and rotation were conducted. The proposed architecture is implemented in TensorFlow and evaluated based on training and validation accuracy. Evaluation of the model based on validation accuracy resulted in 98.6%.

In order to address the issue of miss-diagnosis of brain tumor, Alsubai et al. [21] proposed the implementation of hybrid model which combines CNN and Long Short-Term Memory (LSTM) for accurate classification and prediction of brain tumors from MRIs. The experimental setup is designed based on data curation of MRI images from the MRI brain tumor dataset which is publicly accessible on Kaggle with 253 images (155 tumor and 98 no tumor), followed by pre-processing through cropping in order to remove noise, resizing and feature extraction using CNN. The performance evaluation of the proposed CNN-LSTM resulted in an accuracy, precision, recall and F1-measure of 99.1%, 98.8%, 98.9%, and 99.0%, respectively.

The study proposed by Kolla et al. [22] developed an autonomous brain tumor segmentation and detection model based on CNN that included a Local Binary Pattern (LBP) and a multi-layered Support Vector Machine (SVM). The study acquired Kaggle medical image database dataset. The images are processed using several image filtering techniques, followed by image intensity normalization, patch extraction and feature extraction, conversion from RGB to grayscale and lastly the use of LBP. The extracted features are trained using CNN, CNN-SVM and proposed multi-SVM + CNN. Evaluation of the models and comparison has shown that proposed multi-SVM + CNN achieved the highest accuracy with 99.23%, sensitivity of 95.73%, and specificity of 97.12% and precision of 97.34%.

Arora et al. [23] proposed the application of DL-based techniques for the binary classification of MRI images into tumor and no tumor. The research is structured based on data collection, medical image processing, pattern analysis, enhancement, segmentation and binary classification of brain MRI. In terms of data collection, the study acquired the BRATS2015 which comprises of 3764 images of benign and malignant cases. The acquired images were pre-processed based on several techniques which include feature selection, cleaning, standardization, normalization and resizing. The pre-processed images are subsequently fed into several models which include classical ML models (SVM and Random Forest classifiers) and pre-trained models which include VGG16, InceptionV3 and ResNet. Performance evaluation and comparative analysis has shown that VGG16 outperformed other models with 90.54% testing accuracy followed by ResNet with 81.92% testing accuracy.

The study proposed by Khan et al. [24] implemented 2 DL based mode for binary classification of MRIs into normal and abnormal cases and multiclass classification into meningioma, glioma, and pituitary brain tumors. In order to achieve accurate classification, the study curated two datasets which include CE-MRI Figshare dataset with 3064 images and Harvard Medical Dataset (HMD) with 152 MRI images, respectively. In order to enlarge the training set, data augmentation techniques which include shearing, flipping (horizontal and vertical), zooming, rotation, height shift range and width shift range are implemented. The images are used to train models which include 23-layer CNN and proposed TL model based on VGG16 combined with proposed CNN to address the issue of overfitting. Performance evaluation of the two techniques and datasets has shown that the proposed 23-layer CNN achieved 97.8% accuracy, 96.5% average precision, 96.4% average recall and 97.4% F1-Score using the Figshare dataset for multiclass classification while the proposed CNN model achieved 100% accuracy, recall, precision and F1-Score using the HMD (binary classification).

The use of balanced binary Tree CNN (BT-CNN), CNN-KNN for the binary classification of brain tumor using large-scale dataset curated from different domains and studies is reported by Chauhan et al. [25]. The research is designed based on data collection, data pre-processing, data split, model implementation and performance evaluation. The study acquired 4 datasets which include Figshare Brain Tumor Dataset (Figshare, 2024), Br35H dataset (Brain Tumor Detection, 2024); Sartaj (2024) combined together to form a single dataset with 16,350 images. The dataset undergoes pre-processing in order to eliminate noise, the use of Contrast Limited Adaptive Histogram Equalization (CLAHE) to solve the issue of over-amplification of contrast, cropping and scaling. Performance evaluation of the model resulted in 96.06% testing accuracy. The summary of related work is summarized in Table 1.

Limitation of Existing Studies

Despite the fact there are ample studies in the literature that proposed the deployment of AI-driven models for the binary discrimination of brain tumors into tumor and no tumor or benign and malignant; however, majority of existing studies deployed models for onsite detection. Moreover, plethora number of existing studies only deployed single models or customized models, which may not serve as the best architecture that can efficiently learn underlying patterns, and features of brain tumors. Thirdly, few articles proposed real-time classification of MRI into tumor and no tumor. Therefore, this study will attempt to address these limitations by developing AI/IoT-platform that can discriminate between MRI and non-MRI as well as tumor and no tumor using pre-trained models.

Methodology

The overall methodology of this research revolves around collection of MRI images from publicly accessible domain, followed by image pre-processing via labelling, resizing and normalization, image partitioning into training, validation and testing sets, data augmentation via flipping, and zooming and training of the models. Moreover, the methodology also covers the description of pre-trained models. The overall methodology of this study is summarized in Fig. 1.

Data Collection

Training DL models require extensive amount of data in order to achieve optimum prediction. Thus, in this study, we utilized the Brain tumor multimodal image (CT and MRI), which is publicly accessible on Kaggle through the following link: https://www.kaggle.com/datasets/murtozalikhon/brain-tumor-multimodal-image-ct-and-mri. The dataset comprises of both MRI and CT scan images. The dataset comprises of two folders. The first folder contains 5000 MR mages (2000 healthy vs 3000 tumor). The second folder contains 4618 CT Scan: images (2300 healthy and 2318 tumor) as shown in Fig. 2. Repetitive and redundant images were removed which led to 8000 images (4000 MRI and 4000 CT scans). The rationale behind using CT scans in the first step is to discriminate between MRI and non-MRI cases. Therefore, the inclusion of non-MRI cases contributes to the CNNs efficiencies in recognition of MRI cases and also helps in minimizing false positive and false negative results in practice where some of non-MRI images might appear.

Smaller brain tumors (less than 1 cm) are often asymptomatic but critical to detect early, as they may be malignant (e.g., gliomas) or benign (e.g., meningiomas). Early detection improves surgical outcomes and survival rates. One of the challenges of conventional MRI/scans is that it may miss small tumors due to limited resolution (typically ~ 1–3 mm for clinical MRI), partial volume effects, or low contrast. Considering the fact that the dataset retrieved comprises of MRI images acquired from 9 databases, this ensures diversity of tumors ranging from small to big as well as benign, malignant or glioma, meningioma and pituitary. Among the 3 types, glioma is the largest (2–10 cm, depending on low grade or high grade), followed by meningioma (2–7 cm), while pituitary is the smallest (≤ 1 cm or 10 mm); however, their sizes on MRI scans can vary depending on the tumor type, growth rate, and location. The description of the dataset utilized is summarized in Table 2.

Table 1 Summary of related work

References	Models	Number of images	Classes	Accuracy (%)
[17]	Customized CNN	3000	2	96.33
[18]	ResNet-18 (OMRES)	253	2	98.67
[19]	Customized CNN	2065	2	80.77
[20]	Le-Net	1800	2	98.60
[21]	CNN-LSTM	253	2	99.1
[22]	Multi-SVM + CNN	-	2	99.23
[23]	VGG16	3764	2	90.54
[24]	23-layer CNN	3064	3	97.80
[25]	BT-CNN	16,350	2	96.06

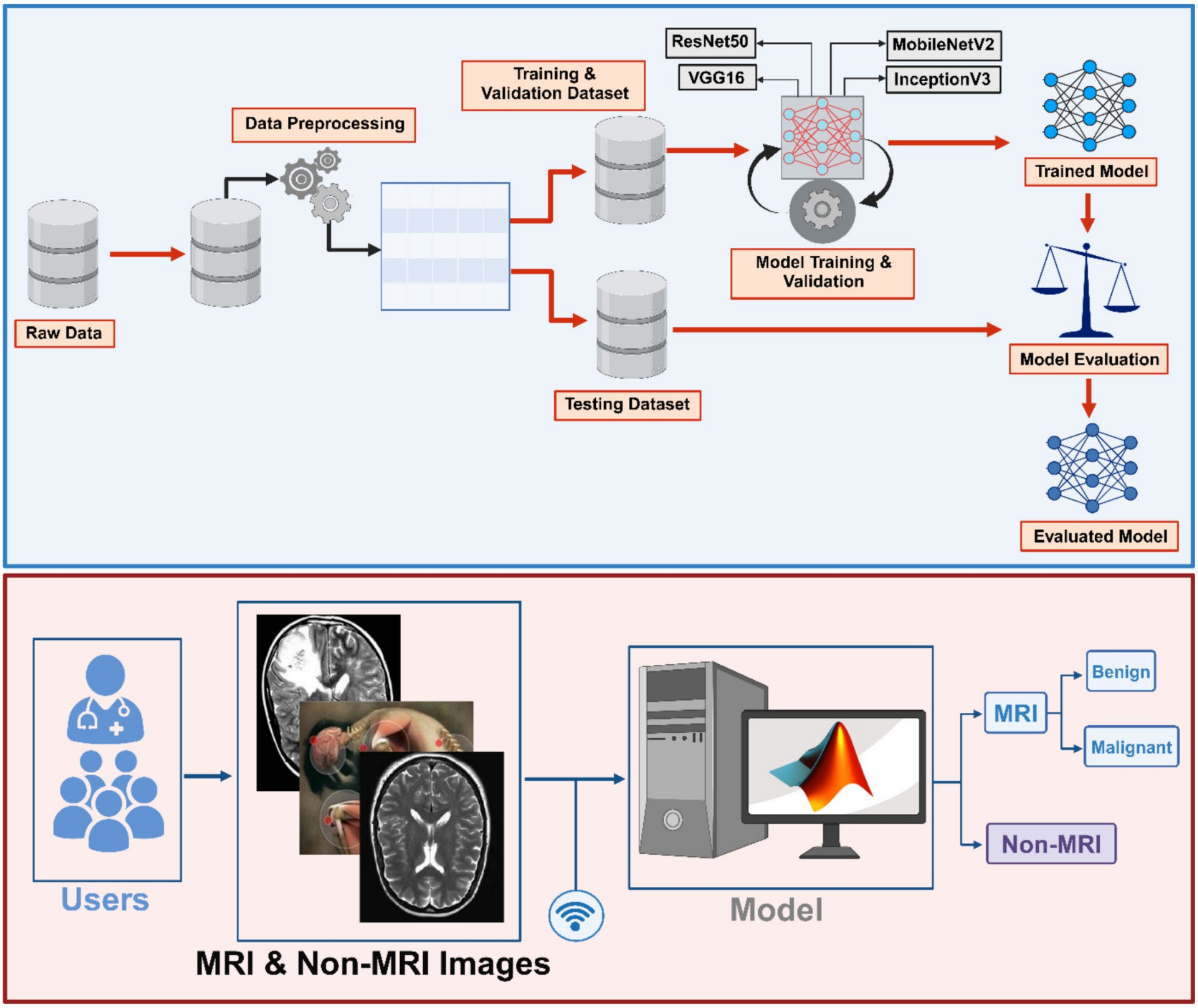

Fig. 1 Overall Methodology

Data pre-Processing

Several data pre-processing steps were carried out which include renaming, labeling, resizing and normalization. In the first step, the entire folder "MRI and CT" is relabeled as MRI and non-MRI, while the MRI sub-folder is relabeled as tumor vs no tumor. This step was carried out so that the flow generator in the code could pick up these new identifiers as training classes for the models. The images were resized to 224 × 224 to fit VGG16, ResNet50 and MobileNet input sizes. For InceptionV3, the input images are resized to 299 × 299 which is specific for the model architecture. Moreover, the images were normalized to values between 0 and 1 to reduce the variance between the dataset and to help achieved faster training. However, images did not go through any form of denoising as the images were sharp and clear and the nature of the images themselves require them to be clear and visible for review.

Data Split and Data Augmentation

Splitting data for training and evaluation of the model using testing set is very crucial for generalizability. Several data split ratios have been reported in the literature which include 2-split technique such as 60:50; 70:30, 80:20 for training and testing respectively and 3-split technique which include 60:20:20, 70:15:15 and 80:10:10 for training, validation and testing respectively. Thus, in this study we opted for 3-split technique which include training: validation: testing. Initially, the images are split into training (80%) and testing (20%). From the training set, a soft split of 80% for training and 20% for validation.

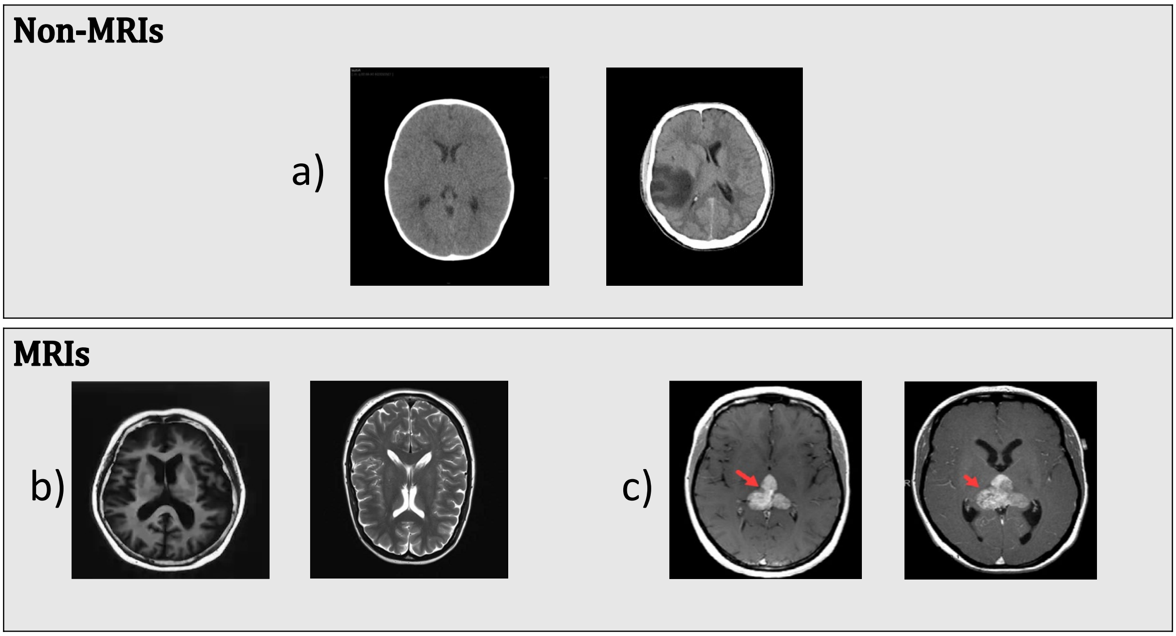

Fig. 2 Sample of MRI and non-MRI. (a) Non-MRI (CT scans). (b) No Tumor. (c) Tumor

Table 2 Description of dataset

Type	Classes	Number of images
MRI	Tumor	2000
	No tumor	2000
	Total	4000
Non-MRI	CT scans	4500

Table 3 Data split

Dataset	Split	Percentage (%)	Number of images
MRI vs non-MRI	Training set	64	5120
	Validation set	16	1280
	Test set	20	1600
	Total	100	8000
Tumor vs non-tumor	Training set	64	2560
	Validation set	16	640
	Test set	20	800
	Total	100	4000

This method of splitting ensures that the test set is not entangled in any way with the training data, and if there was any shuffling, none of the test data would leak into the model. The final split ratios are 64%, 16%, and 20% for training, validation and testing. The data split as shown in Table 3.

In the case of data augmentation, light data augmentation was conducted on the training set to increase the number of images available. The images were randomly flipped horizontally or vertically and then zoomed by a factor of 0.2 (20%). This led to the generation of 2 new image variations, thereby, tripling the sample size of the training set. Thus, MRI vs non-MRI dataset increased from 5120 to 15,360, while the tumor vs no tumor dataset increases from 2560 to 7680.

Deep Learning Models

ResNet

Residual Network also known as ResNet is deep CNN developed for image recognition and classification tasks. The network was introduced by H. Kaiming in an article titled: inspired by the 2015 paper titled "Deep Residual Learning for Image Recognition". After the introduction, the model became the best model in the ILSVRC 2015 competition with 3.57% error rate on the ImageNet dataset [26, 27].

The concept of ResNet was conceived as a solution toward solving the problems suffer by stacking layers known as vanishing gradient. Due to the success of AlexNet and VGGNETs, scientists started stacking more layers which result in poor performance. ResNet addressed this issue by introducing “skip connections” which enables the network to skip certain layers, especially if they do not contribute to the efficiency of the network in terms of performance. In terms of architecture, ResNet comprises of residual blocks or bottleneck which consists of several layers such as convolutional layers, batch normalization, and activation functions [26–28]).

There are several variants of ResNets which include ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152. Thus, the number assigned to each variant correspond to the total number of layers in the network. In this study, we implemented ResNet-50. The network is made of 50 layers which can be sub-divided into five blocks. The first block comprises of convolutional layer followed by pooling layer. The next block comprises of 2 convolutional layers and a skip connection. The third, fourth, and fifth blocks are made of several convolutional layers. The last layer is a fully connected layer which uses SoftMax for classification of output [27].

VGGNet

Visual Geometry VGG is deep learning CNN developed and introduced in 2012 by scientists known as K. Simonyan and A. Zisserman working at Oxford University. Their work is published using the title “Very Deep Convolutional Networks for Large-Scale Image Recognition”. Since the introduction of VGGNETs, the model has become one of the most popular networks along with ResNet and GoogleNet. The network was ranked second in the ILSVRC competition with top-5 test accuracy of 92.7% in ImageNet dataset which comprises of more than 14 million images belonging to 1000 classes. The model is designed as a modification to previously introduced AlexNet by replacing big kernel filters (11×11 and 5×5) with small filters (3×3). The model has several variants but the most common ones are the VGG16 and VGG19. Other less common variants include VGG11 and VGG13 [29].

In terms of architecture, the model accepts image with an input size of 224×224 RGB. VGG16 comprises of 16 layers in which 13 are convolutional layers and 3 fully connected layers with SoftMax as the final layer use for classification. The network can be divided into 6 units separated by pooling layers. The first and second units comprises of 2 convolutional layers with pooling in between and after the second layer. The third, fourth and fifth layers all comprise of 3 convolutional layers separated by pooling layers. The fifth unit is followed by pooling layer and subsequently 3 dense or fully connected layers [29, 30].

Inception

The concept of inception network or GoogleNet goes back to 2014 where scientists presented an article titled “Going deeper with convolutions”. Since the introduction of the model, it has become one of the most sought out network for image detection and classification. Moreover, the network was ranked top best model in the ILSVRC14 competition. The main difference between the previous models such as VGGNET, AlexNet, Le-Net and inception is the introduction of sparse matrices and 1×1 convolutions. The introduction of sparse matrices contributes to computational resources-saving while the addition of 1×1 convolution addresses the issue of downsampling where information is loss (i.e., due to over-downsampling or accumulation over time), thus, 1×1 convolution addresses the issue by pooling across the channel unlike pooling inside a single channel [31, 32].

Previous models attempted to improve performance through stacking of layers which led to several problems such as vanishing gradient, increasing parameters and computational cost and overfitting. Therefore, inception networks addressed these issues by developing sparsely connected architectures instead of densely connected architectures. In terms of architecture, inception architectures are made of multiple parallel convolutions with different filter sizes. The main idea behind the network is the addition of 1×1 convolutional layers before adding convolutional layers with larger filters such as 5×5 and 3×3 which are made of the structure inception modules as the building blocks of the model [31, 33].

MobileNet

MobileNet is introduced by Andrew G Howard in 2017. The network can be described as a CNN developed for mobile and embedded applications. The network was designed based on streamlined architecture that applies depth-wise separable convolution to construct lightweight deep neural networks with low latency for embedded and mobile devices. The model architecture comprises of 28 layers including both depth-wise and point-wise convolution. All layers are designed using batch normalization and Rectified Linear Unit (ReLu) nonlinearity except the last fully connected layer which use SoftMax for classification [34]. The model consists of several operations which include normal convolution, batch normalization and ReLu nonlinearity to the factorized layer with depth-wise convolution, followed by 1×1 pointwise convolution, batch normalization and ReLu after every convolution. Strider convolution is applied to downsample the output in both the depth-wise convolution and the first layer. A final average pooling operation is applied to decrease the spatial resolution to 1 before feeding it into the fully connected layer [34, 35].

Model Training

The dataset was divided into training, validation and testing sets, following a typical ratio of 64:16:20 to make sure that the models had adequate amount of data to learn and simultaneously avoid overfitting based on using a different test set. The models were trained for 20 epochs with a batch size of 4. The training process enables the models to encounter a wide variety of image variations, significantly enhancing its ability to generalize and perform well on unseen data. For optimization of the model, the Adam optimizer was used widely in deep learning with the learning rate set to 0.0001. All the models in the study were pretrained models. The use of transfer learning helped immensely in reaching high accuracies. However, the images had to be trained with some of the layers unfrozen. For each architecture, we have detailed the number of layers that were frozen and unfrozen. For Inception V3 with 311 layers, only the last 2 were unfrozen and then 2 dense layers were added. For MobileNetv2 with 88 layers in total. None of the layers were unfrozen, as it seemed to pick up really good patterns in the images; however, 2 dense layers were added. For ResNet50 with 50 layers, 2 were unfrozen at the end and 2 dense layers were added. And lastly, for VGG16 with 16 valid layers in total, none of the layers were unfrozen; however, 2 dense layers were added.

All models were trained on a computer Dell with Intel® Core™ i9-11750H with a processor speed of 3.60 GHz, a Ram of 32 GB, 1 TB SSD and, a 64-bit operating system. Other specifications include the Nvidia GeForce RTX 2080 TI driver which is one of the most powerful GPUs in NVIDIA's lineup, offering impressive performance and features. It has 11 GB of GDDR6 memory and 4352 CUDA cores, which allows it to handle complex graphics and computational tasks with ease.

Result and Discussion

The field of CAD which incorporates image processing, AI and its subsidiaries such as ML and DL are changing the landscape of disease diagnosis into a more robust, automated, smart, accurate, precise, fast and minimal error method. The integration of AI-aided approaches has shown to assist medical experts in making decision about prognosis. In line with this, we implemented 4 pre-models which include MobileNetV2, InceptionV3, ResNet-50 and VGG16 for the classification of MRI and non-MRI cases as well as tumor and no tumor cases. The models are trained using original and augmented data and evaluated based on test set using several metrics which include accuracy, sensitivity, specificity, recall, precision Mathew Correlation Coefficient (MCC), Area under the Curve (AUC) and F1-score.

Accuracy

Accuracy in ML and DL can be described as the measure of how accurate (i.e., sensitive and precise) a ML-model makes correct prediction or classification. Moreover, accuracy can also be defined as the sum of correctly predicted outcome over total outcome. For example, if a model classified 900 images out of 1000 irrespective of tumor and no tumor (i.e., 500 tumor and 400 no tumor), the accuracy is 90 over 100 which is equal to 0.9 or 90%. The formula of accuracy derived from confusion matrix is represented mathematically as

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Specificity

Specificity is another common evaluation metrics that seeks to answer the question: how many patients are actually or correctly predicted or classified as negative cases. Specificity can be represented mathematically as

$$Specificity = \frac{TN}{TN + FP} \tag{2}$$

Recall

Recall based on medical image classification can be described as true positive rate (TPR) or the number of positive cases that are predicted or classified as positive. Recall can be represented mathematically by the following equation:

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

Precision

Precision or True Negatives (TNs) is the opposite of recall. In this regard, precision can be described as the number of TN cases that are actually or correctly predicted or classified as negative by the model. Precision can be represented mathematically by the following equation:

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

F1-Score

F1 score is a ML evaluation metric that is applied mostly for classification especially when there is class imbalance. It combines both precision and recall as its basic foundation to form a single metric. F1 score can also be defined mathematically as the harmonic mean of recall and precision. The formula of F1-score is:

$$F1 = 2 * \frac{PrecisonXRecall}{Precison + Recall} OR \frac{2TP}{(2TP + FP + FN)} \tag{5}$$

Mathew Correlation Coefficient (MCC)

MCC is a statistical tool developed by Mathew in 1975 which is used to compute the difference between actual and predicted value. The metric is also considered as one of the best performance metrics for classification tasks which helps to summarize confusion metrics based on the following 4 elements: TP, TN, FP and FN. MCC is mathematically represented in the equation below:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{6}$$

Area Under the Curve (AUC)

AUC present the cumulative measure of performance across all classification thresholds. The metric is widely used to evaluate the capability of model to discriminate between binary classes; positive and negative. The value of AUC ranges between 0 and 1. A model with an AUC of 1 indicates 100% correct predictions while 0 indicates 0% prediction. The training curves of models deployed for discrimination between MRI vs non-MRI and also tumor vs no tumor are presented in Fig. 3 and Fig. 4 respectively.

Performance Evaluation of Models Deployed for Binary classification of MRI vs Non-MRI

The performance evaluation of MobileNet has shown that the model achieved 99.94% accuracy, 99.88% precision, 100.00% recall, 100.00% sensitivity, 99.88% specificity, 99.94% F1-score, 99.88% MCC and 100.00% AUC. The

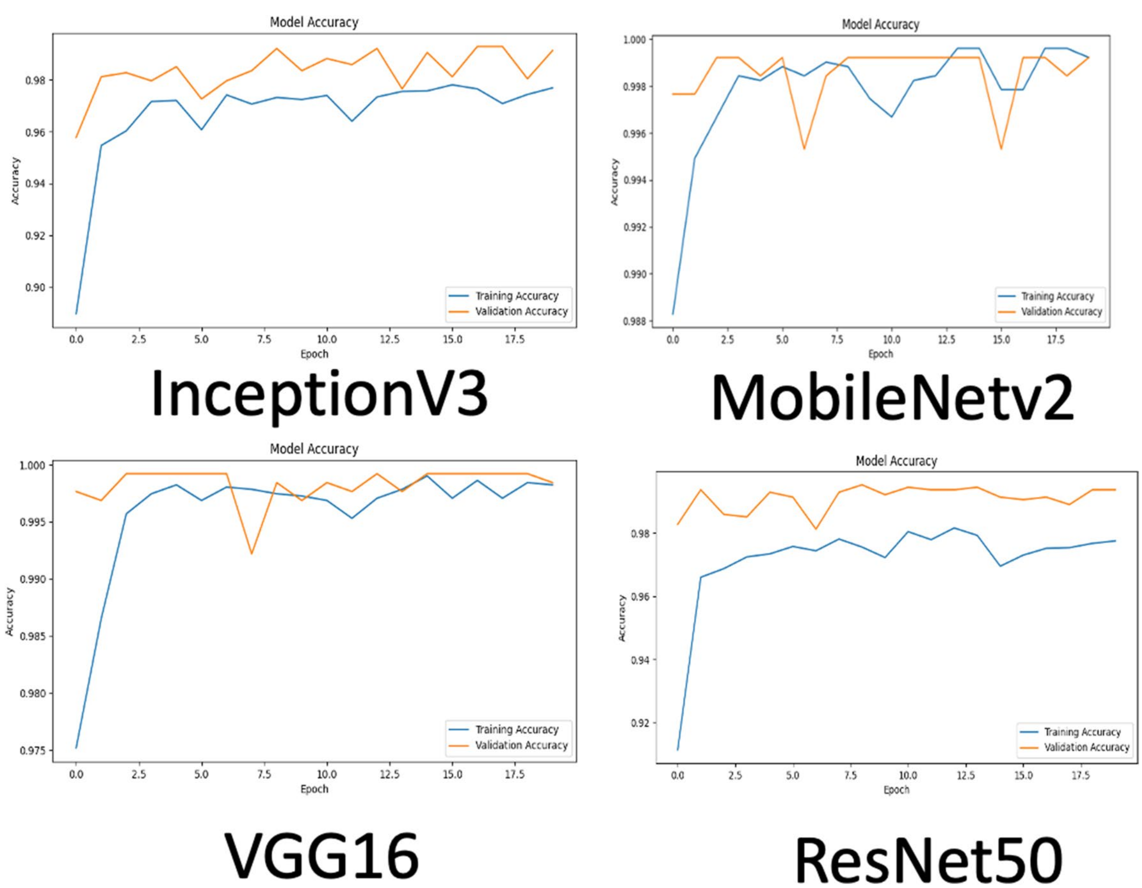

Fig. 3 Training curves of pre-trained models for binary classification of MRI and non-MRI

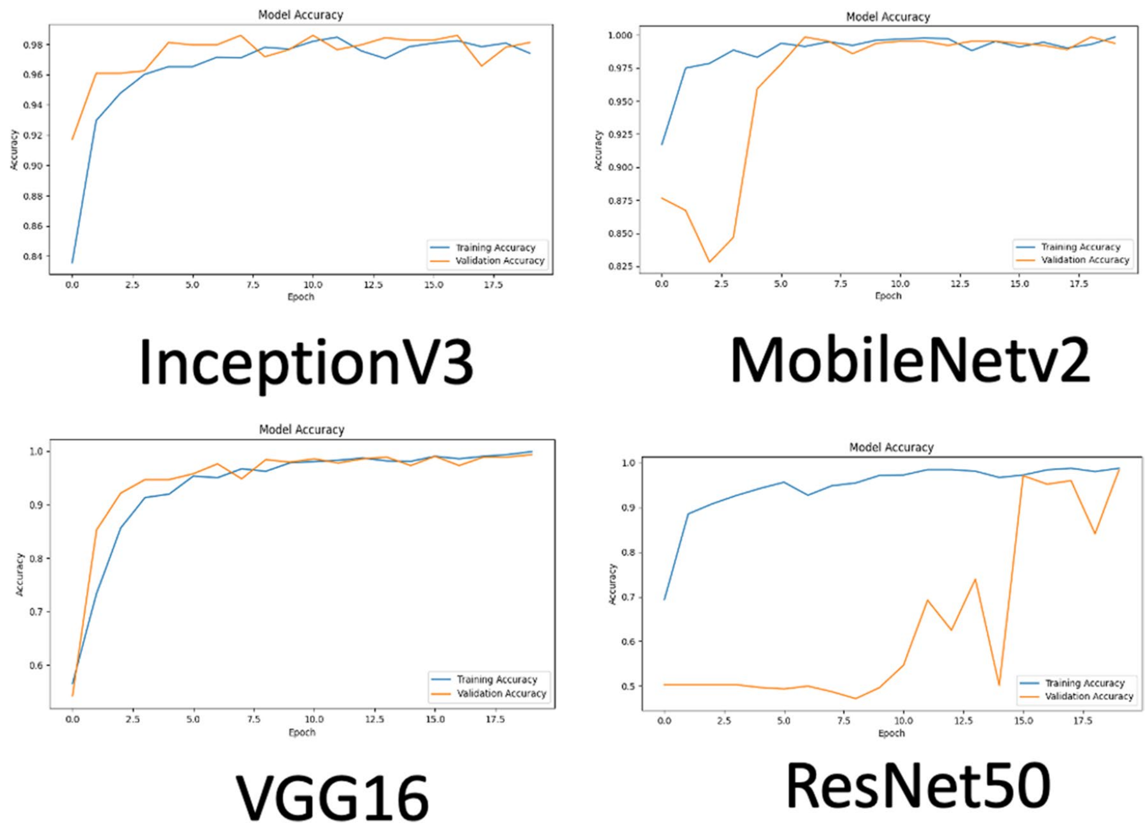

Fig. 4 Training curves of pre-trained models for binary classification of tumor vs no tumor

Table 4 Performance evaluation of models

Models	Accuracy (%)	Precision (%)	Recall (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)	MCC (%)	AUC (%)
MobileNetV2	99.94	99.88	100.00	100.00	99.88	99.94	99.88	100.00
InceptionV3	98.50	98.14	98.88	98.88	98.12	98.51	97.00	99.93
VGG16	99.87	99.88	99.88	99.88	99.88	99.88	99.75	99.99
ResNet50	98.25	98.86	97.62	97.62	99.88	98.24	96.51	99.92

performance evaluation of InceptionV3 on the testing set resulted in 98.50% accuracy, 98.14% precision, 98.88% recall, 98.88% sensitivity, 98.12% specificity, 98.51% F1 score, 97.00% MCC and 99.93% AUC, as shown in Table 4. While performance evaluation of VGG16 resulted in 99.87% accuracy, 99.88% precision, 99.88% recall, 99.88% sensitivity, 99.88% specificity, 99.88% F1-score, 99.75% MCC and 99.99% AUC. Lastly, evaluation of the performance of ResNet50 has shown that the model attained 98.25% accuracy, 98.86% precision, 97.62% recall, 97.62% sensitivity, 99.88% specificity, 98.24% F1-score, % MCC and 99.92% AUC as summarized in Table 4. The ROC/AUC is presented in Fig. 5.

Confusion Matrix Models Deployed for Binary Classification of MRI vs Non-MRI

The confusion matrices as shown in Fig. 6, showcase how many images the models got correct and ones they mislabeled. To test the models, 1600 cases (800 from each class) are used. The result shows that MobileNet accurately classify 799/800 MRI cases, 800/800 non-MRI cases. ResNet50 accurately classified 791/800 MRI cases and 781/800 non-MRI cases. While VGG16 accurately classified 799/800 MRI cases, 799/800 non-MRI cases. Finally, the InceptionV3 model accurately classified 785/800 MRI cases, 791/800 non-MRI cases.

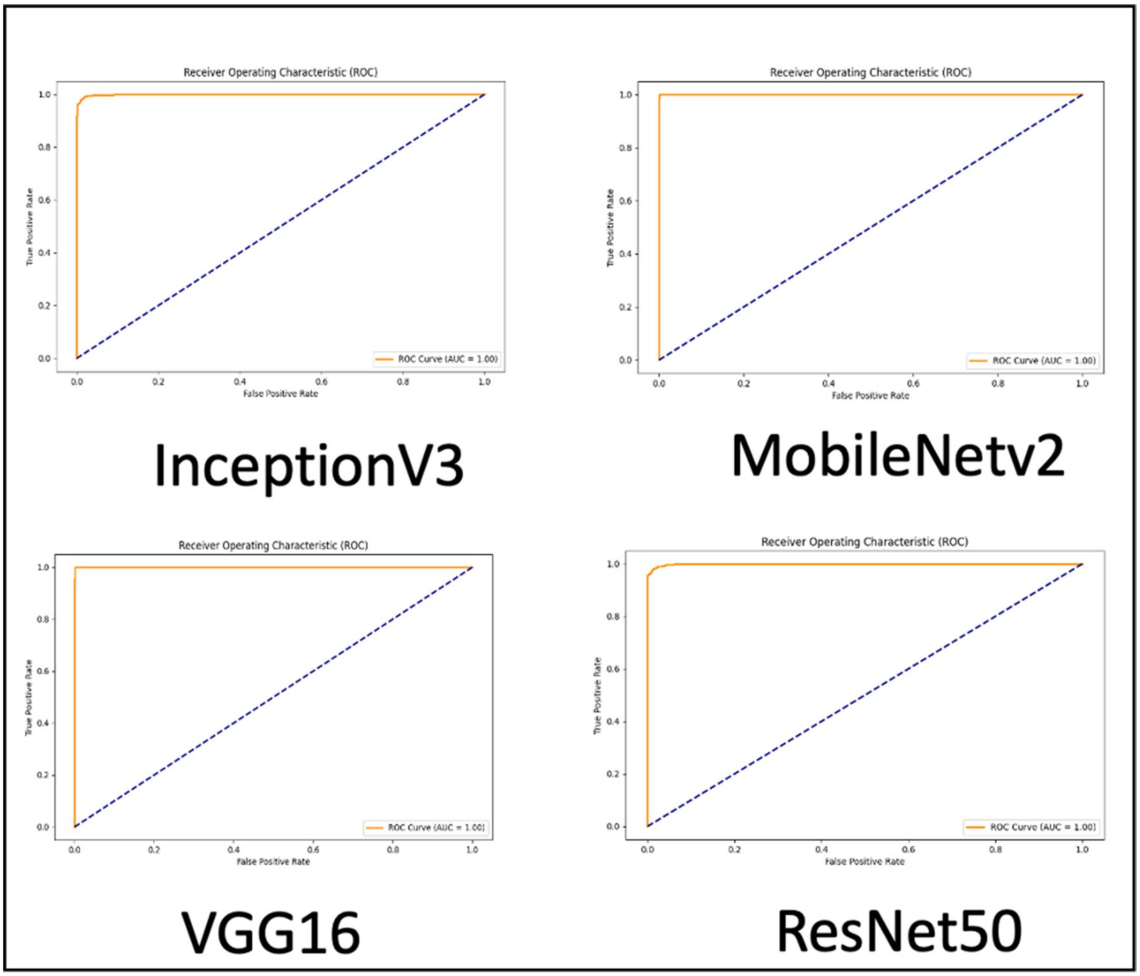

Fig. 5 ROC/AUC curves of pre-trained models for binary classification of MRI and non-MRI

Performance Evaluation of Models Deployed for Binary Classification of Tumor vs No tumor

The performance evaluation of MobileNet has shown that the model achieved 99.00% accuracy, 99.00% precision, 99.00% recall, 99.00% sensitivity, 99.00% specificity, 99.00% F1-score, 98.00% MCC and 99.88% AUC. The performance evaluation of InceptionV3 on the testing set resulted in 98.25% accuracy, 99.74% precision, 96.75% recall, 96.75% sensitivity, 99.75% specificity, 98.22% F1 score, 96.54% MCC and 99.66% AUC. While performance evaluation of VGG16 resulted in 98.62% accuracy, 99.49% precision, 97.75% recall, 97.75% sensitivity, 99.50% specificity, 98.61% F1-score, 97.26% MCC and % AUC. Lastly, evaluation of the performance of ResNet50 has shown that the model attained 98.37% accuracy, 98.01% precision, 98.75% recall, 98.75% sensitivity, 98.00% specificity, 98.38% F1-score, % MCC and 96.75% AUC as summarized in Table 5. The ROC/AUC is presented in Fig. 7.

Confusion Matrix for Binary Classification of Tumor vs No tumor

The confusion matrices as shown in Fig. 8, showcase how many images the models got correct and ones they mislabeled. To test the models, 800 cases (400 from each class) are used. The result shows that MobileNet accurately classify 396/400 MRI cases and 396/400 non-MRI cases. ResNet accurately classified 392/400 MRI cases and 395/400 non-MRI cases. While VGG16 accurately classified 398/400 MRI cases and 391/400 non-MRI cases. Finally, the InceptionV3 model accurately classified 399/400 MRI cases, and 387/400 non-MRI cases.

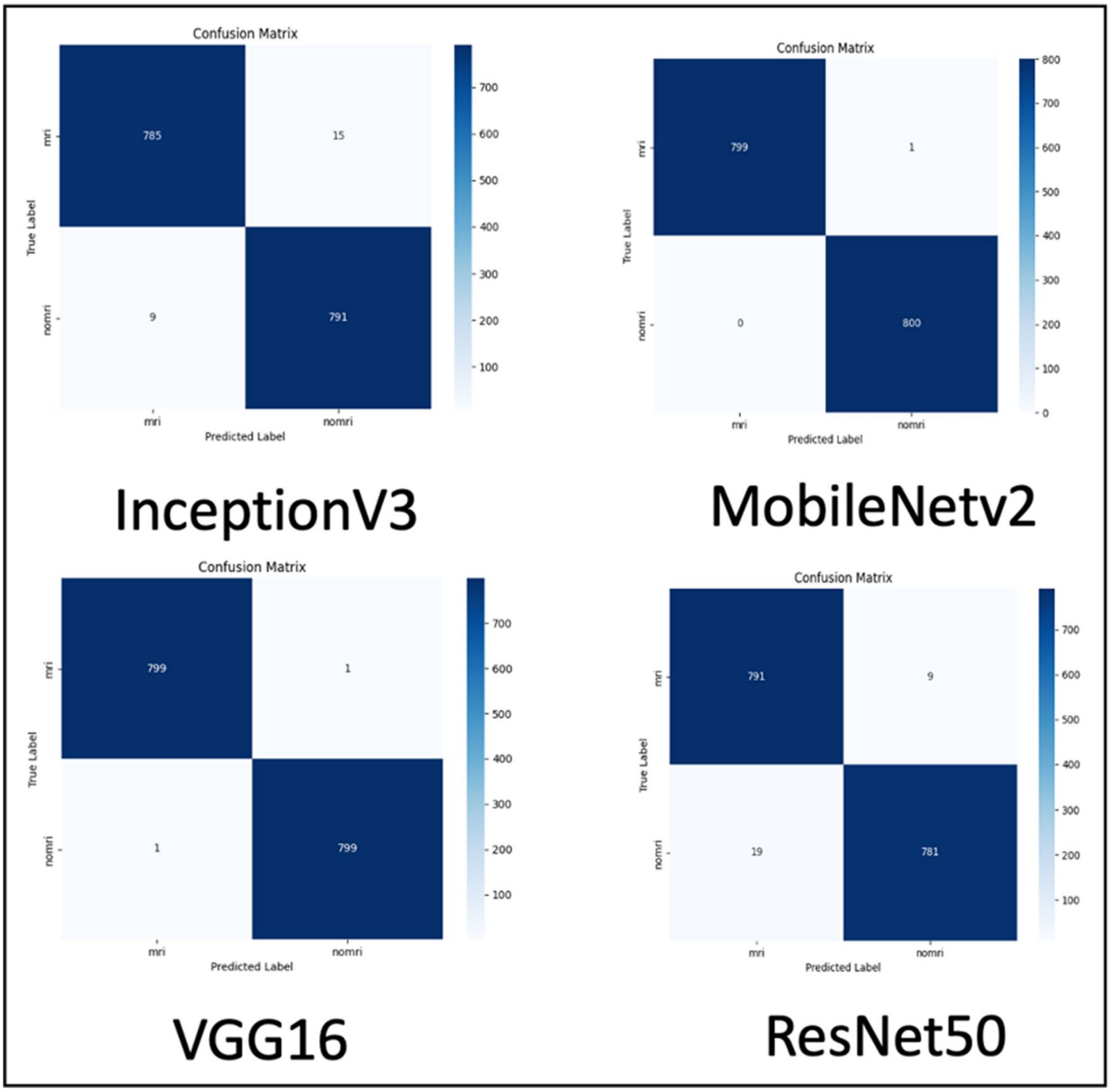

Fig. 6 Confusion Matrix for binary classification of MRI and non-MRI

Table 5 Performance evaluation of models

Models	Accuracy (%)	Precision (%)	Recall (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)	MCC (%)	AUC (%)
MobileNetV2	99.00	99.00	99.00	99.00	99.00	99.00	98.00	99.88
InceptionV3	98.25	99.74	96.75	96.75	99.75	98.22	96.54	99.66
VGG16	98.62	99.49	97.75	97.75	99.50	98.61	97.26	99.78
ResNet50	98.37	98.01	98.75	98.75	98.00	98.38	96.75	99.69

Discussion

Computer-aided detection (CAD) powered by AI-based techniques are transforming the field of medical imaging [36]. Several CAD platforms have been developed for the accurate classification of diseases. However, it was not until recently that scientists integrated IoT-based techniques in medical field (IoMT) for real-time detection of diseases [37, 38]. AI-based techniques aid neurologists, oncologist and radiologists to perform accurate diagnosis by offering deeper insights into the brain tumor pathology. Moreover, AI and DL approach support neuro-oncologists and neuroradiologists by leveraging medical images

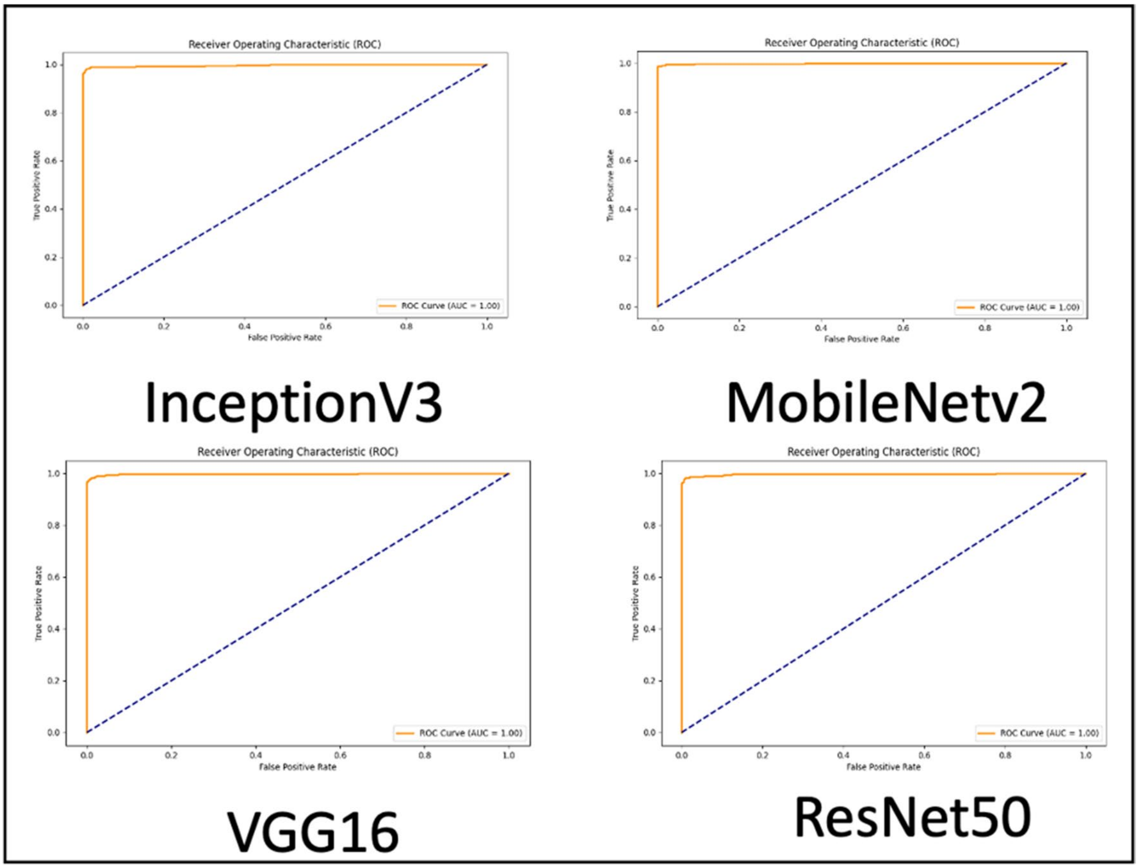

Fig. 7 ROC/AUC curves of pre-trained models for binary classification of MRI and non-MRI

such as MRI for accurate and automated classification, segmentation and detection of brain cancer. Due to the intricate nature and complexity of the brain, brain tumor is perceived as one of the most fatal type of cancer that is difficult to diagnose and treat. Healthcare experts rely on biopsy as general procedure for the screening of different type of cancer and tumors. However, due to invasive nature of this technique, scientists opted for MRI which is an invasive and affordable procedure. However, it has its own drawbacks which include inaccuracy in classifying different grade of brain tumor, high workload which is prone to error and miss-interpretation. To address these issues, scientists such as biomedical and computer scientists integrate DL-based approach which can achieve human or expert level, less error and faster.

The main aim of this study is to develope non-invasive AI-based technique for the discrimination between MRI and non-MRI cases and the detection and classification of MRI images into tumor no tumor. By retrieving the Brain tumor multimodal image dataset and the selection of 4000 MRI and 4000 CT scan (non-MRI), the images were processed based on resizing to fit the models input size requirement. The images are further split into training, validation and testing sets based on 64:16:20 ratio, while the training set is augmented to enlarge the dataset via 2 augmentation techniques. The images are trained and tested using 4 pre-trained models which include MobileNetV2, InceptionV3, ResNet50 and VGG16.

The comparative evaluations of the models for the discrimination between MRI and non-MRI as shown in Table 4 and Figs. 5 and 6 indicated that MobileNetV2 achieved the best scores across all metrics with 99.94% accuracy, 99.88% precision, 100.00% recall, 100.00% sensitivity, 99.88% specificity, 99.94% F1-Score, 99.88% MCC and 100.00% AUC. While, the comparative evaluations of the models for the discrimination between tumor and no tumor as shown in Table 5 and Figs. 7 and 8 indicated that MobileNetV2 achieved the best scores across all metrics with 99.00% accuracy, 99.00% precision, 99.00% recall, 99.00% sensitivity, 99.00% specificity, 99.00% F1-Score, 98.00% MCC and 99.88% AUC.

Web Development

As a way to improve portability, accessibility, and cost-effective solution for early detection of BC, scientists have

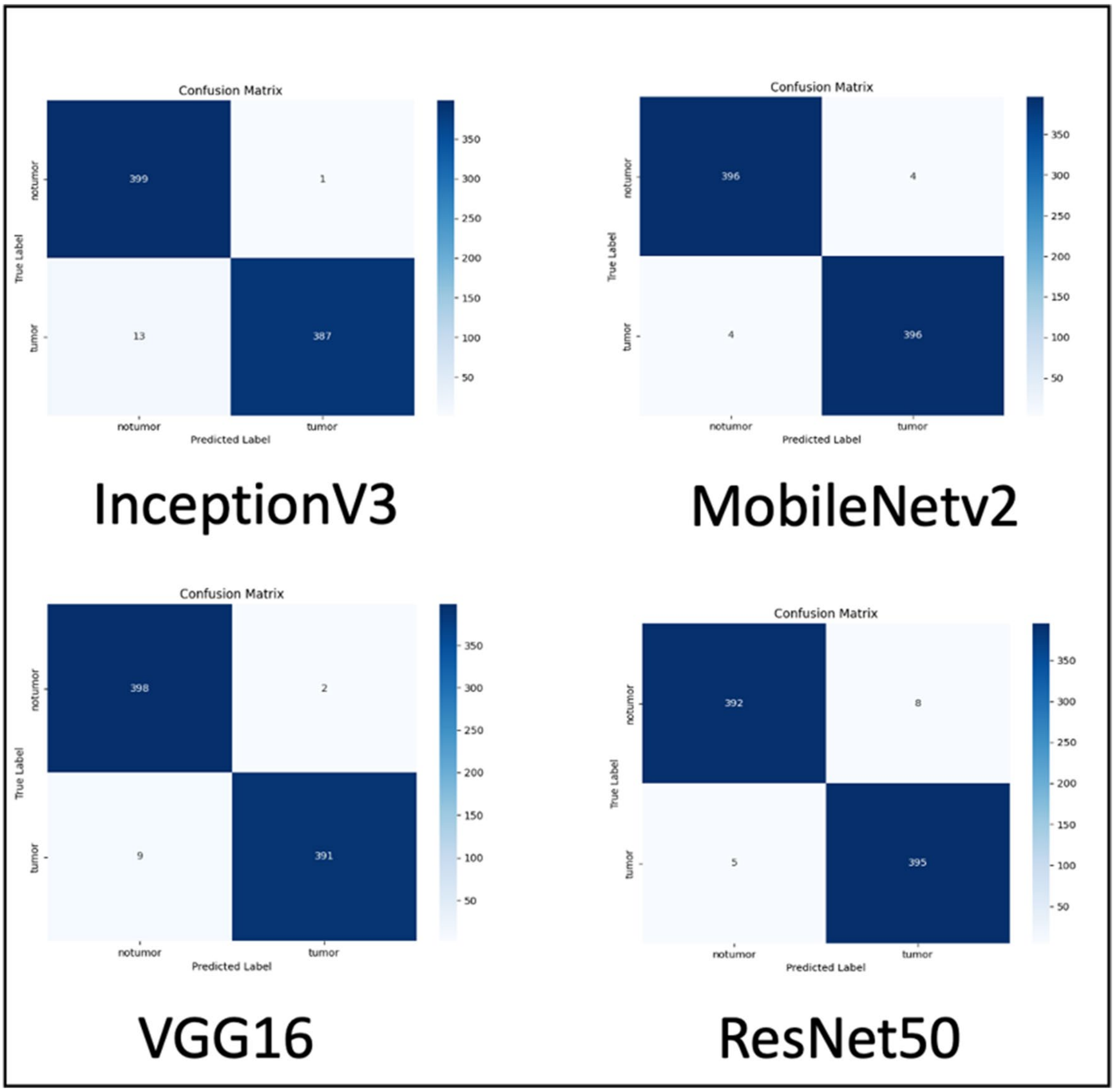

Fig. 8 Confusion Matrix for binary classification of tumor vs no tumor. **a** Inceptionv3. **b** MobileNet. **c** VGG16. **d** ResNet50

proposed the integration of AI with IoT [39–40]. In order to enable remote classification of brain tumors, we developed a user-friendly Stream lit application to help with the classification of the brain tumor images. The application uses the Tensor Flow Lite model for the real-time inference so that users can easily upload MRI images and get result in real-time. The process revolves around 3 main steps which include user log in, image upload and selection of MRI vs non-MRI followed by classification between tumor and no tumor. The first step allows users to log in using username and password. The second step involves uploading of MRI or non-MRI cases using a strong internet connection where the image can be pre-processed through resizing in order to fit the network input size requirement. The next step involves selection of classification and result in a matter of seconds. The overall process (from upload to classification) can be achieved in less than a minute. The step-by-step process is illustrated in Fig. 9and in a video submitted as a Supplementary file 1. The website is accessible via this link: https://ibrain.streamlit.app.

Conclusion

Brain tumor is considered as one of the deadliest forms of tumor due to its location and the vital function played by the brain. Thus, accurate screening of patients suspected of brain tumor is crucial for suitable treatment. In order to accurately diagnose patients, medical experts rely on 2 techniques which include biopsy and MRI. Despite the progress and transformation in medical diagnosis of brain tumor, there are still lingering challenges which limit

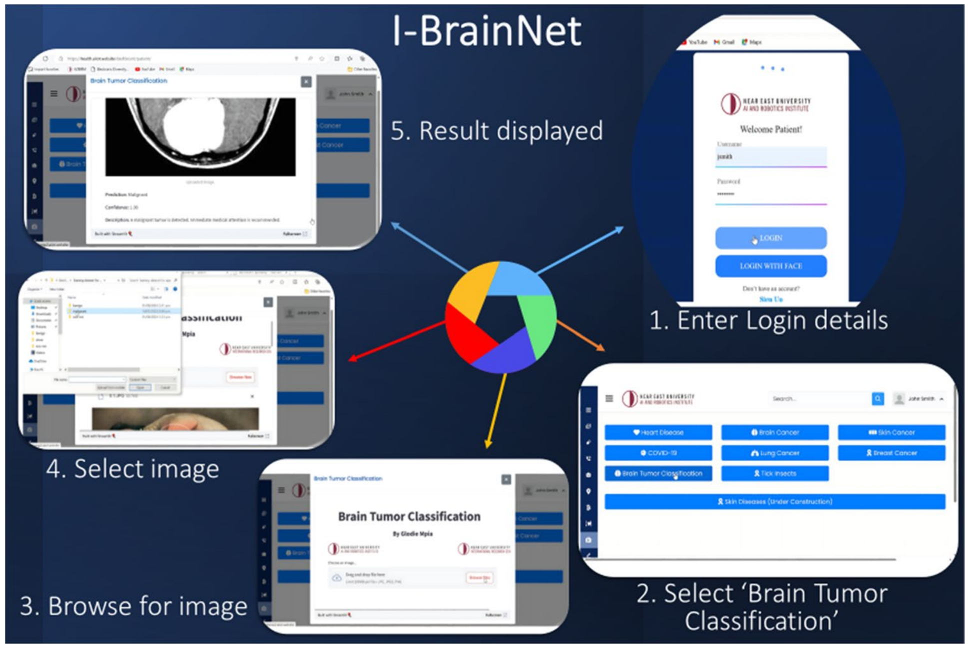

Fig. 9 The step-by-step process for real-time classification of brain tumors

accurate detection and classification of brain tumors. Majority of these limitations revolves around variations such as shape, size and location. Other technical issues include the high workload faced by radiologists and pathologists in terms of manual interpretation which is prone to error. This calls for the need to develop efficient, precise and automated techniques that can address these issues. One of the leading approaches employed by scientists to counter these limitations is the use of AI-based models which learn underlying features from training data (supervised) and attempt to classify unseen or reserve cases to ascertain the model performance.

Despite the fact that there are several studies that reported the development of customized CNN or implementation of pre-trained models, however, studies that proposed remote or real-time classification of tumors are very scarce. Therefore, in this study we proposed the integration of DL models and IoT-based framework for the accurate and real-time classification of brain tumors from MRIs. The experiment is set up based on data acquisition from online domain, data pre-processing, data split, data augmentation, training and testing of models and web development. Evaluation of the model performances and comparative analysis has shown that MobileNetv2 achieved the best result across all metrics for both discrimination between MRI and non-MRI cases as well as tumor and no tumor cases. A website is developed which integrates DL models for accurate and remote classification. Despite the high performance achieved using MobileNetv2, the study can be improved by acquiring more datasets, implementing ensemble learning by combining both DL models and classical ML classification algorithms such as SVM, DT, KNN, etc.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10278-025-01470-1.

Author Contribution All authors contributed to the study conception and design. Material preparation were performed by [Abdullahi Umar Ibrahim] and [Glodie Mpia Engo], Data collection and analysis were performed by [Ibrahim Ame] and [Chidi Wilson Newkwo], Supervision and Proofreading were performed by [Abdullahi Umar Ibrahim] and [Fadi Al-Turjman]. The first draft of the manuscript was written by [Abdullahi Umar Ibrahim] and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Declarations

Ethics Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication Not Applicable.

Competing Interests The authors declare no competing interests.

References

1. Khalighi S, Reddy K, Midya A, Pandav KB, Madabhushi A, Abedalthagafi M: Artificial intelligence in neuro-oncology: advances and challenges in brain tumor diagnosis, prognosis, and precision treatment. NPJ Precis Oncol, https://doi.org/10.1038/s41698-024-00575-0, 2024
2. National Brain Tumor Society (NBTS) (2024). Brain tumor fact. Available at https://braintumor.org/brain-tumors/about-brain-tumors/brain-tumor-facts/. Accessed on 19 September 2024.
3. Cancer Research UK, UK (2024). Brain, other CNS and intracranial tumours statistics. Available at https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/

brain-other-cns-and-intracranial-tumours. Accessed on 19 September 2024.
4. Alentorn A, Hoang-Xuan K, Mikkelsen T: Presenting signs and symptoms in brain tumors. Handbook of Clin Neurol, 134:1, 19–26, https://doi.org/10.1016/B978-0-12-802997-8.00002-5, 2016
5. Alther B, Mylius V, Weller M, Gantenbein AR: From first symptoms to diagnosis: initial clinical presentation of primary brain tumors. Clin Transl Neurosci, https://doi.org/10.1177/2514183X20968368, 2020
6. Van den Bent MJ, Weller M, Wen PY, Kros JM, Aldape K, Chang S: A clinical perspective on the 2016 WHO brain tumor classification and routine molecular diagnostics. Neuro-oncology 19, 614-624. https://doi.org/10.1093/neuonc/now277,2017
7. Banan R, and Hartmann C: The new WHO 2016 classification of brain tumors—what neurosurgeons need to know. Acta neurochirurgica, 159, 403-418https://doi.org/10.1007/s00701-016-3062-3,2017
8. Komori T: The 2021 WHO classification of tumors, central nervous system tumors: the 10 basic principles. Brain Tumor Pathol, 39, 47-50https://doi.org/10.1007/s10014-022-00428-3,2022
9. Butowski NA: Epidemiology and diagnosis of brain tumors. CONTINUUM: Lifelong Learning in Neurology, 21, 301–313, https://doi.org/10.1212/01.CON.0000464171.50638.fa, 2015
10. Abd-Ellah MK, Awad AI, Khalaf AA, Hamed HF: A review on brain tumor diagnosis from MRI images: practical implications, key achievements, and lessons learned. Magn Reson Imaging, 61, 300–318, https://doi.org/10.1016/j.mri.2019.05.028, 2019
11. Peddinti AS, Maloji S, Manepalli K: Evolution in diagnosis and detection of brain tumor–review. In J Physics: Conference Series, https://doi.org/10.1088/1742-6596/2115/1/012039, 2021
12. Shankar GM, Balaj L, Stott SL, Nahed B, Carter BS: Liquid biopsy for brain tumors. Expert Rev Mol Diagn, 17, 943–947, https://doi.org/10.1080/14737159.2017.1374854, 2017
13. Yi Z, Qu C, Zeng Y, Liu Z: Liquid biopsy: early and accurate diagnosis of brain tumor. J Cancer Res Clin Oncol, 148, 2347–2373, https://doi.org/10.1007/s00432-022-04011-3, 2022
14. Bonm AV, Ritterbusch R, Throckmorton P, Graber JJ: Clinical imaging for diagnostic challenges in the management of gliomas: a review. J Neuroimaging, 2, 139–145, https://doi.org/10.1111/jon.12687, 2020
15. Eibl RH, Schneemann M: Liquid biopsy and primary brain tumors. Cancers 13:21, 5429, https://doi.org/10.3390/cancers13215429, 2021
16. Nakhate V, Gonzalez Castro LN: Artificial intelligence in neuro-oncology. Front Neurosci, https://doi.org/10.3389/fnins.2023.1217629, 2023
17. Lamrani D, Cherradi B, El Gannour O, Bouqentar MA, Bahatti L: Brain tumor detection using MRI images and convolutional neural network. Int J Adv Comput Sci Appl, 13:7, https://doi.org/10.14569/IJACSA.2022.0130755, 2022
18. El-Feshawy SA, Saad W, Shokair M, Dessouky M: IoT framework for brain tumor detection based on optimized modified ResNet 18 (OMRES). J Supercomput, 79:1, 1081–1110, https://doi.org/10.1007/s11227-022-04678-y, 2023
19. Brindha PG, Kavinraj M, Manivasakam P, Prasanth P: Brain tumor detection from MRI images using deep learning techniques. In IOP conference series: materials science and engineering 1, 012115–131. IOP Publishing, https://doi.org/10.1088/1757-899X/1055/1/012115, 2021
20. Sawant A, Bhandari M, Yadav R, Yele R, Bendale MS: Brain cancer detection from MRI: a machine learning approach (tensorflow). Brain, 5, 2089-2094, 2018
21. Alsubai S, Khan HU, Alqahtani A, Sha M, Abbas S, Mohammad UG: Ensemble deep learning for brain tumor detection. Front Comput Neurosci, 16, 1005617–31, https://doi.org/10.3389/fncom.2022.1005617, 2022
22. Kolla M, Mishra RK, Zahoor Ul Huq S, Vijayalata Y, Gopalachari MV, Siddiquee KEA: CNN-based brain tumor detection model using local binary pattern and multilayered SVM classifier. Comput Intell Neurosci 1, 9015778–98, https://doi.org/10.1155/2022/9015778, 2022
23. Arora S, Sharma, M: Deep learning for brain tumor classification from MRI images. In 2021 Sixth International Conference on Image Information Processing (ICIIP) 6, 409–412. IEEE, https://doi.org/10.1109/ICIIP53038.2021.9702609, 2021
24. Khan MSI, Rahman A, Debnath T, Karim MR, Nasir MK, Band SS, Mosavi A, Dehzangi I: Accurate brain tumor detection using deep convolutional neural network. Comput Struct Biotechnol J, 20, 4733–4745, https://doi.org/10.1016/j.csbj.2022.08.039, 2022
25. Chauhan S, Cheruku R, Reddy Edla D, Kampa L, Nayak SR, Giri J, Mallik S, Aluvala S, Boddu V, Qin H: BT-CNN: a balanced binary tree architecture for classification of brain tumour using MRI imaging. Front Physiol, 15, 1349111, https://doi.org/10.3389/fphys.2024.1349111, 2024
26. He F, Liu T, Tao D: Why ResNet works? Residuals generalize. IEEE Transact Neural Net Learn Sys, 12, 5349–5362, https://doi.org/10.1109/TNNLS.2020.2966319, 2020
27. Wu Z, Shen C, Van Den Hengel A: Wider or deeper: revisiting the ResNet model for visual recognition. Pattern recognition, 90, 119-133, 2019
28. He K, Zhang X, Ren S, Sun J: Identity mappings in deep residual networks. In Computer Vision–ECCV: 14th European Conference, Amsterdam, Netherlands, Proceedings, 14, 630–645. Springer International Publishing, https://doi.org/10.1007/978-3-319-46493-0_38, 2016.
29. Simonyan K, Zisserman A: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014
30. Sengupta A, Ye Y, Wang R, Liu C, Roy K: Going deeper in spiking neural networks: VGG and residual architectures. Front Neurosci, 13, 95–112, https://doi.org/10.3389/fnins.2019.00095, 2019
31. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z: Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition 1, 2818-2826, 2016
32. 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, https://doi.org/10.1609/aaai.v31i1.11231, 2017
33. Alom MZ, Hasan M, Yakopcic C, Taha TM, and Asari VK: Inception recurrent convolutional neural network for object recognition. Machine Vision and Applications, 32, 1–14. https://doi.org/10.1007/s00138-020-01157-3, 2021
34. Howard AG: Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861, 2017
35. Sinha D, El-Sharkawy M: Thin mobilenet: an enhanced mobilenet architecture. In 2019 IEEE 10th annual ubiquitous computing, electronics & mobile communication conference (UEMCON), 0280–0285, https://doi.org/10.1109/UEMCON47517.2019.8993089, 2019
36. Umar Ibrahim A, Al-Turjman F, Ozsoz M, Serte S. Computer aided detection of tuberculosis using two classifiers. Biomed Eng/Biomedizinische Technik, 67, 513–24. https://doi.org/10.1515/bmt-2021-0310, 2022
37. Irkham I, Ibrahim AU, Nwekwo CW, Al-Turjman F, Hartati YW. Current technologies for detection of COVID-19: biosensors, artificial intelligence and internet of medical things (IOMT). Sens, 3, 426. https://doi.org/10.3390/s23010426, 2022

38. Manickam P, Mariappan SA, Murugesan SM, Hansda S, Kaushik A, Shinde R, Thipperudraswamy SP. Artificial intelligence (AI) and internet of medical things (IoMT) assisted biomedical systems for intelligent healthcare. Biosens, 12, 562–591. https://doi.org/10.3390/bios12080562, 2022
39. Ibrahim AU, Nwaneri IO, Vubangsi M, Al-Turjman F. I-Brainer: artificial intelligence/Internet of Things (AI/IoT)-powered detection of brain cancer. Curr Med Imaging. https://doi.org/10.2174/0115734056333393250117164020, 2025