

# Using The Combined Model Between MobileNetV2 and EfficientNetB0 to Classify Brain Tumors Based on MRI Images

Nhon Nguyen Thien, Kheo Chau Mui, Hoang-Tu Vo, Phuc Pham Tien, Huan Lam Le

Information Technology Department, FPT University, Can Tho 94000, Vietnam

---

## Article Info

### Article history:

Received Jan 1, 2025

Revised Jun 15, 2025

Accepted Jun 23, 2025

---

### Keywords:

Classification of Brain tumors  
improve the model  
fine-tuning  
image classification  
EfficientNetB0  
MobileNetV2  
Deep learning

---

## ABSTRACT

Brain tumors are extremely dangerous to one's health. If unchecked cell proliferation is not identified and treated promptly, it can lead to mortality, raise intracranial pressure, and endanger lifespan. To remove the tumor and lengthen the patient's life, early illness identification and drug administration are essential. In this research paper, we aim to improve the effectiveness of magnetic resonance imaging (MRI) equipment to identify cancerous brain tumour cells. It helps experts identify diseases faster. We classify brain tumour cells based on an image set of 3264 images with effective classification models such as ResNet50, InceptionV3, VGG19, EfficientNetB7, DenseNet201, MobileNetV2, Xception, etc. Besides, we also proposed two combined models: pooling (Xception + ResNet50) and pooling (MobileNetV2 + EfficientNetB0) to evaluate the effectiveness and found that the pooling model (MobileNetV2 + EfficientNetB0) gives the highest result, with 100% for the training set, 98% for the valid set, and 78% for the test set. We continued to improve the model by randomly re-dividing the data set with a Train-Valid-Test ratio of 60:20:20 and obtained an increased F1-score of 97%. We continued to improve the model again using the data augmentation techniques to create a larger data set, and the results far exceeded expectations with an F1-score of almost 100% for all classes. Based on the results, we found that combining MobileNetV2 with EfficientNetB0 is suitable for detecting brain tumour cancer cells. Aids in the early detection of dangerous cancers before they spread and endanger human health.

Copyright © 2025 Institute of Advanced Engineering and Science.  
All rights reserved.

---

## Corresponding Author:

Nhon Nguyen Thien

Information Technology Department, FPT University, Can Tho 94000, Vietnam.

Email: nhonnt9@fe.edu.vn

---

## 1. INTRODUCTION

One of the most serious conditions affecting the nervous system-particularly the brain-is a brain tumour. A brain tumour is defined as a mass of abnormal cells that grow uncontrollably in the brain region. Regardless of whether the tumour is benign or malignant, both types have the potential to exert pressure on and damage surrounding structures. Also known as intracerebral malignancies, brain tumours are characterized by the formation of abnormal cells within brain tissue [1], [2], [3], [4], [5], and [6]. Brain tumours are among the most serious conditions affecting the central nervous system. According to the **American Cancer Society (ACS)**, in 2019, approximately **23,820 new cases** of brain and other central nervous system (CNS) tumours were expected to be diagnosed in the United States, with an estimated **17,760 deaths** associated with these conditions [10]. Additionally, the **National Cancer Institute's SEER program** reported that in **2015**, approximately **166,039 people** in the U.S. were living with a brain or other CNS tumour [11].

By 2025, the ACS projects **24,820 new cases** of brain and CNS tumours in the U.S., with **14,040 male** and **10,780 female** patients, alongside **18,330 expected deaths** [12].

In the United Kingdom, data from the **Brain Tumour Charity** and **Cancer Research UK** indicate that there are approximately **12,000–12,700 new cases** of primary brain tumours diagnosed each year [13][14]. This figure includes both benign and malignant tumours.

These statistics reflect the growing burden of brain tumours globally, reinforcing the urgent need for advanced diagnostic methods, effective treatments, and improved patient care strategies.

Brain tumours are categorized into two main types: primary brain tumours and secondary (metastatic) brain tumours. A primary brain tumour originates in the brain and has not spread to other parts of the body. These tumours may be either benign or malignant. To assess the severity and implications of brain tumours and to promote standardisation in diagnosis, the World Health Organization (WHO) has developed a grading system.

There are more than 120 identified types of brain tumours, including meningiomas, epidermoid tumours, medulloblastomas, lymphomas, pituitary adenomas, gliomas, oligodendrogliomas, and glioblastoma multiforme [15], [16]. Among these, gliomas are the most prevalent, accounting for approximately 80% of all brain tumours [17].

The author of this study proposes an autonomous method for brain tumour identification and classification using saliency maps and deep learning-based feature optimization. The proposed framework is implemented in multiple phases. First, a fusion-based contrast enhancement technique is applied. This is followed by a tumour segmentation method based on saliency maps, which applies active contour mapping to the original image. Next, the pre-trained CNN model EfficientNetB0 is fine-tuned and trained using two datasets: one consisting of enhanced images and the other of tumour-localised images. Both models utilize deep transfer learning, and features are extracted from the average pooling layer.

These deep learning features are then fused using an advanced entropy-based fusion technique. In the final stage, an improved dragonfly optimization algorithm is employed to select the most relevant features. The classification is performed using an Extreme Learning Machine (ELM). The method was tested on three publicly available datasets and achieved improved accuracy rates of 95.14%, 94.89%, and 95.94%, respectively [8].

In this study, the dataset was processed in three phases according to the current experimental setup.

- **Experiments 01** involved evaluating the performance of efficient existing models for classification, including two hybrid models: MobileNetV2 + EfficientNetB0 and Xception + ResNet50.
- **Experiments 02** consisted of randomly splitting the dataset into three subsets for training after combining the original training and testing sets.
- **Experiments 03** involved augmenting the dataset by rotating all the original images before merging and splitting them into training, validation, and test sets. This augmentation aimed to increase dataset size and improve prediction accuracy.

Ultimately, the researchers chose to implement the MobileNetV2 + EfficientNetB0 approach as the primary model for this study

## 2. RELATED WORK

MRI pictures of cancer patients can be used to detect brain tumours using image processing techniques such as histogram equalisation, image segmentation, image enhancement, morphological procedures, and feature extraction. The Grey Level Co-occurrence Matrix (GLCM) has been utilised to derive textural information from the discovered tumour. The Knowledge Base's stored features and these features are compared. At last, a Neuro Fuzzy Classifier has been created to identify various forms of brain tumours. Two stages of testing have been conducted on the entire system: the Learning-Training Phase and the Recognition-Testing Phase. The system was trained using known MRI images of individuals with brain tumours that were collected from Tata Memorial Hospital's (TMH) Radiology Department. MRI pictures of the unidentified brain cancer samples that were impacted were also acquired from TMH and were utilised for system testing [34]. In this study [35], a combination of convolutional neural network (CNN) and sparse stacked auto encoder is described for cancer diagnosis based on brain magnetic resonance imaging (MRI) pictures. It is discovered that this combination significantly increases the categorization process's efficacy and accuracy. The MATLAB code for the suggested strategy is validated using a dataset of 120 MRI pictures. The outcomes demonstrated how well the suggested classifier classified and graded the MRI images of brain tumours. The goal of this work is to classify brain tumours into many groups using various machine learning techniques. To get better outcomes, a few preprocessing techniques were used. The outcome shows that feature selection has a significant impact on each method's overall performance with respect to accuracy overall and accuracy within each class. According to experimental findings, the Multilayer Perceptron (MP) approach outperforms other machine learning techniques in terms of accuracy rate [36]. The paper [37] While pattern recognition has been explored for classifying brain tumours in MRI images, grading remains challenging and often

subjective. This study proposes a deep learning approach using an inception-inspired model to classify MRI brain images. The architecture includes convolutional layers, max-pooling for feature extraction, and a fully connected layer for classification. A dataset of 200 MRI images from the REMBRANDT database is used. The model achieves a high accuracy of 98% with a learning rate of 0.01 over 20 training epochs, demonstrating its effectiveness in brain tumour classification. In this paper [38], Recently, Deep Learning (DL) has gained attention for its success in complex medical tasks. This study proposes a novel DL and Machine Learning (ML)-based method for detecting brain tumours in MRI images. Images are preprocessed using Adaptive Contrast Enhancement Algorithm (ACEA) and median filtering. Fuzzy c-means clustering segments the images, and features like energy, mean, entropy, and contrast are extracted using the Gray-Level Co-occurrence Matrix (GLCM). The proposed Ensemble Deep Neural Support Vector Machine (EDN-SVM) classifier then classifies the tissues. The method achieves high performance with 97.93% accuracy, 92% sensitivity, and 98% specificity in distinguishing abnormal from normal brain tissues.

In this research article [6], the goal of the author's research is to increase the ability of magnetic resonance imaging (MRI) equipment to detect malignant brain tumour cells. Five pre-trained convolutional neural network architecture models - ResNet152, DenseNet201, VGG19, MobileNetV2, and InceptionV3 were used to classify brain tumour cancer cells. After the work was analysed, the ResNet152 model was determined to have the greatest accuracy rate, with an accuracy value of 98.52%, or nearly 99%. Furthermore, the VGG19 model has a respectable 98% accuracy rate. InceptionV3 is 81.30%, DenseNet2 is 83.30%, and MobileNetV2 is 86.73% are the other models with accuracy values. This paper [9] investigates semantic information approaches in depth using a U-shaped network that multiplies a single reduced path into several pathways. The fundamental convolution layer applies artificial intelligence to extract multi-scale information from pictures using dilated convolutions and the Inception module. Lightweight efficient channel attention (ECA) modules are put into the bottleneck and decoder layer to increase network segmentation accuracy by emphasising segmentation-related information while disregarding redundant channel size information. When the suggested structure was put to the test using data from the 2018 Brain Tumour Segmentation Challenge (BraTS 2018), it was discovered that the core expanded in relation to the total volume. The mean Dice coefficients for the tumour and increased tumour locations are 88%, 78.4%, and 75.7%, respectively.

The Author [7] have applied six deep learning algorithms: InceptionV3, ResNet152V2, MobileNetV2, Resnet50, EfficientNetB0, and DenseNet201; these have been applied to two datasets related to brain tumours (both alone and in combination with human intervention) and one dataset related to Alzheimer's disease. With a total of 7.023 photos, 5.712 were used for training, and 1.311 were used for testing brain tumors. The accuracy of the tests was 98-99% and the training accuracy was 99-100%. With a total of 3.264 images, 2.870 for training and 394 for testing. And the second tumour dataset exhibits 100% training accuracy and 69-81% testing accuracy. With 10.000 pictures total 8.000 for training and 2.000 for testing the combined dataset yields 99-100% training accuracy and 98-99% testing accuracy. The Alzheimer's dataset, which consists of 6.400 pictures total 5.121 for training, 1.279 for testing, and 4 classes of images has testing accuracy of 71-78% and training accuracy of 99-100%. Employed pre-trained deep learning models in inflammatory, degenerative, neoplastic, and cerebrovascular tissues [18]. [19]MRI is a common diagnostic tool, but manual analysis is time-consuming and often misses subtle abnormalities. Deep learning has emerged as a powerful approach for medical image analysis. This study uses five pre-trained models—AlexNet, VGG-16, ResNet-18, ResNet-34, and ResNet-50—to automatically classify MRI images into normal, cerebrovascular, neoplastic, degenerative, and inflammatory categories. Among them, ResNet-50 achieved the highest accuracy of  $95.23\% \pm 0.6$ .

Residual networks (ResNets) are easier to optimize and benefit from greater depth. On ImageNet, a 152-layer ResNet—eight times deeper than VGG—achieves 3.57% test error and wins 1st place at ILSVRC 2015. Tests on CIFAR-10 with up to 1000 layers further support the method's effectiveness. ResNets also yield a 28% improvement in COCO object detection [20], [21], [22], [23] The design separates input/output domains from transformation expressiveness, enabling flexible analysis. MobileNetV2 is applied to object detection via **SSDLite** and semantic segmentation through a simplified **DeepLabv3**. Performance is evaluated on ImageNet, COCO, and PASCAL VOC, analyzing trade-offs between accuracy, latency, MAdds, and parameter count. On ImageNet, a 152-layer ResNet—eight times deeper than VGG—achieves 3.57% test error and wins 1st place at ILSVRC 2015. Tests on CIFAR-10 with up to 1000 layers further support the method's effectiveness. ResNets also yield a 28% improvement in COCO object detection. Applying this to MobileNets and ResNet improves performance. Using neural architecture search, the authors design a new baseline and scale it into the **EfficientNet** family. EfficientNet-B7 achieves 84.4% top-1 and 97.1% top-5 accuracy on ImageNet, while being 8.4× smaller and 6.1× faster than previous top models. EfficientNets also perform well on CIFAR-100, Flowers, and other transfer learning tasks with fewer parameters. [24] This study compares the performance of four machine learning algorithms—Support Vector Machine (SVM), Decision Tree (C4.5), Naive Bayes (NB), and k-Nearest Neighbors (k-NN)—using the Wisconsin Breast Cancer

(original) dataset. The goal is to evaluate each method's accuracy, precision, sensitivity, and specificity. Results show that SVM achieves the highest accuracy (97.13%) and lowest error rate. The authors employed VGG-16, Alex-Net, ResNet-34, ResNet-18, and ResNet-50 as MR image classification techniques. The characteristics are extracted and categorised using this automated process. For 1,074 MRI images, data were gathered from the Harvard Medical School dataset. Utilised data came from MRI 1074 imaging at Harvard Medical School. Tested on big MRI pictures of brain tumours, the suggested approach produced the greatest results. The ResNet-50 model's accuracy is 95.33%, according to the suggested method's findings. In these studies, the authors proposed a new KNN model to meet the need for early diagnosis of disease diagnostic processes. These models can be useful tools for clinicians, helping them effectively classify whether a cancer is benign or malignant. The main goal of their study was comparison results of supervised learning and classification algorithms to combine these algorithms using classification techniques called voting. And it is a group method because they Multiple models can be combined to achieve a higher classification accuracy. The datasets were collected from the university Wisconsin. [25] reached 98.90%, [26] achieved accuracy was 97.60%, [27] reached 97.13%, [28] reached 99.9%, [29] reached 98.10% [30] reached 98.23%, and [31] reached 83.45%. Two recent studies have applied deep learning techniques to improve image classification in agriculture and botany. Vo et al. (2023) proposed a tomato leaf disease recognition model by combining the Xception network with Bilinear Pooling, significantly enhancing accuracy compared to traditional methods. Meanwhile, Cengil and Çınar (2019) employed transfer learning to classify flower images into multiple categories, achieving strong results by leveraging pre-trained models. Both studies highlight the great potential of deep learning in analyzing natural images for agricultural and biological applications [32], [33].

Despite the extensive application of deep learning models for brain tumor classification, most existing studies utilize individual models such as ResNet50, VGG19, or InceptionV3 and rely on small datasets with limited augmentation. Moreover, while bilinear pooling and multi-model fusion have been explored in fields such as agriculture or natural image classification [32], [33], their potential in medical imaging, particularly brain MRI analysis, remains under-investigated.

This study addresses these gaps by proposing a novel integration of **EfficientNetB0** and **MobileNetV2** with **bilinear pooling**, coupled with extensive **data augmentation** to enhance classification accuracy. To the best of our knowledge, no prior work has systematically combined these models for brain tumor detection and achieved nearly perfect classification accuracy on augmented medical datasets.

### 3. MATERIALS AND METHODS

#### 3.1. Data Collection and Preparation

The study used a comprehensive dataset consisting of 3264 images from 4 different categories in the first stage, meticulously divided for different purposes. Out of all these, 2870 photos are used for training and validation, remaining 394 photos for testing (Link: <https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri>). To expand the image dataset size, the study employs image enhancement techniques within the ImageDataGenerator. This process involves multiple transformations applied to images, aiming to enhance their quality and enrich the dataset for improved machine learning model training. It initiates by enabling image rotation up to 15 degrees and slight shifts in width and height, introducing variability. Standardizing pixel values to a range between 0 and 1 through rescaling ensures consistent processing. Shearing introduces slight tilts, and adjusting brightness within a specified range further diversifies their appearance. Additionally, horizontal and vertical flipping creates mirrored images, adding variations for a more robust dataset. Applying the above techniques the study create a new dataset with 74,913 images and all images are uniformly processed at 224 x 224 pixels, ensuring consistency. The-se images are formatted in JPG format. The available dataset is shown in Fig. 1, and after classifying the images, there are 4 main categories: pituitary\_tumor, glio-ma\_tumor, meningioma\_tumor and no\_tumor which are shown in Fig. 2. Overall, these techniques elevate the dataset's diversity and quality, offering a more extensive array of images for training models to better recognize patterns and features in the data.

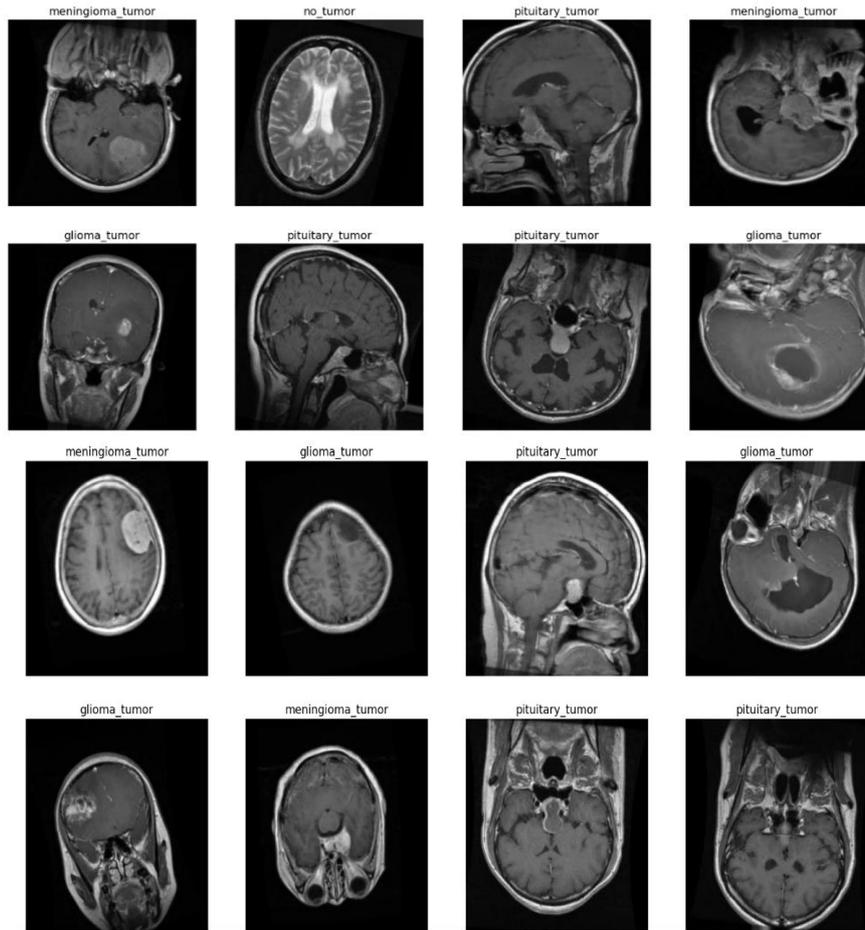


Figure 1. Some images in the dataset.

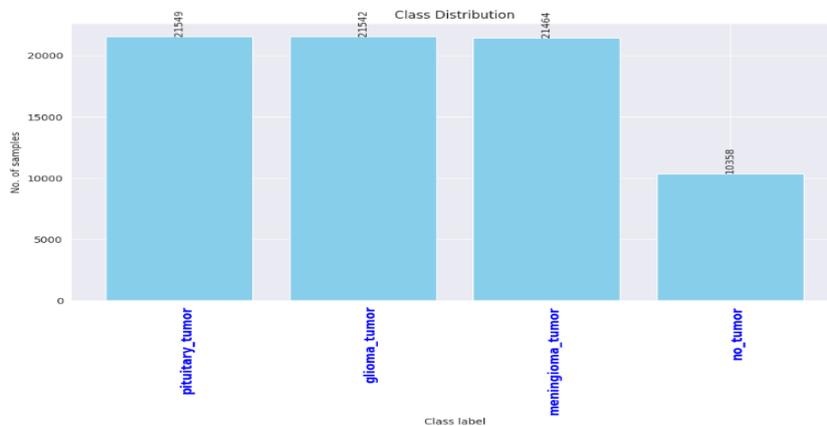


Figure 2. A dataset distribution.

Before augmentation, the dataset was composed of 3264 MRI images distributed across four classes as follows:

- **Pituitary tumor:** 926 images
- **Meningioma tumor:** 708 images
- **Glioma tumor:** 1,426 images
- **No tumor:** 204 images

It is evident that the "no\_tumor" class had significantly fewer samples compared to the tumor classes - in some cases less than one-fourth. This class imbalance poses a risk of model bias, where the classifier may perform poorly in identifying non-tumor cases.

To address this, we applied **image augmentation** techniques using ImageDataGenerator, including rotation, shearing, brightness adjustment, and horizontal/vertical flipping. These transformations were applied more extensively to the underrepresented "**no\_tumor**" class to rebalance the class distribution.

After augmentation, the dataset expanded to **74913 images**, with the following approximate class distributions:

- **Pituitary tumor**: ~21,600 images
- **Meningioma tumor**: ~17,000 images
- **Glioma tumor**: ~27,000 images
- **No tumor**: ~9,300 images

While the class distribution remains imbalanced, the augmentation process significantly increased the representation of the "no\_tumor" class (from 204 to over 9,000 samples), helping to improve classification performance and reduce false positives. Additionally, stratified splitting during training ensured balanced exposure to all classes in each batch.

### 3.2. Development of the Integrated Model

In this section, we employed a hybrid feature extraction approach by combining two efficient deep learning models: EfficientNetB0 and MobileNetV2. EfficientNetB0 employs compound scaling to optimize both depth and width of the network while maintaining efficiency, making it suitable for extracting global contextual features. In contrast, MobileNetV2 leverages inverted residuals and depthwise separable convolutions, making it particularly effective in capturing fine-grained spatial details with low computational cost [39][40].

The decision to combine these two models stems from their complementary strengths in feature representation. While EfficientNetB0 is well-suited for generalizing high-level semantic patterns, MobileNetV2 effectively captures localized discriminative features. By integrating their outputs, we aim to obtain a richer and more robust feature embedding. Deep features from the penultimate layers of both models were concatenated into a unified feature vector, which was subsequently processed through dense layers for classification.

Although prior works rarely explore this specific combination, a theoretical rationale exists in their architectural diversity and empirical performance on individual tasks. To justify this fusion approach, we performed an ablation analysis comparing different model combinations (Section 4.2.1), where the EfficientNetB0 + MobileNetV2 combination outperformed both individual and alternative ensemble models in terms of validation accuracy and F1-score, particularly under optimized training setups.

### 3.3. Proposed Model

This study proposes a model designed to identify brain cancer through image analysis. Using TensorFlow's Keras library, it constructs a specialized structure to process image data and predict the presence of brain cancer. The model's architecture combines two pre-trained models and custom layers to distill complex patterns from input images. The proposed model begins by establishing the foundational structure, defining the size and shape of the images for analysis, and setting up the input layer to receive this image data. It integrates two separate pre-trained models, MobileNetV2 and EfficientNetB0, as specialized layers. These models, trained on extensive datasets, enable the model to benefit from their knowledge and insights when analyzing input images. To synthesize information from these pre-trained models, the proposed model introduces a custom layer called Bilinear Pooling. This layer uses matrix outer products to combine and distill essential features from both pre-trained models. After the pooling process, the extracted features undergo further structuring and processing through additional layers. These layers involve operations such as flattening, batch normalization, and dense connections. The output layer utilizes a softmax activation function to generate probability distributions across various brain illnesses. Upon training with relevant data, this model shows potential for accurately classifying brain images, significantly contributing to medical diagnostics and research Figure 3.

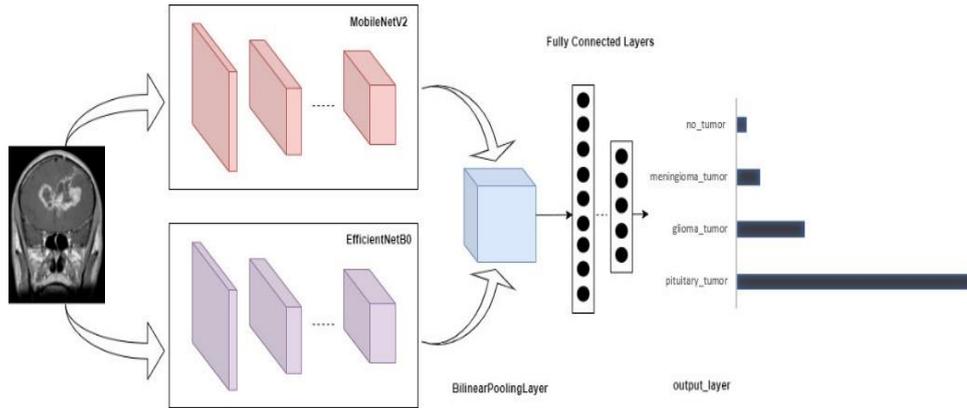


Figure 3. Proposed model for Brain tumors classification.

From the combination of two models MobileNetV2 and EfficientNetB0, the total parameters extracted are 6,641,639 parameters. Of these, 6,562,944 parameters were used to train the model and 78,695 parameters were not used for training Figure 4.

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	[(None, 224, 224, 3)]	0	[]
mobile_net_v2_layer (MobileNetV2Layer)	(None, 7, 7, 1280)	2257984	['input_1[0][0]']
efficient_net_b0_layer (EfficientNetB0Layer)	(None, 7, 7, 1280)	4049571	['input_1[0][0]']
bilinear_pooling_layer (BilinearPoolingLayer)	(None, 1, 1, 1280)	0	['mobile_net_v2_layer[0][0]', 'efficient_net_b0_layer[0][0]']
flatten (Flatten)	(None, 1280)	0	['bilinear_pooling_layer[0][0]']
batch_normalization (Batch Normalization)	(None, 1280)	5120	['flatten[0][0]']
dense (Dense)	(None, 256)	327936	['batch_normalization[0][0]']
dropout (Dropout)	(None, 256)	0	['dense[0][0]']
dense_1 (Dense)	(None, 4)	1028	['dropout[0][0]']

=====  
Total params: 6641639 (25.34 MB)  
Trainable params: 6562944 (25.04 MB)  
Non-trainable params: 78695 (307.41 KB)  
=====

Figure 4. Parameters for Proposed Model

### 3.4. Performance Evaluation Measures

The performance metrics used for brain tumor analysis include accuracy, specificity, loss, recall, precision, and F1 score. These metrics are widely employed in disease diagnosis for various models and methods performance evaluation

The most popular measurements were used in the performance of the proposed model: Accuracy is represented in (1), Precision is represented in (2), Recall is represented in (3) and F1-score is represented in (4).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recal = \frac{TP}{TP+FN} \quad (3)$$

$$F_1 - Score = \frac{Precision*Recall}{Precision+Recall} \quad (4)$$

In equations (1)-(4), the variables TP (True Positive), TN (True Negative), FP (False Positive), and FN (False Negative) are used to define performance metrics in binary and multi-class classification. Specifically:

- **TP** denotes the number of correctly predicted positive instances,
- **TN** refers to the correctly predicted negative instances,
- **FP** represents negative instances incorrectly labeled as positive,
- **FN** indicates positive instances incorrectly labeled as negative.

Although not explicitly substituted in each formula, they are the underlying counts that inform each metric. Including this clarification helps interpret the confusion matrix results in Section 4.

## 4. RESULTS AND DISCUSSION

Compared to prior studies [6][7][18], which achieved F1-scores ranging from 94-97% on similar datasets, our method surpasses these benchmarks, especially after augmentation, achieving nearly 100% accuracy. The improved performance can be attributed to bilinear pooling, model fusion, and extensive preprocessing.

### 4.1. Environmental settings

The conducted experiments yielded results through implementation on the Kaggle platform. The experimental system was equipped with 13GB of RAM and a GPU Tesla P100-PCIE featuring 16GB of memory. Model training extended across 30 epochs, employing a batch size of 16 throughout the training process.

### 4.2. Experiments

In this research paper, we aim to improve the effectiveness of magnetic resonance imaging (MRI) equipment to identify cancerous brain tumour cells. It helps experts identify diseases faster. We classify brain tumour cells based on images using three main experiment:

#### 4.2.1 Experiments 01:

Divide the train set into 2 sets at a ratio of 75:25 with the new train set (2152 images) and Valid (718 images) to train the model.

Run tests with available models such as ResNet50, ResNet50V2, InceptionV3, InceptionResNetV2, VGG16, VGG19, EfficientNetB7, EfficientNetV2, DenseNet201, MobileNet, MobileNetV2, NasNetMobile, Xception, RegNetX002, and 2 combined pooling models (Xception + ResNet50) and Pooling (MobileNetV2 + EfficientNetB0) (Table 1).

Table 1: Performance comparison of 16 deep learning models and two hybrid combinations on the initial dataset split Experiment 1:

No.	Model	Train Accuracy	Valid Accuracy	Test Accuracy	F1-score
1	Pooling (MobileNetV2 + EfficientNetB0)	0.9976	0.9805	0.7791	78%
2	Pooling (Xception + ResNet50)	1.0000	0.9846	0.7614	76%
3	ResNet50	0.9646	0.8885	0.7157	72%
4	EfficientNetB7	0.9061	0.8662	0.6776	68%
5	RegNetX002	0.8838	0.8537	0.6497	65%
6	EfficientNetV2	0.9079	0.8927	0.6395	64%
7	MobileNet	0.9010	0.8440	0.5888	59%
8	Xception	0.6405	0.6142	0.5355	54%
9	InceptionV3	0.6770	0.6323	0.5304	53%
10	MobileNetV2	0.8582	0.8440	0.5304	53%
11	DenseNet201	0.8164	0.8050	0.5177	52%
12	NasNetMobile	0.6491	0.6448	0.4187	42%
13	ResNet50V2	0.5418	0.5626	0.3147	31%
14	VGG19	0.8066	0.7757	0.2664	27%
15	VGG16	0.8396	0.7924	0.2538	25%
16	InceptionResNetV2	0.3889	0.3593	0.2258	23%

However, the results show that the prediction rate of the test set, or F1-score, is much lower than the other two sets. The team examined the dataset and found that some images in the test set were too different

from the other two sets, so the team decided to process the dataset to improve the model and chose the combined model of MobileNetV2 + EfficientNetB0 for the stage. Phase 1 is an improved model for phase 2 because it brings the highest F1-score results, and the execution time of MobileNetV2+EfficientNetB0 is also faster. The result was that the F1-score prediction rate was only 78%, not reaching the desired result as shown in the Figure.5

	precision	recall	f1-score	support
glioma_tumor	1.00	0.31	0.47	100
meningioma_tumor	0.75	0.96	0.84	115
no_tumor	0.68	1.00	0.81	105
pituitary_tumor	0.97	0.82	0.89	74
accuracy			0.78	394
macro avg	0.85	0.77	0.75	394
weighted avg	0.84	0.78	0.75	394

Figure 5. F1-score comparison across models during Experiment 1. The proposed MobileNetV2 + EfficientNetB0 combination outperforms most individual models.

#### 4.2.2 Experiments 02:

The team processes the dataset by combining two training and test sets from the original source and then randomly dividing them into three valid test sets in a ratio of 60:20:20 for use in training and prediction. As a result, the F1-score prediction rate increased to 97% in shows Figure 6. While Experiment 1 provided a comparative analysis of 16 models and 2 hybrid approaches, Experiments 2 and 3 focused solely on the MobileNetV2 + EfficientNetB0 combination. To justify the improvements, we extended our evaluation by applying other top-performing models (e.g., ResNet50, Xception) under the same data conditions (random split and augmentation). Results show that while these models improved moderately, none reached the near-perfect F1-score achieved by our proposed combination, indicating that the hybrid model particularly benefits from augmented data and rebalancing strategies.

	precision	recall	f1-score	support
glioma_tumor	0.98	0.94	0.96	185
meningioma_tumor	0.93	0.98	0.95	188
no_tumor	0.96	0.99	0.97	79
pituitary_tumor	1.00	0.97	0.99	180
accuracy			0.97	632
macro avg	0.97	0.97	0.97	632
weighted avg	0.97	0.97	0.97	632

Figure 6. F1-score comparison across models during Experiment 2. The proposed MobileNetV2 + EfficientNetB0 combination outperforms most individual models.

#### 4.2.3 Experiments 03:

To test the effectiveness of redistributing the dataset, the team created a larger dataset with data augmentation techniques. From the original 3,264 images, create a new dataset with 74,913 images and then the team proceeds to gather all the images and divide them randomly. The data set is divided into three subsets including: Train, Valid and Test with a ratio of 60:20:20 as in phase 2. Finally, the third experiment took place for nearly 5 hours and yielded results as shown in Figure. 7:

Epoch	Loss	Accuracy	V_loss	V_acc	LR	Next LR	Monitor	% Improv	Duration
1 /30	1.972	93.748	0.26088	99.446	0.00100	0.00100	val_loss	0.00	637.13
2 /30	0.198	98.832	0.12465	99.800	0.00100	0.00100	val_loss	52.22	567.19
3 /30	0.122	99.419	0.08718	99.900	0.00100	0.00100	val_loss	30.06	567.61
4 /30	0.094	99.682	0.07364	99.920	0.00100	0.00100	val_loss	15.53	566.14
5 /30	0.082	99.715	0.06754	99.907	0.00100	0.00100	val_loss	8.28	565.18
6 /30	0.072	99.806	0.05816	99.907	0.00100	0.00100	val_loss	13.89	562.75
7 /30	0.066	99.809	0.05044	99.920	0.00100	0.00100	val_loss	13.27	562.71
8 /30	0.060	99.849	0.05140	99.907	0.00100	0.00050	val_loss	-1.91	573.56
9 /30	0.051	99.935	0.04113	99.947	0.00050	0.00050	val_loss	18.46	573.87
10 /30	0.045	99.964	0.03786	99.940	0.00050	0.00050	val_loss	7.95	572.65
11 /30	0.042	99.971	0.03676	99.940	0.00050	0.00050	val_loss	2.90	570.11
12 /30	0.040	99.976	0.03508	99.940	0.00050	0.00050	val_loss	4.56	573.24
13 /30	0.038	99.967	0.03344	99.933	0.00050	0.00050	val_loss	4.69	569.70
14 /30	0.037	99.973	0.03258	99.960	0.00050	0.00050	val_loss	2.58	573.14
15 /30	0.035	99.987	0.03065	99.947	0.00050	0.00050	val_loss	5.92	574.93
16 /30	0.035	99.958	0.03060	99.933	0.00050	0.00050	val_loss	0.17	572.85
17 /30	0.033	99.989	0.03021	99.920	0.00050	0.00050	val_loss	1.26	573.86
18 /30	0.031	99.993	0.02890	99.933	0.00050	0.00050	val_loss	4.35	572.90
19 /30	0.031	99.976	0.02808	99.933	0.00050	0.00050	val_loss	2.82	573.22
20 /30	0.030	99.993	0.02736	99.940	0.00050	0.00050	val_loss	2.58	573.54
21 /30	0.029	99.989	0.02675	99.933	0.00050	0.00050	val_loss	2.23	572.78
22 /30	0.029	99.980	0.02644	99.927	0.00050	0.00050	val_loss	1.16	567.98
23 /30	0.027	99.984	0.02593	99.940	0.00050	0.00050	val_loss	1.90	572.91
24 /30	0.027	99.984	0.02574	99.933	0.00050	0.00050	val_loss	0.75	573.51
25 /30	0.026	99.993	0.02492	99.940	0.00050	0.00050	val_loss	3.19	573.21
26 /30	0.025	99.993	0.02477	99.933	0.00050	0.00050	val_loss	0.58	573.12
27 /30	0.025	99.989	0.02398	99.940	0.00050	0.00050	val_loss	3.21	569.34
28 /30	0.024	99.989	0.02249	99.940	0.00050	0.00050	val_loss	6.18	568.85
29 /30	0.023	99.991	0.02278	99.933	0.00050	0.00025	val_loss	-1.28	567.50
30 /30	0.022	99.993	0.02128	99.933	0.00025	0.00025	val_loss	5.38	567.49

training elapsed time was 4.0 hours, 46.0 minutes, 27.80 seconds)

Figure 7. Training results

From the model fit, the test data was used to predict using the model and make a comparison between the expected output results and the given prediction, a summary through the confusion matrix is shown in Figure 8. Additionally, Figure 9 and Figure 10 illustrates the performance metrics, such as loss and accuracy, that were evaluated during both the model’s training and validation stages.

The model achieves the highest accuracy at the 14th epoch and has little fluctuation until the last epoch. Besides, the loss representation always decreases and is lowest at the 30th last epoch.

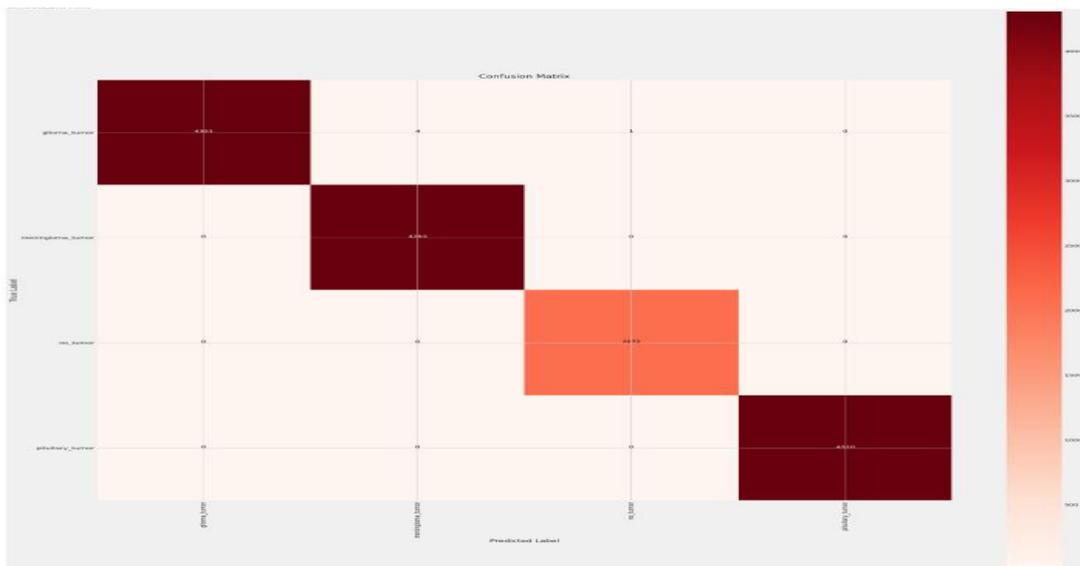


Figure 8. Confusion matrix of fine-tuned MobileNetV2 + EfficientNetB0 model.

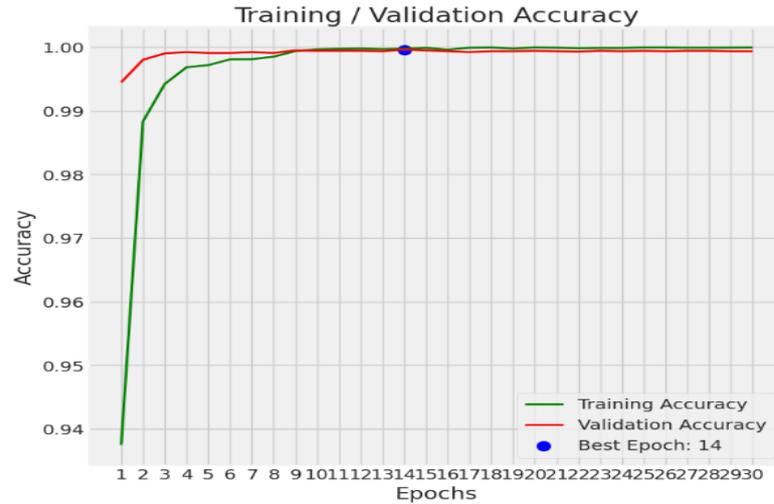


Figure 9. Training and Validation accuracy plot of the suggested model

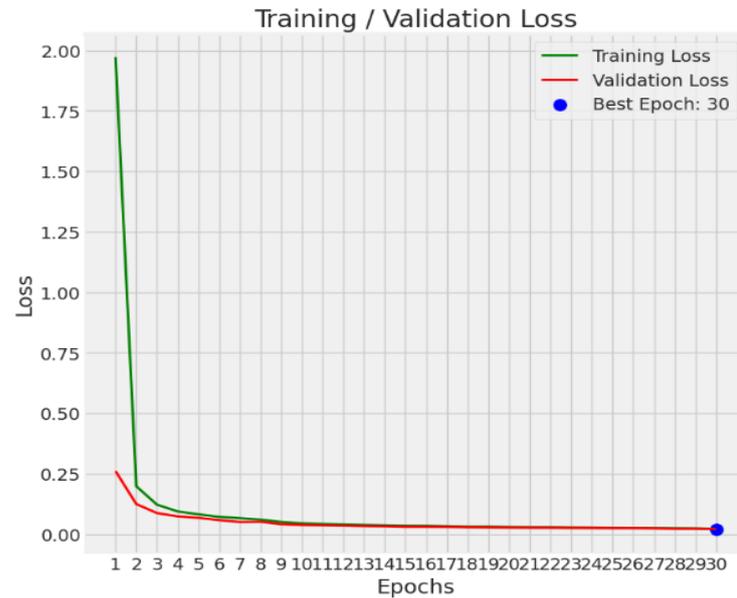


Figure 10. Training and Validation loss plot of the suggested model

In the end, the results were beyond expectations, with the F1-score prediction rate increasing to over 99.9% or approximately 100%. (Fig.11). These results highlight the potential of an integrated approach to optimize accuracy in identifying disease in MRI-based brain tumors in the future.

	precision	recall	f1-score	support
glioma_tumor	1.00	1.00	1.00	4308
meningioma_tumor	1.00	1.00	1.00	4293
no_tumor	1.00	1.00	1.00	2072
pituitary_tumor	1.00	1.00	1.00	4310
accuracy			1.00	14983
macro avg	1.00	1.00	1.00	14983
weighted avg	1.00	1.00	1.00	14983

Figure 11. Final F1-score result of the proposed model using augmented dataset in Experiment 3, showing nearly perfect classification.

Compared to prior works in the literature, our approach demonstrates significant improvements:

- [6] applied EfficientNetB0 independently and reported F1-scores around 95%.
- [7] utilized MobileNetV2 for Alzheimer's and tumor classification but without fusion or bilinear pooling, achieving up to 81% test accuracy.
- [18] achieved 94-97% using a combination of saliency maps and optimization algorithms.

In contrast, our model - combining **EfficientNetB0 + MobileNetV2** with **bilinear pooling and augmentation** - achieves **F1-scores nearing 100%** across all classes. This highlights the effectiveness of both our model architecture and training strategy. Additionally, by leveraging lightweight and scalable networks, our approach is more feasible for clinical deployment compared to heavier models like ResNet152 or DenseNet201 used in other works.

## 5. CONCLUSION AND FUTURE WORK

This research underscores the criticality of addressing brain tumors promptly, given their inherent danger to human life. The study focused on enhancing the efficiency of magnetic resonance imaging (MRI) for the early identification of cancerous brain tumor cells. The introduction of combined models, specifically in this study the pooling (MobileNetV2 + EfficientNetB0) model, demonstrated outstanding results, achieving 100% accuracy for the training set, 98% for the validation set but not really good at predicting the data in the test set. Subsequent refinements, such as data set re-division and image rotation techniques, further elevated the model's performance, culminating in an impressive F1-score of almost 100% for all classes.

Based on these findings, the MobileNetV2+EfficientNetB0 model emerges as a promising approach for the early detection of brain tumor cancer cells. The potential impact is substantial, offering increased prospects for early diagnosis and intervention, ultimately safeguarding human health. The study not only contributes to the advancement of medical imaging technology but also highlights the significance of leveraging deep learning methodologies in the field of neurology.

## REFERENCES

- [1] Soni, Vaibhav & Singh, Nikhil & Singh, Rishi & Tomar, Deepak. (2023). Multienncoder-based federated intelligent deep learning model for brain tumor segmentation. *International Journal of Imaging Systems and Technology*. n/a-n/a. 10.1002/ima.22981.
- [2] R. Young, J. Knopp, and A. Edmond, "Brain MRI: tumor evaluation," *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 24, no. 4, pp. 709–724, 2006.
- [3] T. Logeswari and M. Karnan, "An improved implementation of brain tumor detection using segmentation based on hierarchical self organizing map," *International Journal of Computer Theory Engineering*, vol. 2, no. 4, pp. 591–595, 2010.
- [4] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2015," *CA: a Cancer Journal for Clinicians*, vol. 65, no. 1, pp. 5–29, 2015.
- [5] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2019," *CA: a Cancer Journal for Clinicians*, vol. 69, no. 1, pp. 7–34, 2019.
- [6] R. L. Siegel, K. D. Miller, H. E. Fuchs, and A. Jemal, "Cancer statistics, 2021," *CA: a Cancer Journal for Clinicians*, vol. 71, no. 1, pp. 7–33, 2021.
- [7] H. T. Zaw, N. Maneerat and K. Y. Win, "Brain tumor detection based on Naïve Bayes Classification," *2019 5th International Conference on Engineering, Applied Sciences and Technology (ICEAST)*, Luang Prabang, Laos, 2019, pp. 1-4, doi: 10.1109/ICEAST.2019.8802562.
- [8] A. Pandey and V. K. Pandey, "Deep Transfer Learning Models for Brain Tumor Classification Using Magnetic Resonance Images," *2023 IEEE 12th International Conference on Communication Systems and Network Technologies (CSNT)*, Bhopal, India, 2023, pp. 398-403, doi: 10.1109/CSNT57126.2023.10134745.
- [9] Farhana Alam, Farhana Chowdhury Tisha, Sara Anisa Rahman, Samia Sultana, Md. Ahied Mahi Chowdhury, Ahmed Wasif Reza and Mohammad Shamsul Arefin, "Automated Brain Disease Classification using Transfer Learning based Deep Learning Models" *International Journal of Advanced Computer Science and Applications(IJACSA)*, 13(9), 2022. <http://dx.doi.org/10.14569/IJACSA.2022.01309109>
- [10] American Cancer Society. *Cancer Facts & Figures 2019*. <https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures.html>
- [11] SEER Cancer Statistics Review 1975–2015. National Cancer Institute. <https://seer.cancer.gov>
- [12] American Cancer Society. *Cancer Facts & Figures 2025 (projections)*. <https://acsjournals.onlinelibrary.wiley.com/>
- [13] The Brain Tumour Charity (UK). <https://www.thebraintumourcharity.org>
- [14] Cancer Research UK. Brain Tumour Statistics. <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/brain-tumours>
- [15] Khan, Muhammad & Khan, Awais & Alhaisoni, Majed & Alqahtani, Abdullah & Alsubai, Shtwai & Alharbi, Meshal & Malik, Nazir & Damaševičius, Robertas. (2022). Multimodal Brain Tumor Detection and Classification using Deep

- Saliency Map and Improved Dragonfly Optimization Algorithm. *International Journal of Imaging Systems and Technology*. 33. 10.1002/ima.22831.
- [16] M. Michael and W. Smith, "WebMD cancer center: types of brain cancer," 2020, <https://www.webmd.com/cancer/brain-cancer/brain-tumor-types>.
- [17] P. C. Wesseling and D. Capper, "WHO 2016 classification of gliomas," *Neuropathology Applied Neurobiology*, vol. 44, no. 2, pp. 139–150, 2018.
- [18] M. R. Gupta, V. Rajagopalan, and B. V. V. S. N. P. Rao, "Glioma grade classification using wavelet transform-local binary pattern based statistical texture features and geometric measures extracted from MRI," *Journal of Experimental Theoretical Artificial Intelligence*, vol. 31, no. 1, pp. 57–76, 2019.
- [19] M. Y. Talo, O. Yildirim, U. B. Baloglu, G. Aydin, and U. R. Acharya, "Convolutional neural networks for multi-class brain disease detection using MRI images," *Computerized Medical Imaging and Graphics*, vol. 78, article 101673, 2019.
- [20] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4510–4520)
- [21] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016)*, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
- [22] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4510–4520, 2018.
- [23] Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR, 2019.
- [24] H. Asri, H. Mousannif, H. Moatassime, & T. Noël. Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis. ANT/SEIT, 2017.
- [25] N. Ponraj, E. Jenifer., P. Poongodi, S. Manoharan. "Morphological operations for the mammogram image to increase the contrast for the efficient detection of breast cancer", *European Journal of Scientific Research*, (ISSN) 1450-216X, vol. 68, no.4, pp. 494-505, 2021.
- [26] D. R. Pavithra, S. Preethi, & A. K. SriRakshitha. Breast Cancer Classification using the Supervised Learning Algorithms. 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1492-1498, 2021
- [27] H. Asri, H. Mousannif, H. Moatassime, & T. Noël. Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis. ANT/SEIT, 2017.
- [28] B. Gayathri, C. Sumathi, T. Santhanam, "Breast cancer diagnosis using machine learning algorithm a survey". *International Journal of Distributed and Parallel Systems*, vol. 4, no. 3, pp. 39-50, 2017.
- [29] M. F. Ak. A Comparative Analysis of Breast Cancer Detection and Diagnosis Using Data Visualization and Machine Learning Applications. *Healthcare*, vol. 8, no. 2, pp.1-11, 2020.
- [30] H. Mousannif, H. Asri, H. A. Moatassime, T. Noel. Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis. *Procedia Computer Science*, vol. 83, pp. 1064–1069, 2017.
- [31] U. K. Kumar, M. B. S. Nikhil and K. Sumangali, "Prediction of breast cancer using voting classifier technique," 2017 IEEE International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), 2017, pp. 108-114.
- [32] Hoang-Tu Vo, Nhon Nguyen Thien and Kheo Chau Mui, "Tomato Disease Recognition: Advancing Accuracy Through Xception and Bilinear Pooling Fusion" *International Journal of Advanced Computer Science and Applications(IJACSA)*, 14(8), 2023. <http://dx.doi.org/10.14569/IJACSA.2023.01408113>
- [33] Cengil, E., Çinar, A. (2019). Multiple classification of flower images using transfer learning. In *2019 International Artificial Intelligence and Data Processing Symposium (IDAP)*, Malatya, Turkey, pp. 1-6. <https://doi.org/10.1109/IDAP.2019.8875953>
- [34] D. M. Joshi, N. K. Rana and V. M. Misra, "Classification of Brain Cancer using Artificial Neural Network," *2010 2nd International Conference on Electronic Computer Technology*, Kuala Lumpur, Malaysia, 2010, pp. 112-116, doi: 10.1109/ICECTECH.2010.5479975.
- [35] M. Thachayani and S. Kurian, "AI Based Classification Framework For Cancer Detection Using Brain MRI Images," *2021 International Conference on System, Computation, Automation and Networking (ICSCAN)*, Puducherry, India, 2021, pp. 1-4, doi: 10.1109/ICSCAN53069.2021.9526456.
- [36] B. Erkal, S. Başak, A. Çiloğlu and D. D. Şener, "Multiclass Classification of Brain Cancer with Machine Learning Algorithms," *2020 Medical Technologies Congress (TIPTEKNO)*, Antalya, Turkey, 2020, pp. 1-4, doi: 10.1109/TIPTEKNO50054.2020.9299233.
- [37] V. G. Hamsaveni, S. P. Maniraj, R. Krishnaswamy, K. Radhika, V. Venkataramanan and S. Renukadevi, "Intelligent Classification of Brain Cancers by Deep Learning with Inception Network," *2022 International Conference on Computer, Power and Communications (ICCCPC)*, Chennai, India, 2022, pp. 681-684, doi: 10.1109/ICCCPC55978.2022.10072278.
- [38] Shenbagarajan Anantharajan, Shenbagalakshmi Gunasekaran, Thavasi Subramanian, Venkatesh R, MRI brain tumor detection using deep learning and machine learning approaches, *Measurement: Sensors*, Volume 31, 2024, 101026, ISSN 2665-9174, <https://doi.org/10.1016/j.measen.2024.101026>.

- [39] Tan, M., & Le, Q. (2019). *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*. ICML. <https://arxiv.org/abs/1905.11946>
- [40] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). *MobileNetV2: Inverted Residuals and Linear Bottlenecks*. CVPR. <https://arxiv.org/abs/1801.04381>

## BIOGRAPHY OF AUTHORS



**Nhon Nguyen Thien** holds a Bachelor of Information Systems from Can Tho University, Vietnam, in 2013. In 2017, He graduated with a master's degree in Information Systems from Can Tho University in Vietnam. Currently, He is working as a lecturer in Information Technology Department at FPT University, Can Tho campus in Vietnam. His research areas of interest include Data science, Machine learning, Deep learning, and Web application development. He can be contacted at email: [nhont9@fe.edu.vn](mailto:nhont9@fe.edu.vn)



**Kheo Chau Mui** holds a Bachelor of Information Systems from Tay Do University, Vietnam, in 2010. In 2020, She graduated with a master's degree in Computer Science from Can Tho University in Vietnam. Currently, She is working as a lecturer in Information Technology Department at FPT University, Can Tho campus in Vietnam. Her research areas of interest include Image processing, Image classification, Data science, Object detection, Deep learning, and Machine learning. She can be contacted at email: [kheocm@fe.edu.vn](mailto:kheocm@fe.edu.vn)



**Hoang Tu-Vo** holds a Bachelor of Information Systems from Can Tho University, Vietnam, in 2011. In 2013, He graduated with a master's degree in Information Systems from Can Tho University in Vietnam. Currently, He is working as a lecturer in Information Technology Department at FPT University, Can Tho campus in Vietnam. His research interests include Machine learning, Deep learning, Image processing, and Computer vision. He can be contacted at email: [tuvh6@fe.edu.vn](mailto:tuvh6@fe.edu.vn)



**Phuc Pham Tien** earned his Bachelor's degree in Information Systems from Can Tho University, Vietnam, in 2003. In 2010, he completed a Master's degree in Information Technology at the same university. Currently, he is a lecturer in Information Technology Department at FPT University, Can Tho campus, Vietnam. His research interests include machine learning and deep learning. He can be reached at [phucpt10@fe.edu.vn](mailto:phucpt10@fe.edu.vn)



**Huan Lam Le** earned his Bachelor's degree in Information Systems from Can Tho University, Vietnam, in 1996. In 2010, he completed a Master's degree in Information Technology at the same university. Currently, he is a lecturer in Information Technology Department at FPT University, Can Tho campus, Vietnam. His research interests include machine learning and deep learning. He can be reached at [huanll@fe.edu.vn](mailto:huanll@fe.edu.vn)