Ensemble Transfer Learning Approach for Efficient Brain Tumour Prediction

Istabraq H. Jasim, Zakariya A. Oraibi*, and Entesar B. Talal

Department of Computer Science, College of Education for Pure Sciences, University of Basrah, Iraq
Email: estabraq.haider@uobasrah.edu.iq (I.H.J.); zakaria_au@uobasrah.edu.iq (Z.A.O.);
entesar.barges@gmail.com (E.B.T.)
*Corresponding author

***Abstract*—Brain tumour detection and classification remain critical tasks in medical imaging. Although Convolutional Neural Networks (CNNs) have demonstrated significant potential, most studies rely on two-dimensional Magnetic Resonance Imaging (MRI) datasets, which may contain duplicated or redundant samples, leading to over-optimistic performance. In this paper, a new approach that utilises ensemble transfer learning is presented, integrating four well-known CNN architectures: EfficientNetB7, MobileNetV2, VGG16, and Xception. Each model is first fine-tuned through transfer learning, and their extracted feature maps are concatenated to form a richer representation before classification using dense layers and a softmax output layer. In addition to reporting the ensemble performance, thorough experiments are conducted to evaluate the contribution of each architecture individually and compare the proposed method against state-of-the-art CNNs and Vision Transformer (ViT)-based models. Results on the Kaggle7023 dataset demonstrate high classification accuracy, although some limitations related to the dataset are emphasised, particularly its 2D nature and potential redundancy, which restrict generalisation. Findings highlight both the promise and the challenges of ensemble deep learning models for uprediction, specifically predicting brain tumours.**

***Keywords*—brain tumour, artificial intelligence, diagnosis, medical imaging**

I. Introduction

Treatment planning and prognosis for brain tumour patients are heavily influenced by accurate brain tumour prediction, a critical task in medical imaging. Convolutional Neural Networks (CNNs) emerged as an efficient technology to provide accurate prediction of various brain tumour classes using Magnetic Resonance Imaging (MRI) scans—the ability to extract hierarchical features using CNNs, making them a perfect fit for accurate tumour identification (Fig. 1). To make CNNs more productive, leveraging large datasets of labelled MRI images is crucial, enabling these neural networks to recognise intricate patterns and distinctive variations that are often difficult for the human eye to capture, thereby facilitating the accurate classification of various brain tumour classes.

Numerous studies in the literature have already explored the application of CNNs for brain tumour prediction, highlighting their importance in improving diagnostic efficiency and accuracy. Reddy *et al.* [1] explored using a fine-tuned Vision Transformer (ViT) for enhanced multi-class brain tumour classification using MRI scans [2]. In their work, a pre-trained ViT model is fine-tuned on a dataset of brain MRI images, achieving a high accuracy of 98.70% [1]. Hammad *et al.* [3] presented an article focused on developing an efficienttumour detection in MRI images. Their work addressed the need for computationally less demanding models suitable for clinical applications, where resource constraints might be a concern—a high accuracy of 96.86%. In addition, Rao *et al.* [4] investigated the use of pre-trained CNN models for efficient brain tumour detection and classification in MRI images. Their study focused on leveraging the feature extraction capabilities of established CNN architectures, likely through transfer learning. Their work achieved an accuracy of 98.0%. Iqbal *et al.* [5] utilised a framework based on EfficientNet-B0 for an intelligent diagnostic system to enhance brain tumour multi-classification within Internet of Medical Things (IoMT) applications. Their approach outperformed the original EfficientNet-B0 baseline and achieved an accuracy of 99.0%. On the other hand, Sarkar *et al.* [6] proposed an approach that was based on a hybrid of technique of CNN and machine learning. Their research focused on utilizing the pre-trained AlexNet architecture to extract relevant features from MRI scans to feed these features into various machine learning algorithms. Their work achieved an accuracy rate of 98.15%. Mehrotra *et al.* [7] proposed a transfer learning approach for brain tumour classification. Their research focused on leveraging pre-trained deep learning models, adapting them to the specific task of brain tumour classification using MRI images. They achieved a score of 99.04%. Results of applying the proposed methodology on Kaggle7023 dataset in Table I.

Due to the need for an efficient automated tool to perfectly diagnose brain tumour types accurately to make sure patients receive the correct treatment instantly, a robust approach is required to meet this demand. In this

Manuscript received July 25, 2025; revised August 14, 2025; accepted October 15, 2025; published April 28, 2026.

article, a novel hybrid approach is proposed that multiple state-of-the-art pre-trained CNN architectures to perform robust brain tumour detection. The proposed approach four architectures: EfficientNetB7, MobileNetV2, VGG-16, and Xception. First, transfer learning is applied to these networks to avoid training them from scratch. After that, a concatenation technique is used to merge the resultant feature maps of each network, forming a bigger feature map. Then, the newly combined model is trained. Finally, testing is done and the classification accuracy is reported. The reason behind adopting this hybrid approach is to provide a comprehensive feature map that captures the most important features, which make the process of robustly classifying brain tumour classes more efficient. The new approach demonstrated the benefit of using different feature maps to capture the required patterns of state-of-the-art CNN models. The new technique was applied to a challenging dataset with four classes and scored a high accuracy of 99.62%. This accuracy is considered the highest among the the other research done on the same dataset.

The rest of the paper is organized as follows: Section II provides a thorough literature review of the previous methods used for brain tumor classification. Section III introduces the four deep learning modes used in this paper with the proposed fusion mechanism. Section IV discusses the database used in the experiments. Sections V and VI introduce the results generatued using the proposed approach and the comparison with state-of-the-art methods. Secion VII covers the discussion related to the proposed approach and the results. Finally, Section VIII concludes the paper.

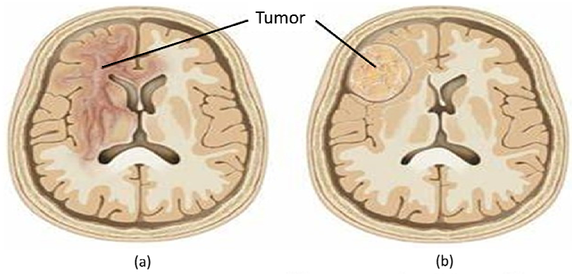

Fig. 1. The difference between the average normal brain and the tumour brain. (a) A tumor invading destroying normal brain tissue. (b) A tumor pressing on normal tissue and causing increasing pressure within the brain.

TABLE I. AN OVERVIEW OF THE PROPOSED METHODS ON KAGGLE7023 DATASET

Reference	Year	Methods	Dataset	Accuracy (%)	Limitation
Reddy *et al*. [1]	2024	FTVT	Kaggle7023	98.62	Low accuracy compared to other methods.
Rao *et al*. [4]	2024	ResNet50, EfficientNet, Fully connected layers	Kaggle7023	98	Improvement can be made to enhance the accuracy using advanced deep learning techniques.
Iqbal *et al*. [5]	2024	LoMT, EfficientNet	Kaggle7023	99	The accuracy can be further improved.
Amran *et al*. [8]	2022	Feature extraction, and a SoftMax, traditional technique, ResNet, MobileNet V2, Alex Net, Squeeze Net, VGG-16	Kaggle7023	99.51	It only captures location-dependent features. Therefore, the categorisation ends up being inaccurate if there is even a little shift in the position of the feature inside the image.
Ullah *et al*. [9]	2023	Gabor+ResNet50+svm	Kaggle7023	95.73	Discussed in the manual are types of techniques for detecting brain tumours at a very early stage.
Wang *et al*. [10]	2024	RanMerFormer	Kaggle7023	98.86	The RanMerFormer method has its limitations and areas for improvement.
Bansal *et al.* [11]	2024	CNN-SVM	Kaggle7023	99	Low accuracy compared to other methods.
Rahman *et al.* [12]	2024	Resize + egrey scale transformation + augmentation + dilated PDCNN + machine learning classifiers + average ensemble	Kaggle7023	98.67	As the number of patients has grown, individually analysing these images has become laborious, disorganised, and frequently incorrect.
Albalawour *et al.* [13]	2024	VGG with Federated Learning	Kaggle7023	98	Federated learning-based CNN classification presents a significant improvement in brain tumour detection, which is essential to address its current limitations and to explore its full potential in the broader context of medical imaging is essential.
Nag *et al*. [14]	2024	TumorGANet	Kaggle7023	99.53	The model was trained and evaluated on a specific dataset of 7,023 MRI images. Its performance “may not generalise across all data types”, meaning results could vary with different scanner types, protocols, or patient populations.

II. LITERATURE REVIEW

In recent years, many methodologies have been introduced to tackle the problem of brain tumour prediction efficiently. The majority of them relied on deep learning. Mehrotra *et al*. [7] proposed a modified deep learning network using the objective dataset and performed classification using a softmax layer. A highest accuracy of 99.04% was achieved. On the other hand, Chinar and Yildirim [15] proposed a CNN model to classify brain tumours using ResNet50, AlexNet, GoogleNet, DensNet201 and InceptionV3 models through transfer learning and achieved 95% accuracy. A big confusion occurred since combining these models generated better results.

Reddy *et al*. [1] introduced a Fine-Tuned Vision Transform (FTVT) model with updated classifier heads, tuned the hyperparameters for efficient brain tumour classification and along the same lines. Pre-trained convolutional neural networks achieved the best accuracy of 99.04%. In addition, Juneja *et al.* [16] introduced a set of traditional algorithms that rely on region-based preprocessing, which achieved accuracies ranging from 71.39% to 94.68%. Tummala *et al.* [2] employ four well-known transfer learning approaches—ResNet152, VGG19, DenseNet169, and MobileNetv3—to create an ensemble model that is efficient to classify the four-class brain tumor dataset. A high accuracy of 99.51% was achieved. Table I shows the results of other methodologies applied on the same Kaggle7023 image brain tumor database.

III. Materials and Methods

In this section, the robust hybrid methodology proposed in this paper is introduced. The new methodology is dependent on four well-known CNN architectures, which are EfficientNetB7, MobileNetV2, VGG-16, and Xception. Transfer learning was applied before the feature maps were extracted and combined from each model.

A. *EfficientNetB7 Architecture*

This model belongs to the EfficientNetB family, and the use of EfficientNetB7 is due to its powerfulness in this series in terms of performance, achieving high classification results and being very accurate, especially for classifying medical images. This large and deep model is considered as a very efficient model compared to models of the same size. It is very accurate, as it can capture fine details, processing images with a resolution of 600×600 pixels. Therefore, it is very suitable for classifying medical images, as the model can be fine-tuned for medical image classification tasks [17]. Fig. 2 shows the structure of the EfficientNetB7 model.

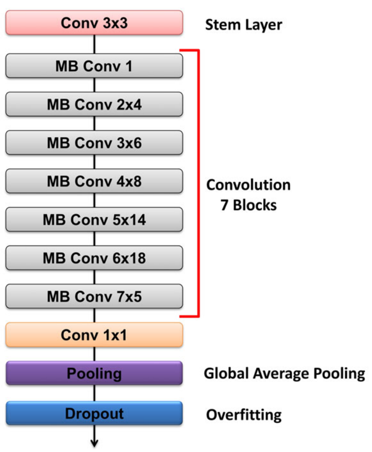

Fig. 2. EfficientNetB7 architecture.

B. *MobileNetV2 Architecture*

It is a convolutional neural network designed to reduce CNN computational complexity, memory consumption, and maintain high classification accuracy. This improved architecture can reduce the computational complexity of convolutional operations. Fig. 3 illustrates the structure of the MobileNetv2 model [5].

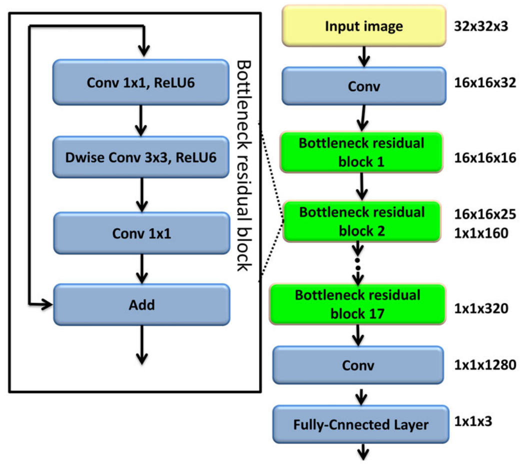

Fig. 3. MobileNetV2 architecture.

C. *VGG16 Architecture*

It is a deep neural network consisting of 16 layers, including 13 convolution layers and threedense layers. Pooling layers are used to improve the network's performance. It prevents bias using dropout. The architecture of VGG-16 is shown in Fig. 4.

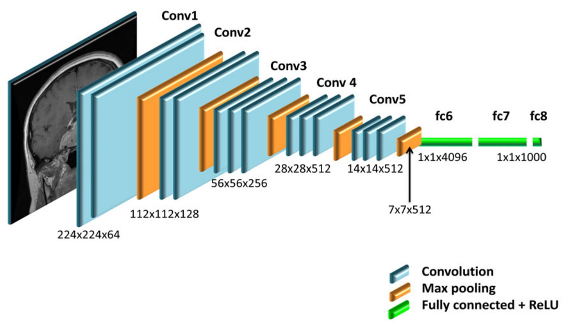

Fig. 4. VGG16 architecture.

D. *Xception Architecture*

It is a convolutional neural network model that uses Keras, an improved model of the Inception architecture. This model is based on vertically separated convolutions, where the convolution process is a two-stage process: applying a filter to the input channel independently, and a stage that combines the output using a convolution between the channels. The performance is superior and provides high classification results. Hence, reducing the number of parameters without compromising accuracy, makes it efficient [18]. Fig. 5 shows a diagram of the Xception architecture.

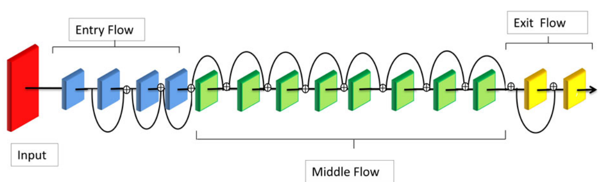

Fig. 5. Xception architecture.

E. Transfer Learning

Transfer Learning (TL) is a machine learning technique that uses pre-trained models to solve new problems. A pre-trained model is loaded and modified to suit the new task. This reduces training time and improves performance. TL allows the reuse of knowledge gained from previous training and its application to the new task. Some layers are frozen, and the last layers are modified to suit the new task. This reduces computational cost and improves performance. TL is an efficient, fast, and accurate solution to some tasks and problems. Fig. 6 shows transfer learning.

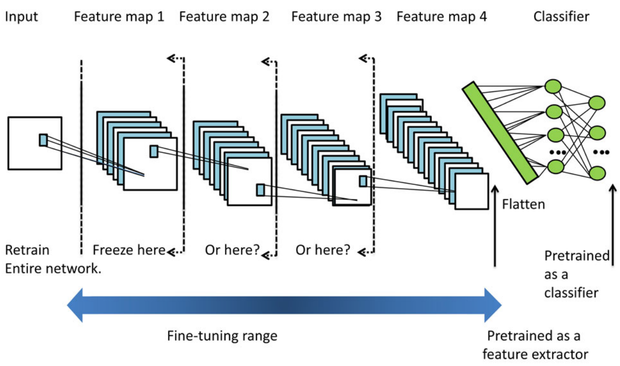

Fig. 6. The process of transfer learning.

F. The Proposed Hybrid Model

As mentioned earlier, the proposed technique relies on merging the feature maps of four robust CNN architectures. First, transfer learning is applied to each standalone model. Then, feature maps are extracted from each model. These feature maps will be combined to form a bigger feature map to be used later for the classification task using the softmax layer. In the experiments, the robustness of the proposed hybrid approach is demonstrated by thoroughly analysing the accuracy of each model alone and the accuracy of merging these models. Fig. 7 shows the hybrid scheme of the new model. All the dimensions of the extracted feature maps are illustrated, and the dimension of the final feature map is also presented.

Table I has been included, presenting the studies utilised in developing the ViT model, detailing for each study the data, methods, and techniques employed, providing a comprehensive overview of the key references informing the current research. The proposed approach as mentioned is based on transfer learning. A partial fine-tuning strategy was utilized. The first 20 layers of the baseline skeleton were frozen to preserve low-level feature representations (i.e., color, shape, etc.), while the last 30 layers were refrozen and fine-tuned on the target dataset to learn brain tumor-specific features. This approach allowed the model to adapt high-level features to the task of brain tumor classification while avoiding overfitting and high computational costs. Fast, robust, and high-classification accuracy models were used, as were lightweight models. This contributed to building a model capable of classifying brain tumors with high accuracy. Table II lists the hyperparameters used in the experiments.

TABLE II. LIST OF HYPERPARAMETERS UTILIZED DURING TRAINING

Hyper parameter	Value
Learning Rate	0.0001
Epochs	55
Batch-Size	32
Optimizer	Adam
Shuffel	True
Loss-Function	Categorical Cross-entropy

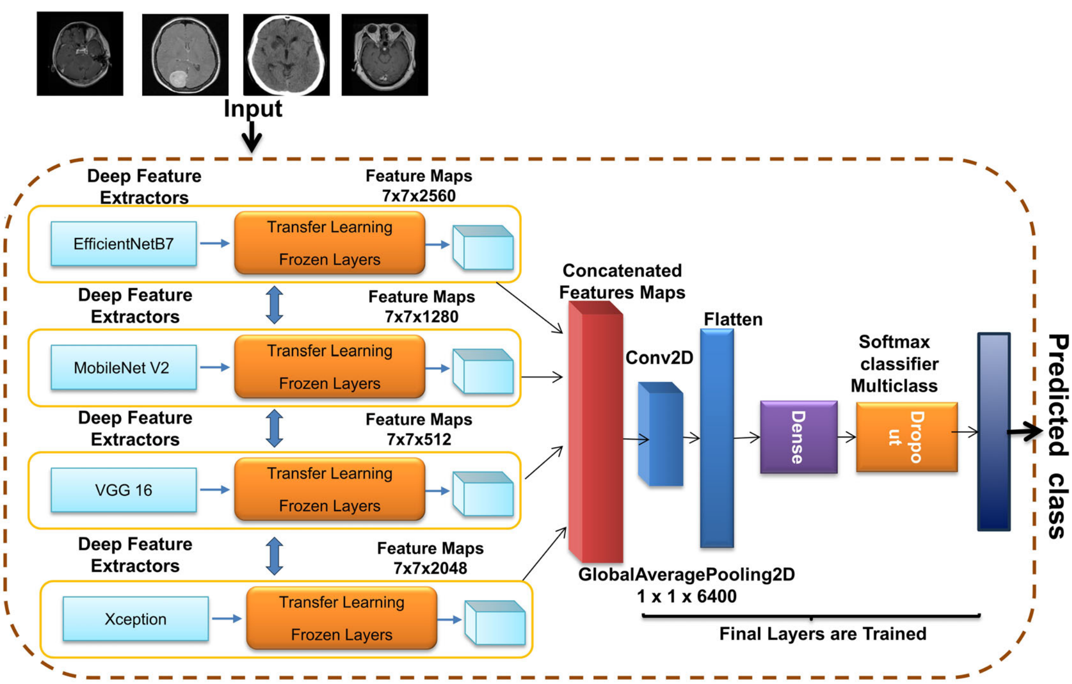

Fig. 7. The proposed hybrid model applied using transfer learning.

Feature splicing, or concatenating multiple feature maps of different CNN architectures, is effective because it has the ability to make the model more diverse, leading to a lower error rate. Distinct architectures like EfficientNet and VGG possess different inductive biases, meaning they extract diverse, low-correlation features from the same input. Splicing these representations provides the final classifier with a richer, more robust feature space, effectively reducing the collective variance and uncorrelated errors of the individual models for superior generalisation.

IV. DATASET

The brain tumour dataset (kaggle7023: https://www.kaggle.com/datasets/mohamadabouali1/mriz brain-tumor-dataset-4-class-7023-images.) used in this study consists of 7023 MRI images. The dataset was obtained from Kaggle, an open-access website that contains a dataset for scientific research purposes. It consists of four types of brain tumours: Glioma, Meningioma, No tumour, and Pituitary. The number of images per class varies. In addition, the dataset was randomly split into training images, validation images, and test images for the experiments. The training set contained 4922 images, the testing set contains 1406 images, and the validation set contains 704 images [19]. Furthermore, the database used in the experiments to assess the proposed features concatenation model is a public set of images in the Kaggle repository [15].

It has four classes: meningioma 1645 images, glioma 1621 images, pituitary tumour 1757 images, and no tumour 2000 images in that dataset. The total number of samples available is 7023. Sample images of the dataset are shown in Fig. 8. The images in this datasetwere resized to fit the model requirements to 224×224×3. This allows the meta learner to leverage a richer set of learned representations from various perspectives before the final classification. Let $Fi(x)$ denote the feature map extracted from an intermediate layer of the model for an input image x. Each $Fi(x)$ will have a shape of (Hi, Wi, Ci), where Hi represents height and width, while Ci represents the number of channels. To write this as an equation, the following formula is used:

$$Fcombined(x) = Concatenate\left(F1(x), F2(x), ..., FN(x)\right) \quad (1)$$

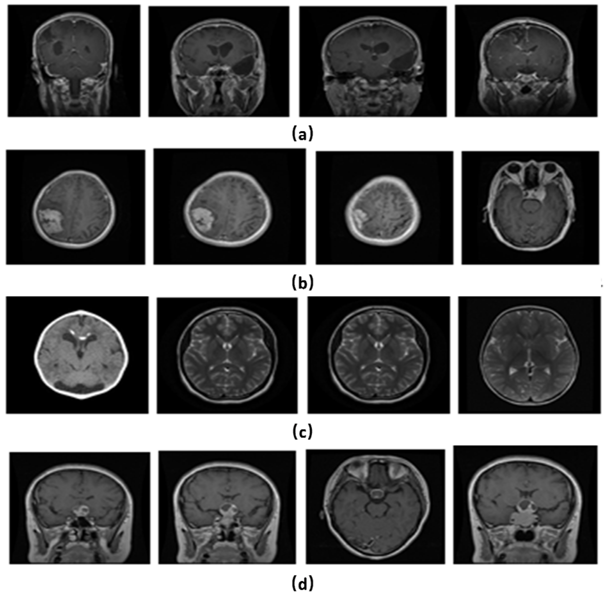

Fig. 8. Sample of images from the Kaggle7023 dataset. (a) Sample brain tumor dataset with glioma tumor. (b) Sample brain tumor dataset with meningioma tumor. (c) Sample brain tumor dataset with no tumor. (d) Sample brain tumor dataset with pituitary tumor.

V. EXPERIMENTAL RESULTS

In this section, the results of the proposed methodology are presented after applying the hybrid approach to the Kaggle7023 dataset. During the experiments, 80% of the data was used for training; hence, the remaining 20% of samples were applied for testing. Our used hyperparameters, including the Adam optimiser, a batch size of 32, a learning rate of 0.0001, a cross-entropy loss function, and a number of epochs of 55 for the hybrid model. In Table III, the performance of the proposed methodology in terms of accuracy metrics is presented.

TABLE III. RESULTS OF APPLYING THE PROPOSED METHODOLOGY ON THE KAGGLE DATASET

Method	Accuracy	Precision	F1-Score	Recall
EfficientNetB7	97.56%	97.5%	97.5%	97.25%
MobileNetV2	82%	80.75%	79.75%	81%
VGG16	98.78%	98.5%	98.5%	98.5%
Xception	85.66%	85.75%	84.5%	84.5%
EfficientNetB7 + MobileNetV2	99.16%	99.25%	99.25%	99.25%
EfficientNetB7 + MobileNetV2 + VGG16	99.39%	99.25%	99.25%	99.25%
EfficientNetB7 + MobileNetV2 + VGG16 + Xception	99.62%	99.75%	99.75%	99.5%

To program the deep learning models, Keras library has been used along with TensorFlow. To classify the machine learning classifiers, scikit-learn has also been utilized, which uses Softmax and ReLU to train four models: EfficientNetB7, MobileNetV2, Xception, and VGG16. All experiments were performed on Google Colab Pro+ platform since the resultant model requires big resources. To validated the proposed model, model's performance results using various performance metrics: confusion matrix, accuracy, recall, F1-Score, and loss. The training MRI images, including brain tumor images for four disease classes: meningioma, pituitary tumor, glioma, and tumor-free (benign). The TensorFlow library was used and the layered cascade architectures Dropout, MaxPooling2D, Flatten, Dense, and Conv2D. We then discussed the model's performance using performance metrics and compared our results with previous studies that used the same dataset. We analyzed the model's performance using the confusion matrix, as shown in Fig. 9.

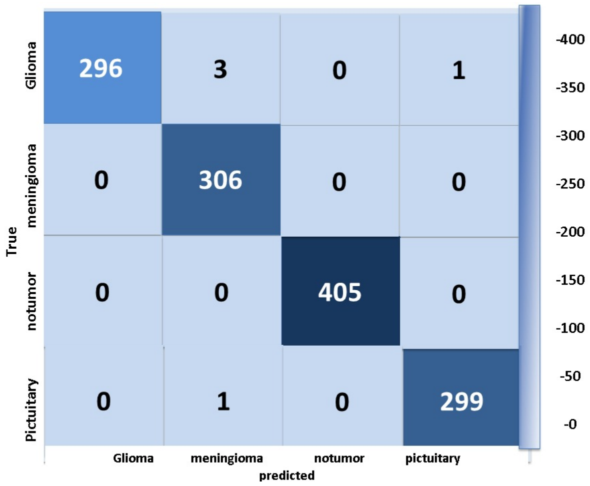

Fig. 9. The confusion matrix.

Fig. 10 shows the Receiver Operating Characteristic (ROC)-AUC, a measure of classification accuracy of the categories. It illustrates the model's ability to distinguish between the categories, with an accuracy of 1, meaning that the model is capable of distinguishing between the categories excellently. The plotted loss and accuracy curves for the dataset, showing that the model achieved high accuracy, as shown in Fig. 11. In addition, the F1-Score was calculated for each class, as shown in Table III.

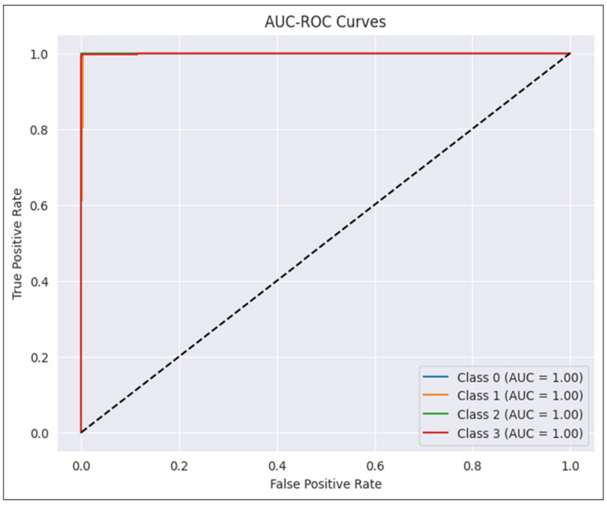

Fig. 10. The AUC-ROC curves.

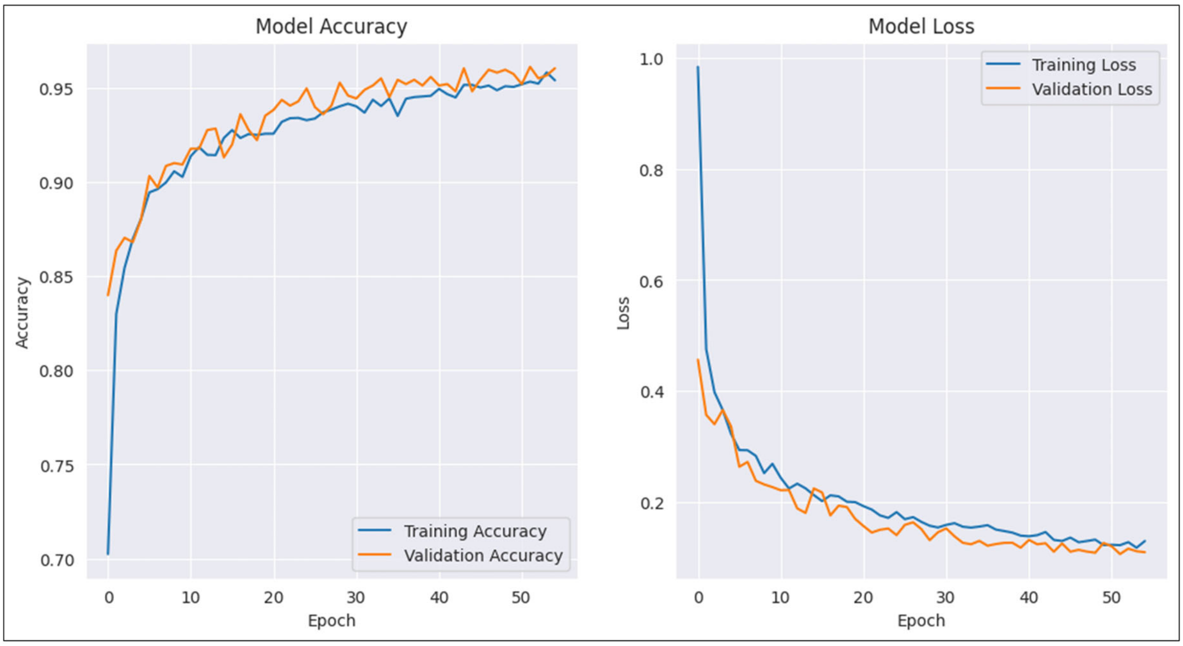

Fig. 11. Model accuracy and model loss after 55 epochs.

VI. Comparison with State-of-the-Art Results

To show how powerful the proposed approach in this paper is, a comparison has been made between the proposed methodology and the state-of-the-art approach from the literature applied on the same dataset. The classification report is presented in Table IV.

To compare the results of the proposed approach with literature review results, Rao *et al.* [4] proposed RanMerFormer, which can achieve outstanding classification performance with an overall accuracy of around 98% for the Kaggle7023. The RanMerFormer is a useful CAD system for brain tumour classification. The ViT ranMerFormer model provides for classifying images of brain tumours by 2D MRI. ViT models perform better in computer vision, but at a cost compared to CNN models. The mechanism of action is to integrate a new code from the trans to the previous blocks to reduce redundancy and improve computational cost.

TABLE IV. CLASSIFICATION REPORT

Class	Precision	Recall	F1-Score	Support	AUC-ROC
Glioma	1.00	0.98	0.99	300	0.99
Meningioma	0.97	0.98	0.97	306	0.99
Notumor	1.00	0.99	0.99	405	
Pituitary	0.99	0.99	1.00	1.00	
Accuracy	0.99	0.99	1.00	1311	
macro avg	1.00	1.00	1.00	1311	1
weighted avg	1.00	1.00	1.00	1311	0.99
Average AUC-ROC					99.97

To guide the final prediction, RVFL was used. DS1 has four categories of images: no tumour, glioma, pituitary gland and meningioma. In comparison, DS2 has three classes for muscle mass/meningioma, pituitary and meningioma. The model achieved 99% accuracy for both groups in the future possible collection of more three-dimensional magnetic resonance imaging [20]. Deep learning and FTVT models from previous studies, as presented in Table I, illustrate the importance of fine-tuning made in this study. It contains a large number (7053) of MRI scan images of the brain, and various preprocessing techniques such as data augmentation, normalisation, and rigorous training which are used for model training. This study has achieved remarkable accuracy for classifying brain tumors which helps professionals detect and classify tumours in the early stages. Models with different evaluation scales to improve analysis and comparison individually and across each tumour category. FTVT models achieved greater accuracy than other deep learning models, ResNet-50, EfficientNet-b0, and MobileNet-V2. Also, FTVT models individually and across tumour classes outperform these deep learning models. The model achieved 98.62% accuracy [1].

Rao *et al*. [4] used a CNN and achieved a 97% accuracy without data augmentation. As the data increased, the validation model scored 85% accuracy and 99% training accuracy. To classify the tumour, transformational learning models were used. The Recent50 model yielded 96% test accuracy with 9.4% test loss, 97% accuracy and 98% callback, while the EfficientNet model had 98% test accuracy with 5.09% test loss, 98.5% accuracy and 99% callback. It is necessary to collect more data to increase the classification accuracy [4]. Classification of medical images, especially brain tumours, is performed using deep CNN structures and transfer learning. The ImageNet dataset classifies brain tumours into glioma, meningioma, and pituitary tumour, using the Figshare dataset. VGGNet, GoogleNet, and AlexNet models are used on MRI images. To explore and evaluate the performance of deep networks, transfer learning was used through freezing and fine-tuning, achieving an accuracy of 98.69% [21]. A pre-trained convolutional neural network based on deep learning was used to classify brain tumour images into benign and malignant tumours. Transfer learning techniques are used, and the AlexNet transfer learning model was used, achieving an accuracy of 99.04% [7]. Table V shows a comparison between the proposed model and previous studies that used the same dataset.

TABLE V. COMPARISON BETWEEN THE PROPOSED HYBRID FEATURES TECHNIQUE OF EFFICIENTNETB0, MOBILENET, XCEPTION AND THE BEST TECHNIQUES FROM LITERATURE

Reference	Methodology	Dataset	Year	Accuracy (%)
Reddy *et al*. [1]	FTVT	Kaggle7023	2024	98.62
Rao *et al*. [4]	EfficientNet	Kaggle7023	2024	98
Ullah *et al*. [9]	Gabor+ResNet50+Svm	Kaggle7023	2023	95.73
Wang *et al*. [10]	RanMer Former	Kaggle7023	2024	98.86
Bansal *et al*. [11]	CNN-SVM	Kaggle7023	2024	99
Rahman *et al*. [12]	Resize + gray scale transformation + augmentation + dilated PDCNN + machine learning classifiers + average ensemble	Kaggle7023	2024	98.67
Albalawour *et al*. [13]	VGG with federated learning	Kaggle7023	2024	98
Nag *et al*. [14]	Tumor GANet	Kaggle7023	2024	99.53
Proposed	Hybrid features (EfficientNetB7, MobileNetV2, VGG16, Xception)	Kaggle7023	--	99.62

VII. DISCUSSION

The experimental results highlight that the proposed ensemble model achieves superior accuracy compared to each individual CNN baseline. Specifically, EfficientNetB7 and Xception contributed the most to the final ensemble performance, while MobileNetV2 and VGG16 provided complementary but weaker improvements. This demonstrates the value of combining diverse feature representations.

However, it is important to stress that the dataset used (Kaggle 7023) consists of 2D MRI slices and has been reported to include repeated or very similar samples, which may inflate accuracy metrics. Thus, while the reported accuracy exceeds 99%, the generalizability of the model to independent clinical datasets remains uncertain. Moreover, the comparison with Vision Transformer (ViT)-based approaches shows that while CNN ensembles can still achieve competitive results, ViTs are emerging as strong alternatives for medical imaging tasks. Future work should explore combining CNNs with transformer-based models, as well as validating results on larger and more diverse 2D MRI datasets to ensure clinical applicability.

To further assess the robustness of the proposed model, additional experiments were conducted by varying the dropout rate. Specifically, five different dropout values ranging from 0.1 to 0.5 were evaluated. As illustrated in Fig. 12, the lowest performance was observed at a dropout value of 0.5, yielding an accuracy of 99.39%. This minor reduction compared to the highest accuracy achieved with a dropout value of 0.1 demonstrates the stability and efficiency of the proposed ensemble learning model.

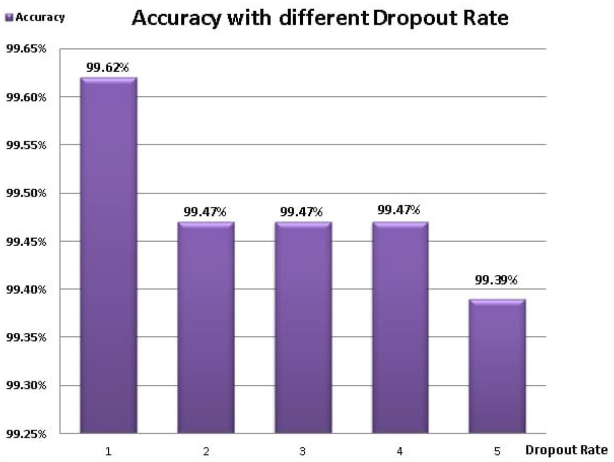

Fig. 12. Accuracy generated using different dropout values.

Although the proposed hybrid deep learning system combining four pretrained models offers superior accuracy, it also struggles with limitations. The main challenges are high computational overhead for training the resultant big model, which could lead to poor interpretability. As a result, future directions will focus on overcoming these barriers through knowledge distillation; implementing dynamic ensembling to save resources by selectively activating models; and developing interpretable fusion techniques, like attention-based weighting, to better understand each model's contribution to the final decision.

Another limitation that arose was the scarcity of image samples of some classes in the provided dataset, which has been overcomed by using augmentation techniques. It is worth mentioning that generalisation in the prediction task could be hindered using 2D images instead of 3D contextual nformation. A hybrid approach of multiple CNN architectures was proposed in this article to generate a high classification accuracy when applied to a challenging four-class brain tumour dataset.

The proposed hybrid ensemble method generates superior accuracy in comparison with other methods from the literature. Amran *et al.* [8] proposed an ensemble learning technique that combines deep and local features with SVM and generated only 95.73% accuracy. On the other hand, Tandel *et al*. [22] proposed a majority voting ensemble technique of five deep learning models, AlexNet, VGG16, ResNet18, GoogleNet, and ResNet50, and achieved an accuracy of 98.88% for classifying FLAIR-MRI brain tumour images. Furthermore, Jain *et al*. [23] proposed a direct ensemble fusion with base models of Pre-trained ResNet50 and SqueezeNet, and their approach achieved 97.9% accuracy.

Removing one CNN model from an ensemble architecture immediately compromises its theoretical power by diminishing model diversity. This loss means the final classifier's spliced feature vector is less descriptive, leading to a significant increase in error correlation among the remaining components. In Table II, removing the Xception model from the ensemble classifier will lead to approximately 0.5% loss in accuracy. Consequently, the ensemble loses its key advantage of variance reduction, resulting in a measurable degradation in predictive accuracy and a less robust system, sacrificing performance gains for marginal improvements in inference speed.

Finally, the new hybrid methodology has the potential to be integrated into clinical practice, thereby revolutionising healthcare. First, many aspects should be addressed carefully, including data, the proposed model, and workflow. Second, the purpose is to focus on moving from manual and time-consuming methods to an accurate, efficient, and reliable machine learning assisted diagnostic process. Hence, data should be prepared carefully with high-quality MRI scans, taking into account adhering to patient privacy regulations. In terms of methodology, the proposed technique has already applied state-of-the-art deep learning methods to ensure robust processing techniques. Moreover, the new hybrid model was tested on a challenging brain tumour dataset, which showed superior performance. Finally, these models and algorithms should be integrated with the hospital infrastructure to provide an efficient user interface.

VIII. CONCLUSIONS

This paper proposed an ensemble transfer learning approach that integrates EfficientNetB7, MobileNetV2, VGG16, and Xception for brain tumour classification. While the hybrid model achieved state-of-the-art accuracy on the Kaggle7023 dataset, further analysis revealed that dataset limitations, such as its 2D format and possible redundancy, may restrict the generalisation of results.

Ablation studies confirmed that EfficientNetB7 and Xception contributed most significantly to the ensemble's performance. Comparisons with recent ViT-based models suggest that hybrid architectures combining CNNs and transformers hold promise for future work. Ultimately, real-world applicability requires validation on larger, multi-centre, 3D MRI datasets to confirm clinical relevance and robustness.

PYTHON CODE AVAILABLE

The Python code written to generate the results in this paper can be found in this link: https://github.com/

astemisalj2024-debug/Code-of-Brain-Tumor-Classification-2025/blob/main/2025__FF_AUC_ROC_onlly_4_99_62__EfficientNetB7%2B_VGG16%2BMobileNetV2%2BXception%2B%2BAdam_2025.ipynb.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Zakariya A. Oraibi provided the idea and helped writing the paper; Istabraq H. Jassim was responsible for building the model and supervises the final results; Entesar B. Talal analyzed the data and helped in writing the paper; all authors had approved the final version.

REFERENCES

[1] C. K. K. Reddy, P. A. Reddy, H. Janapati *et al*., "A fine-tuned vision transformer-based enhanced multi-class brain tumour classification using MRI scan imagery," *Front. Oncol.*, vol. 14, 1400341, 2024.

[2] S. Tummala, "Brain tumour classification from MRI using vision transformers ensembling," *Research Square*, 2022. https://doi.org/10.21203/rs.3.rs-1593662/v1

[3] M. Hammad, M. ElAffendi, A. A. Ateya, and A. A. A. El-Latif, "Efficient brain tumour detection with a lightweight end-to-end deep learning model," *Cancers*, vol. 15, no. 10, p. 2837, 2023.

[4] K. N. Rao, O. I. Khalaf, V. Krishnasree *et al.*, "An efficient brain tumour detection and classification using pre-trained convolutional neural network models," *Heliyon*, vol. 10, no. 17, e36773, 2024.

[5] A. Iqbal, M. A. Jaffar, and R. Jahangir, "Enhancing brain tumour multi-classification using Efficient-Net B0-based intelligent diagnosis for Internet of Medical Things (IoMT) applications," *Information*, vol. 15, no. 8, p. 489, 2024.

[6] A. Sarkar, M. Maniruzzaman, M. A. Alahe, and M. Ahmad, "An effective and novel approach for brain tumour classification using AlexNet CNN feature extractor and multiple eminent machine learning classifiers in MRIs," *J. Sensors*, vol. 2023, no. 1, 1224619, 2023.

[7] R. Mehrotra, M. A. Ansari, R. Agrawal, and R. S. Anand, "A transfer learning approach for AI-based classification of brain tumours," *Mach. Learn. with Appl.*, vol. 2, 100003, 2020.

[8] G. A. Amran, M. S. Alsharam, A. O. A. Blajam *et al.*, "Brain tumour classification and detection using a hybrid deep tumour network," *Electronics*, vol. 11, no. 21, p. 3457, 2022.

[9] S. Ullah, M. Ahmad, S. Anwar, and M. I. Khattak, "An intelligent hybrid approach for brain tumour detection," *Pakistan J. Eng. Technol.*, vol. 6, no. 1, pp. 42–50, 2023.

[10] J. Wang, S. Y. Lu, S. H. Wang, and Y. D. Zhang, "RanMerFormer: Randomised vision transformer with token merging for brain tumour classification," *Neurocomputing*, vol. 573, 127216, 2024.

[11] S. Bansal, R. S. Jadon, and S. K. Gupta, "A robust hybrid convolutional network for tumour classification using brain MRI image datasets," *Int. J. Adv. Comput. Sci. Appl*., vol. 15, no. 4, pp. 576–584, 2024.

[12] T. Rahman, M. S. Islam, and J. Uddin, "MRI-based brain tumour classification using a dilated parallel deep convolutional neural network," *Digital*, vol. 4, no. 3, pp. 529–554, 2024.

[13] E. Albalawi, M. TR, A. Thakur *et al*., "Integrated approach of federated learning with transfer learning for classification and diagnosis of brain tumour," *BMC Med. Imaging*, vol. 24, no. 1, p. 110, 2024.

[14] A. Nag, H. Mondal, M. M. Hassan *et al*., "TumorGANet: A transfer learning and generative adversarial network-based data augmentation model for brain tumour classification," *IEEE Access*, vol. 12, pp. 103060–103081, 2024.

[15] A. Çinar and M. Yildirim, "Detection of tumours on brain MRI images using the hybrid convolutional neural network architecture," *Med. Hypotheses*, vol. 139, 109684, 2020.

[16] M. Juneja, N. Aggarwal, S. K. Saini, S. Pathak, M. Kaur, and M. Jaiswal, "A comprehensive review on artificial intelligence-driven preprocessing, segmentation, and classification techniques for precision furcation analysis in radiographic images," *Multimed. Tools Appl*., vol. 84, no. 19, pp. 21467–21520, 2025.

[17] S. Rokhva and B. Teimourpour, "Accurate & real-time food classification through the synergistic integration of EfficientNetB7, CBAM, transfer learning, and data augmentation," *Food Humanit.*, vol. 4, 100492, 2025.

[18] R. Sathya, T. R. Mahesh, S. B. Khan *et al*., "Employing Xception convolutional neural network through high-precision MRI analysis for brain tumour diagnosis," *Front. Med.*, vol. 11, 1487713, 2024.

[19] H. H. Luong, "Improving brain disease classification accuracy using ensemble learning," in *Proc. International Symposium on Integrated Uncertainty in Knowledge Modelling and Decision Making*, 2025, pp. 197–208.

[20] M. A. B. Siddique, S. Sakib, M. M. R. Khan, A. K. Tanzeem, M. Chowdhury, and N. Yasmin, "Deep convolutional neural networks model-based brain tumour detection in brain MRI images," in *Proc. 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)*, 2020, pp. 909–914.

[21] A. Rehman, S. Naz, M. I. Razzak, F. Akram, and M. Imran, "A deep learning-based framework for automatic brain tumours classification using transfer learning," *Circuits, Syst. Signal Process*., vol. 39, no. 2, pp. 757–775, 2020.

[22] G. S. Tandel, A. Tiwari, and O. G. Kakde, "Multi-class brain tumor grades classification using a deep learning-based majority voting algorithm and its validation using explainable-AI," *J. Imaging Informatics Med.*, vol. 38, pp. 2793–2830, 2025.

[23] S. Jain, V. Jain, and J. M. Chatterjee, "Ensemble based brain tumor classification technique from MRI based on K fold validation approach," *J. Integr. Sci. Technol.*, vol. 13, no. 5, p. 1114, 2025.