Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

A Multi-Model Deep Learning Approach for Classification of Different Types of Brain Tumour

P. Yugandhar Reddy, Prof .E. Sreenivas Reddy
Department of CSE ,ANUCET ,ANU
SCOPE,VIT-AP,AMARAVATHI

ARTICLE INFO

Received: 24 Dec 2024

Revised: 12 Feb 2025

Accepted: 26 Feb 2025

ABSTRACT

This paper examines recent advancements in Deep learning techniques with Classification of brain tumors. Brain tumours pose significant diagnostic challenges, requiring precise and timely intervention for improved patient outcomes. Deep Learning(DL), particularly Convolutional Neural Networks (CNNs), has transformed medical imaging by enabling automated and accurate tumour classification. Key methods discussed include CNNs, transfer learning, & hybrid models, which have shown promising results in improving diagnostic efficiency. Additionally, widely used datasets like TCIA and When evaluating model performance, evaluation metrics like accuracy and AUC-ROC are essential.

DL for brain tumor classification still faces several obstacles despite tremendous advancements. Data scarcity, class imbalance, and model interpretability hinder widespread clinical adoption. Addressing these limitations requires advancements in explainable AI, self-supervised learning future, and multi-model approaches, which integrate diverse data sources for more comprehensive analysis. Research should focus on these areas to enhance reliability and clinical applicability of deep learning-based diagnostic systems. This paper proposes two distinct approaches were implemented and evaluated for brain tumour classification: a Convolutional Neural Network (CNN) designed from scratch, and a Transfer Learning model using EfficientNetB0 pre-trained on ImageNet and the Transfer Learning approach using EfficientNetB0 demonstrated higher accuracy, better generalization, along with instruction behaviour that is more consistent than that of the custom CNN.

Keywords: Deep Learning,Convolutional Neural Network ,Brain tumor, MRI images.

1. INTRODUCTION

Brain tumours are serious medical conditions that require timely and accurate diagnosis. MRI is the standard imaging technique used for detection, but manual interpretation is time-consuming and prone to human error. Differentiating between tumour types like glioma, meningioma, and pituitary tumours is particularly challenging due to overlapping visual features. Traditional machine learning methods rely on manual feature extraction, which limits their effectiveness. Convolutional Neural Networks (CNNs) have exceptional performance in image classification tasks due to their ability to automatically extract hierarchical features from raw data [1]. Traditional CNNs, when designed and trained from scratch, require large amounts of labelled data and computational resources. To address these limitations, **transfer learning** techniques, which reuse pre-trained models on related tasks, have gained popularity in medical image analysis [10], [16], [17]. This paper proposes a CNN-based deep learning model for classifying different types of brain tumours from MRI scans. The goal is to enhance diagnostic accuracy and support clinical decision-making through an automated and efficient approach.

Deep learning which is a subset of machine learning, to develop hierarchical data representations[3,4,5] have several layers which are used in Neural networks. In tasks involving images, Convolutional Neural Networks (CNNs) have shown remarkable performance. Spatial Kernel Selection due to their ability to extract spatial and hierarchical features from images automatically[6,7,8,9]. DL has enabled tumor detection, segmentation, and classification advancements in medical imaging.

Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

Deep learning techniques have evolved over the past two decades, with CNNs becoming prominent in image classification after its breakthrough of AlexNet in 2012. Since then, various architectures such as VGGNet, ResNet, and DenseNet have been explored for medical imaging applications. Initially, medical image processing relied on handcrafted feature extraction techniques, which deep learning has largely automated, leading to increased accuracy and efficiency.

2. LITERATURE REVIEW

Deep learning has significantly impacted various medical imaging applications. One of the most notable contributions is in tumour detection and classification. CNN-based models have shown remarkable performance in detecting and classifying tumours from medical scans such as MRI and CT images. By leveraging deep hierarchical feature extraction, these models can distinguish between benign and malignant tumours with high accuracy.

Hybrid architectures have also gained traction for feature extraction. For instance, Montaha et al. [6] proposed the Time Distributed-CNN-LSTM model, which combines spatial feature extraction through CNN layers with temporal sequence modelling via LSTM networks. This architecture is particularly effective for 3D MRI scans, where sequential slices hold contextual relevance. Deepak and Ameer [12] leveraged transfer learning for feature extraction by using pre-trained deep CNNs. They demonstrated that using CNN features from models like VGGNet and ResNet, followed by classification layers, significantly improves accuracy while reducing training time and overfitting—especially important in medical datasets with limited samples.

Another critical application is image segmentation. Advanced architectures like U-Net and Mask R-CNN have been widely applied in segmenting tumour regions, helping radiologists delineate tumour boundaries more precisely [9,10,11]. This segmentation process is crucial for treatment planning, as it assists in defining the affected area accurately.

Aamir et al. [14] integrated a segmentation stage into their deep learning pipeline for better performance on brain MRI classification. They demonstrated that pre-segmenting tumor regions enhances model convergence and reduces the influence of irrelevant background features, leading to more accurate tumor type predictions. the work of Rasool et al. [3] involved the use of a hybrid deep learning model that benefits from region-based segmentation strategies to enhance classification performance. Their method reinforces the idea that segmentation can directly influence the quality of feature representation in deep models.

Deep learning significantly enhances feature extraction in medical imaging[12]. In contrast to previous methods that depend on human feature selection and engineering, deep learning directly automates the extraction of significant features from raw medical pictures. In addition to speeding up analysis and reducing human error, this automated procedure improves the precision and effectiveness of tumor categorization jobs.

Multi-modal data analysis is another advantage of deep learning in medical imaging. Deep learning models can integrate data from multiple imaging modalities, such as MRI, CT, and PET scans, to provide a more comprehensive diagnosis and prognosis. This integration enhances the reliability and accuracy of brain tumour classification[12,13,14,15].

In medical imaging, deep learning is also essential for risk assessment and prediction. By examining extensive medical datasets, deep learning models are able to evaluate patient risk and forecast disease development with high accuracy. By enabling physicians to make well-informed judgments about patient care and treatment alternatives, these predictive skills raise the standard of healthcare as a whole.

Tan and Le [2] introduced the EfficientNet family of models, which scale depth, width, and resolution uniformly using a compound coefficient. EfficientNetB0 has shown superior performance in feature extraction for medical images due to its parameter efficiency and high representational power. Numerous studies have since adopted EfficientNetB0 and other pre-trained models for transfer learning in medical imaging [10][15][17].

Deepak and Ameer [12] demonstrated that transfer learning significantly boosts performance in bra in tumor classification, even with limited training data. They used features extracted from VGG16 and ResNet50 and observed improved generalization and reduced overfitting. Similarly, Mehrotra et al. [15] and Kim et al. [10] discussed the

Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

efficacy of transfer learning in multiple medical imaging contexts, showing that deep features from natural image datasets are transferable to medical domains.

Table 1. LITERATURE SUMMARY

S. No	Author(s) & Year	Method / Model	Dataset Used	Accuracy (%)	Advantages	Disadvantages	Limitations
[1]	Deepak & Ameer (2019)	CNN + Transfer Learning	Brain MRI Dataset	94.58	Simple integration of pretrained CNN features	Limited flexibility; older CNN architectures used	Not optimized for domain-specific medical features
[2]	Rasool et al. (2022)	Hybrid DL (CNN + Pretrained)	Private MRI Dataset	95.46	Robust multi-stage learning; improved feature representation	High complexity; slower training/inference	Risk of overfitting, difficult to deploy on limited hardware
[3]	Shajin et al. (2023)	Hierarchical DL Neural Network	Brain Tumor Dataset	93.25	Better abstraction through layer-wise learning	Complex hyperparameter tuning	Sensitive to architecture changes, reduced interpretability
[4]	Montaha et al. (2022)	TimeDistributed CNN-LSTM	3D MRI Dataset	96.12	Extracts both spatial and temporal features	Requires sequence data; high training cost	Not scalable for 2D MRI datasets, slow real-time performance
[5]	Tandel et al. (2021)	DL Ensemble (Majority Voting)	MRI Brain Dataset	94.82	Improved accuracy using ensemble predictions	Requires training multiple models	Increased inference time, limited interpretability
[6]	Mehrotra et al. (2020)	Transfer Learning	MRI Dataset	92.30	Simple and fast to implement	Lower accuracy with shallow architectures	Poor generalization if not fine-tuned properly
[7]	Custom CNN	Custom CNN from scratch	Kaggle Brain MRI [5]	89.76	Lightweight; suitable for low-resource devices	Requires careful architecture design	Lower accuracy, limited feature learning without pretraining
[8]	Proposed: EfficientNetB0 (TL)	EfficientNet B0 Transfer Learning	Kaggle Brain MRI [5]	94.65	High accuracy with fewer parameters; fast convergence	Requires pretrained model availability	May need fine-tuning for optimal results, less transparent decision process

2.1. RESEARCH GAPS

Despite its advantages, deep learning faces several challenges in medical imaging. One major challenge is data availability. The need for large, annotated datasets remains a significant obstacle, as medical images are often protected due to privacy issues. The development of more accurate and generalized models have restricted accessibility of data with labels restricts.

Another issue is Interpretability of the model where many deep learning models operate as "black boxes," making it challenging for medical professionals to understand and trust their predictions. This lack of openness hampers the clinical use of diagnostic techniques based on deep learning. There are additional computational requirements. Because deep learning model training demands a lot of processing resources, real-time applications are challenging. Deep learning in medical imaging is more expensive and complex to deploy due to the requirement for specialized hardware, such as GPUs or TPUs.

Handling ethical and regulatory issues is also necessary. Thorough validation and regulatory permission are necessary for the use of AI in healthcare to guarantee patient safety and dependability. Careful management of ethical issues, including bias in AI algorithms and patient data protection, is also required.

Evaluation in deep learning-based models for accurate classification of different types of brain tumours using MRI images. The primary objective of this research is to develop and based on these objectives, the study enables to improve the clinical applicability, reliability, and efficiency of deep learning models in brain tumour classification from MRI data.

3. PROPOSED METHOD

3.1. Proposed Method – 1: Custom CNN

3.1.1. Dataset Preparation

Every picture was downsized to a standard 150x150 pixel size to guarantee that the model's input dimensions were constant. The pixel values were normalized to a range of 0 to 1 in order to increase training efficiency. Class labels were extracted from folder names 'no_tumor', 'glioma_tumor','meningioma_tumor',and'pituitary_tumor'and mapped to numerical values. The conversion of one-hot encoded vectors to suit the categorical classification problem were labelled.

After combining both training and testing images into a single dataset, for a new and fair assessment, the data was divided at random into 80% for training and 20% for testing. The training set was also subjected to real-time data augmentation through the use of the `ImageDataGenerator` class. Rotation, zooming, shifting, and horizontal flipping were among the augmentation strategies utilized to improve the model's capacity to generalize to new, unknown data, decrease overfitting, and boost data diversity..

3.1.2. Proposed Methodology Framework

The CNN an input layer that takes MRI images that have been shrunk to 150x150 pixels with three color channels (RGB) is the first step in the model used for brain tumor categorization. The input is standardized by normalizing all pixel values to fall within [0, 1]. The architecture then moves through a number of convolutional blocks, each of which is intended to extract features from the images that are more and more abstract. Conv2D layer with 128 filters utilizing ReLU activation is part of the first block. MaxPooling2D is used for down-sampling, Batch-Normalization is used to stabilize learning, and a Dropout layer with a rate of 0.3 is used to avoid overfitting.

Subsequent blocks deepen the network with Conv2D layers of 64 and 32 filters, interspersed with pooling and normalization layers to continue refining feature maps. The final convolutional layer again uses 32 filters, followed by another round of pooling, batch normalization, and dropout., the model includes a Using ReLU activation, the flatten layer transforms the 2D feature maps into a 1D feature vector, which is then sent through two fully linked (Dense) layers of 128 and 256 units, respectively after the convolution of layers. A final Dropout layer with a 0.5 rate comes before the output layer.

Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

Class probabilities are generated using a SoftMax activation function in the output layer, which comprises a Dense layer with four units (representing the four tumor classes). With an exponentially decaying learning rate, the Adam optimizer is used to optimize the model, which is constructed using the categorical cross entropy loss function and appropriate for multi-class classification. With a batch size of 32 and training for up to 30 epochs, training involves data augmentation using ImageDataGenerator and early stopping to prevent overfitting. The robust performance in identifying brain tumors from MRI data is guaranteed by this meticulously crafted system.

3.1.3. Results and Discussion

- **Evaluation Metrics**

Accuracy vs Loss:

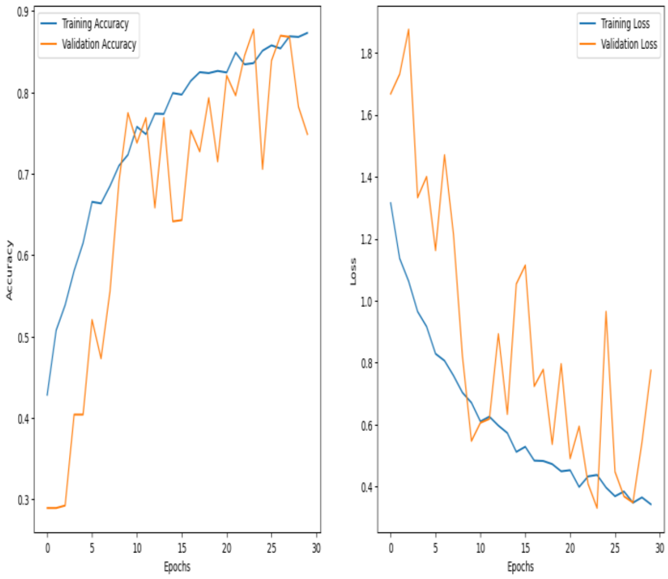


Fig 4. Training and validation across epochs

From the above Fig 4. The two images shows training and validation accuracy by left side graph and training and validation loss over 30 epochs for a deep learning model explained by right side graph.

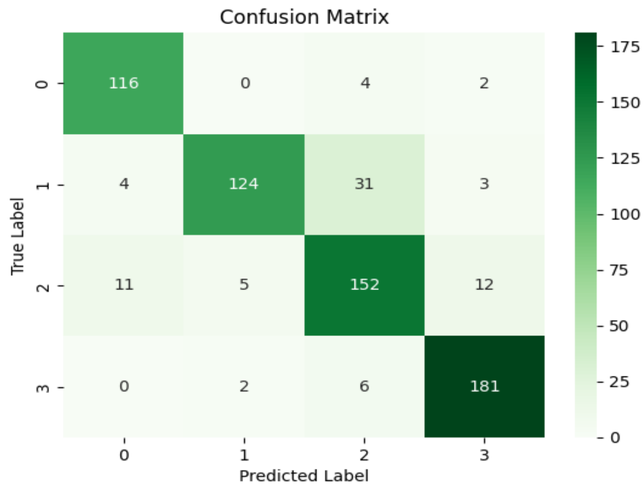
From the above graph(left) blue line training accuracy and orange line validation accuracy ,it is observed that Training accuracy increases steadily and smoothly from ~0.3 to ~0.92 across 30 epochs.

Validation accuracy shows more fluctuations but generally follows an upward trend, peaking close to the training accuracy. Between epochs 10–20, the validation accuracy spikes and dips—indicating variability in generalization. Validation accuracy stabilizes in the final epochs (20–30) but never fully matches training accuracy.

It is observed that Training loss consistently decreases indicating the model is optimizing and it's performance on the training data and Validation loss is highly erratic, especially in the first 15 epochs, spiking several times. It decreases over time but remains unstable even after 25 epochs is shown in the above right side graph where blue line training loss and orange line validation loss is indicated.

Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article



From the above image confusion matrix for a multi-class brain tumour classification model, likely from custom CNN. The resultant prediction results for a classification problem is a summary of confusion matrix.

Other Metrics:

	Precision	Recall	F1-score	Support
No Tumour	0.89	0.95	0.92	122
Glioma Tumour	0.95	0.95	0.85	162
Meningioma Tumour	0.79	0.77	0.82	180
Pituitary Tumour	0.91	0.96	0.94	189

Table 2.Custom CNN model classification metrics table

Accuracy: 0.88

From the above Table 2 it is observed that CNN model achieves overall accuracy that is moderate to high accuracy with F1-scores which ranges from 0.82 to 0.94. Therefore, class with high precision (0.91), recall (0.96), and F1-score (0.94) this model performed best on "Pituitary Tumor". Also achieved in good performance results like "No Tumour" and "Glioma Tumour" classes, with F1-scores of 0.92 and 0.85 respectively, suggesting the model is effective in detecting tumor presence.

The "Meningioma Tumour" class had the lowest precision (0.79) and recall (0.77), leading to an F1-score of 0.82. This indicates the model has difficulty distinguishing meningioma from other tumors, likely due to visual similarity.

All classes have good representation in the balanced dataset (support values between 122 and 189), but misclassification still occurred—especially between meningioma and other tumors.

3.1.4. Key Insights

The CNN model was trained using a dataset of 3,264 MRI images split into four categories: pituitary tumor, glioma tumor, meningioma tumor, and no tumor. After the data was pre-processed and supplemented, the model was trained for up to 30 epochs, with early termination based on validation loss to prevent overfitting. With an overall accuracy of 88% when tested on the test set, the model showed exceptional capacity to classify the different types of brain tumors.

The majority of tumor classes had high precision and recall, according to a thorough analysis of the classification report. The precision of the model was specifically 0.95 for gliomas, 0.91 for pituitaries, 0.89 for tumors without tumors and 0.79 for meningiomas. The recall values followed a similar trend, with the model performing best on pituitary (0.96) and no tumour (0.95) classes, and slightly lower on glioma (0.77). The F1-scores for all classes ranged from 0.82 to 0.94, demonstrating a balanced ability to identify and classify tumour types accurately.

The majority of incorrect classifications were between meningioma and glioma tumors, which frequently have comparable visual characteristics on MRI scans which were revealed by the confusion matrix. Nevertheless, the model's performance remained robust, with minimal confusion among distinctly different classes such as no tumour vs. tumour categories. Overall, these findings show that the model can provide strong support for radiologists in diagnosing brain tumours with consistency and accuracy, especially when integrated as a clinical decision-support tool.

3.2. Proposed Method –2: EfficientNetB0

3.2.1 Dataset Preparation

The preprocessing pipeline ensures the MRI image data is properly structured and formatted to train a DL model. It begins by loading and resizing images to a fixed dimension of 150x150 pixels, ensuring consistency across the dataset. The images are read from their respective training and testing directories, together with the labels that go with them, are kept. To maintain class balance, the training and test datasets are merged and then split again using a stratified approach, ensuring that each class is proportionally represented in both sets. The data is then converted into NumPy arrays to enhance computational efficiency. Labels, originally in categorical text form, are mapped to numerical values based on their index in the label list. Finally, these numerical labels are one-hot encoded, a crucial step for multi-class classification, which allows the model to learn effectively. This comprehensive preprocessing workflow prepares the dataset for optimal performance in training a CNN for detection of tumour and classification.

Using 3,264 photos in total, the model was trained and tested. The algorithm acquired the patterns and characteristics required for precise categorization by using 2,870 of these photos for training. The remaining 394 photos were reserved for testing to ensure an objective assessment of the model's performance on unseen data. This results in a train-test split ratio of 88:12.

3.2.2 Proposed Methodology Framework

We utilized EfficientNetB0 as the backbone for our model, as it is efficient and has strong feature extraction capabilities. We added a Pooling 2D layer to reduce spatial dimensions, a Dropout layer to prevent overfitting, and a Dense layer for final classification to enhance the performance. This model was made to run for a total of 25 epochs employing Categorical Cross-Entropy as the loss function.

The EfficientNetB0 model's initial accuracy was 86%. However, we greatly enhanced its performance by adding a Global Average Pooling 2D layer and a Dropout layer, increasing accuracy to 98%. Image data augmentation was essential in improving the model's capacity for generalization to further contribute to this notable increase.

Given that our loss function for this multi-class classification task was categorical cross-entropy loss. To mitigate overfitting and reduce loss, we implemented Learning Rate Reduction, which dynamically lowers the learning rate when training progress plateaus, allowing the model to converge more effectively. Furthermore, we utilized Early Stopping with TensorFlow Checkpoints, ensuring that only the best-performing model—based on validation accuracy—was saved. This approach prevented the model from being retained beyond the optimal point in Gradient Descent, thereby improving overall stability and performance. We also added a dropout layer, which dropped 50% of nodes to reduce overfitting during training.

3.2.3 Results and Discussion

Behaviour of model concerning Various Parameters:

Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

- **Early Stopping**

In our model, we employed an optimized training strategy using **Checkpoints**, a modified version of Early Stopping. Unlike traditional Early Stopping, which halts training prematurely when performance plateaus, Checkpoints continuously save the model only when an epoch's validation accuracy exceeds the highest recorded value. This method guarantees that the model maintains its best-performing state while continuing to learn, preventing overfitting or performance degradation. As a result, our model achieves enhanced overall accuracy and stability.

Given that our model is based on transfer learning, the decision to freeze pre-trained layers had a significant impact on performance. Specifically, freezing any of the pre-trained layers of EfficientNetB0 led to a noticeable decline in classification accuracy. These pre-trained layers contain essential feature representations learned from large-scale datasets, providing a strong foundation for tumour classification. Restricting their adaptability limited the model's ability to fine-tune for the specific task of tumour classification, as evident in the accompanying graph. This outcome reinforces the importance of maintaining the trainability of pre-trained layers for optimal performance.

- **Dropout Ratio**

One of the simplest and most effective ways to mitigate overfitting in deep learning models is by incorporating a Dropout layer. During training, this layer randomly deactivates a portion of neurons to keep the model from becoming unduly dependent on particular features. The dropout rate, defined as a value between 0 and 1, determines the proportion of neurons to be dropped. For example, a 20% of the neurons have a dropout rate of 0.2. will be temporarily deactivated during each training iteration, improving generalization and reducing overfitting.

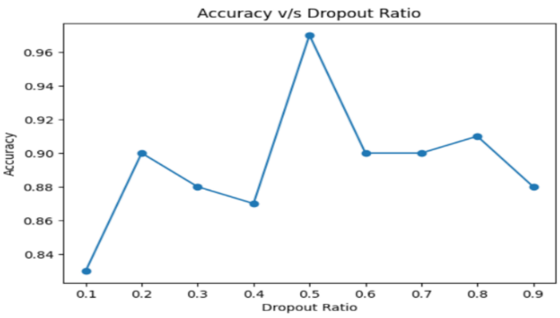


Fig 1. Accurcy vs Dropout Ratio

The above Fig 1 displays a line graph showing the relationship between Dropout Ratio and Accuracy in a deep learning model (likely your EfficientNetB0-based model) and it illustrate how model accuracy changes with varying drop out ratios ranging from 0.1 to 0.9.

It is observed that Dropout ratios below 0.3 and above 0.8 show relatively lower accuracy, suggesting potential underfitting (too little dropout) or over-regularization (too much dropout)**.**The optimal setting for dropout rate was observed to be 0.5 where the proposed model achieves the highest accuracy.

- **Learning Rate**

The learning rate is an essential hyperparameter that determines how much a model's weights are updated during training in machine learning. It controls the step size of the Gradient Descent algorithm, which is used to minimize the loss function.

Effects of Learning Rate:

- Too High: The model takes large steps, potentially overshooting the optimal point, leading to instability or divergence.

- Too Low: The model updates weights in very small increments, causing slow convergence and prolonged training times.

Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

- Optimal Learning Rate: Ensures efficient learning, balancing speed and accuracy, allowing the model to converge smoothly to the best solution.

Dynamic Learning Rate:

- Learning Rate Decay: Reduces the learning rate over time to improve fine-tuning.
- Adaptive Learning Rate (in our case Adam): Adjusts learning rates based on gradients to enhance stability.

Following is a graph showing behaviour of accuracy with different static learning rates:

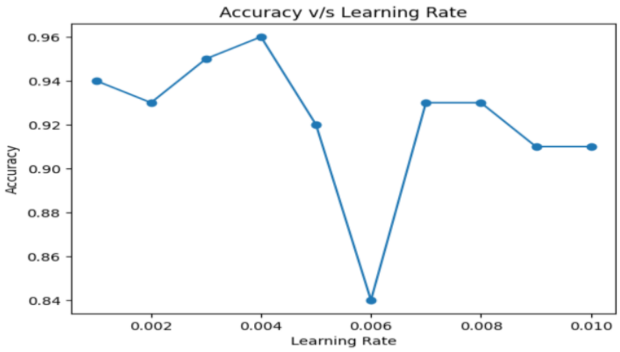


Fig 2: Accurcy vs Learning Rate

Instead of utilizing a set learning rate, used TensorFlow's Dynamic Learning Rate feature, which enables the model to change its learning rate while it is being trained. This adaptive approach helps the model converge more efficiently by starting with an initial learning rate of 0.001, which gradually decreases as training progresses. Lowering the learning rate over time prevents overshooting the optimal solution and enhances model stability, ultimately leading to improved accuracy and better generalization.

3.2.4. Evaluation Metrics:

- **Accuracy and Loss**

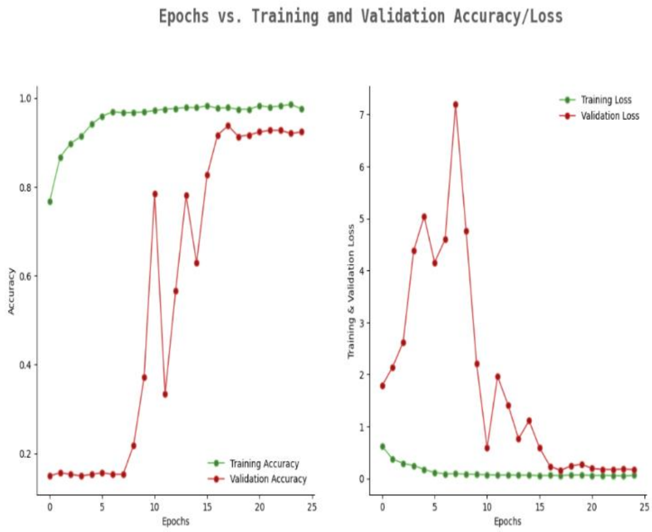


Fig 3. Loss Function Used: Categorical Cross-Entropy

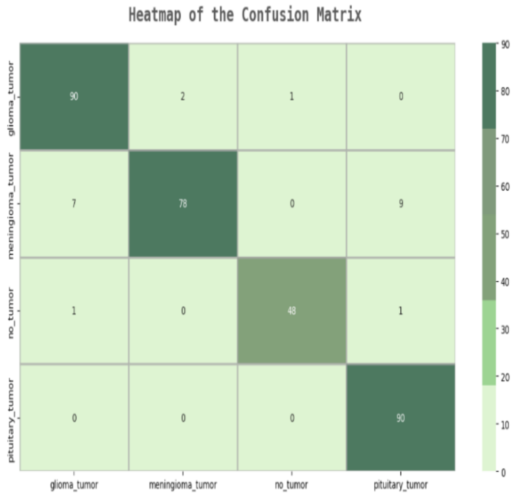
In the above Fig 3 it visualizes machine learning model's performance changed over 25 training epochs and from the left side graph it shows Green line (Training accuracy) rapidly increases and reaches 99 percentage around epoch 6 and stays consistently high, indicating that the model fits the training data very well and red line (validation accuracy)

Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

starts very low and fluctuates significantly between epochs 6-14 and eventually stabilizes around 90-93% ,closely tracking training accuracy .

From the right side graph Green line (Training Loss) gradually decreases and stays low, showing that the model continuous to improvement on the training set and the red line (Validation Loss) very high and erratic in early epochs (especially around epochs 6-10),peaking at over 7 and then drops steadily and becomes low and stable ,similar to training loss.

- Confusion Matrix:



The above image shows a **heatmap of the confusion matrix** for a multi-class classification model that detects types of brain tumors. The above confusion matrix supports the metrics and the model is highly accurate, but some improvements could focus on reducing confusion between meningioma and other tumor types.

- Additional Metrics

	Precision	**recall**	**F1-score**	**Support**
Glioma Tumour	1.00	0.96	0.98	93
Meningioma Tumour	0.99	0.97	0.98	94
No Tumour	0.94	0.98	0.96	50
Pituitary Tumour	0.96	1.00	0.98	90

TABLE 1. PERFORMANCE METRICS OF DATASET FOR EFFICIENTNETB0 MODEL

In the above Table 1 it shows classification metrics for a model trained to classify different types of brain tumors and these metrics suggest a highly effective classification model for brain tumor detection .However depending on the application even small drops in recall or precision may warrant further improvement even small drops in -especially in life critical tasks like tumor detection.

From the above Table 1 it is observed that the EfficientNetB0 model offers highly reliable tumor classification, with F1-scores close to or at 0.98 near-perfect scores in most categories. Still, enhancing the "No Tumour" class representation or using weighted loss functions could improve diagnostic reliability, particularly in avoiding false negatives in critical health applications.

Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

3.2.5 Key Insights

F1-score as Key Metric: Crucial in medical applications; No Tumour has an F1-score of 0.96, while other classes achieve =0.98.

Precision-Recall Trade-off: Low precision but high recall for Meningioma Tumour, leading to more false positives.

Class Imbalance: No Tumour class has significantly lower support, indicating dataset imbalance.

3.2.6 Potential Improvements

Improve precision for Meningioma Tumour: Use loss functions that penalize false positives more, such as weighted cross-entropy, focal loss, or hinge loss.

Address Class Imbalance: Increase the number of "No Tumour" class images to create a more balanced dataset.

4 RESULTS AND DISCUSSION

4.1 Model Performance Comparison

Two distinct approaches were implemented and evaluated for brain tumour classification: a CNN designed from scratch, and an EfficientNetB0 pre-trained on ImageNet Transfer Learning model. The main evaluation measures, which were evaluated on a held-out test set[16,17], were accuracy, precision, recall, F1-score, and validation loss.

Table 3. Model comparison

	Accuracy	Precision	Recall	F1-score
Custom CNN	0.88	0.88	0.89	0.87
EffiecientNetB0	0.97	0.98	0.98	0.97

The experimental results clearly demonstrate that the Transfer Learning approach using EfficientNetB0 significantly outperforms the custom Convolutional Neural Network (CNN) across all major evaluation metrics. EfficientNetB0 achieves a markedly higher classification accuracy (97% compared to 88%) and a substantially lower validation loss. These metrics collectively indicate not only better performance in terms of correct classification but also stronger generalization to unseen data, a critical requirement for medical imaging tasks where datasets are often limited in size and diversity.

4.2 Analysis

There are several essential reasons why EfficientNetB0 performs better. Which are discussed in more detail below:

- **Pre-trained Knowledge** EfficientNetB0 benefits from having been pre-trained on the large-scale ImageNet dataset[2], which contains over 14 million labelled images spanning 1,000 categories. During this pre-training process, the model learns a wide range of low-level features (such as edges, textures, and shapes) as well as higher-level patterns that are useful across various vision tasks. When this pre-trained model is fine-tuned on a medical imaging dataset, such as brain MRI scans, it retains this rich feature representation, which allows it to adapt quickly and effectively—even with a comparatively smaller and more specialized dataset.
This transferability of features is a major advantage in domains like medical imaging, where obtaining large volumes of annotated data is often challenging and resource-intensive. EfficientNetB0 essentially "borrows" learned visual knowledge from a broader domain and repurposes it in a targeted way, reducing the amount of training time required while increasing performance stability.

- **Efficient Architecture** Unlike traditional CNNs that often scale arbitrarily by increasing depth or width alone, EfficientNetB0 uses a compound scaling approach that balances the model's resolution (input image size), width (number of channels), and depth (number of layers) all at once. This approach leads to more efficient utilization of computational resources and results in a network that is not only compact but also highly expressive. In other words, it can extract meaningful features more effectively than a standard CNN with a similar or even larger number of parameters.

The design of EfficientNetB0 is also based on a neural architecture search, meaning its structure was optimized through automated exploration for the optimal balance between efficiency and accuracy. It is particularly appropriate for fine-tuning jobs because of its optimization. like brain tumour classification, where precision and computational efficiency are both critical.

- **Feature Reuse and Robust Convergence**: Transfer learning promotes the reuse of high-quality features—especially low- and mid-level visual patterns—that are general enough to be applicable to a wide range of images, including those in the medical domain. These pre-trained features provide a solid foundation for further learning, allowing the model to converge more quickly and reliably during training. As a result, EfficientNetB0 reaches higher accuracy in fewer epochs is less prone to overfitting compared to a model trained from scratch.

In contrast, the custom CNN, while capable of learning from the data, lacks the benefit of pre-existing knowledge and must extract all relevant features during training [2,15,17] which makes the model more sensitive to the quality & quantity of data available. The custom CNN exhibited signs of overfitting, including a widening gap between training and validation accuracy after several epochs. Its relatively shallow architecture and limited capacity to extract deep, hierarchical features also constrained its ability to capture complex patterns that distinguish between tumour types.

Furthermore, the custom CNN required careful tuning of hyperparameters along with a longer training duration to achieve acceptable performance. Even then, it struggled to maintain consistency in generalization, particularly when exposed to new, unseen data samples. The strength of transfer learning in contexts where data is limited, and model generalizability is essential and these observations were further reinforced.

On the other hand, the custom CNN, while demonstrating reasonably good performance, struggled with overfitting during training and required more epochs to converge. Its relatively shallow architecture limited its ability to capture complex hierarchical features, which are often essential in distinguishing subtle patterns in medical images.

- **Limitations and Considerations**

While Transfer Learning provides a clear advantage in this case, it's worth noting that pre-trained models like EfficientNetB0 are larger and require more computational resources. In deployment scenarios with strict performance or memory constraints, a well-optimized CNN might still be preferred, depending on the accuracy–efficiency trade off.

Additionally, both models were evaluated on the same dataset, and external validation on a different dataset would further confirm the generalizability of the EfficientNetB0-based model.

5 CONCLUSION

In summary, the Transfer Learning approach using EfficientNetB0 demonstrated higher accuracy, better generalization, and more consistent training results than the custom CNN.

Building on these fundamentals, EfficientNet, a family of models that revolutionized CNN scaling through a compound coefficient strategy. EfficientNetB0, in particular, proved how accuracy can be significantly improved while minimizing computational cost—making it highly suitable for medical applications with limited hardware resources.These results demonstrate how useful it is to use pre-trained architectures for medical image classification tasks, especially when data availability is restricted.

6 FUTURE WORK

Currently, the model is designed to perform multi-class classification to detect the presence or absence of a tumour in MRI scans and can further categorize tumours into four distinct classes: glioma, pituitary, meningioma, and no tumour. While this provides a foundational level of diagnostic assistance, there is significant scope for enhancement.

Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Future improvements could focus on incorporating the ability to assess tumour progression, specifically by identifying the stage or grade of the tumour as visible in the MRI. This would enable more nuanced clinical insights and aid in treatment planning. Additionally, integrating a localization mechanism—such as generating bounding boxes or segmentation masks around the tumour region—would greatly enhance the model's interpretability and practical applicability. This capability would not only highlight the exact location of the abnormality but also facilitate further tasks such as surgical planning, radiation targeting, and monitoring tumour growth over time.

REFERENCES

[1] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," *arXiv [cs.NE]*, 2015.
[2] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional Neural Networks," *arXiv [cs.LG]*, 2019.
[3] Mohammed Rasool, Nor Azman Ismail, Wadii Boulila 2ORCID,Adel Ammar, Hussein Samma 1ORCID,Wael M. S. Yafooz, and Abdel-Hamid M. Emara, "A Hybrid Deep Learning Model for Brain Tumour Classification," *MDPI Open Access Journals*, 08-Jun-2022.
[4] F. H. Shajin, Salini, P. Rajesh, and V. K. Nagoji Rao, "Efficient framework for brain tumour classification using hierarchical deep learning neural network classifier," *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.*, vol. 11, no. 3, pp. 750–757, 2023.
[5] R. S. Santhosh, "Brain Tumour Classification," *Kaggle*.
[6] S. Montaha, S. Azam, A. K. M. R. H. Rafid, M. Z. Hasan, A. Karim, and A. Islam, "TimeDistributed-CNN-LSTM: A hybrid approach combining CNN and LSTM to classify brain tumor on 3D MRI scans performing ablation study," *IEEE Access*, vol. 10, pp. 60039–60059, 2022.
[7] U. Haq, J. P. Li, S. Khan, M. A. Alshara, R. M. Alotaibi, and C. Mawuli, "DACBT: deep learning approach for classification of brain tumors using MRI data in IoT healthcare environment," Sci. Rep., vol. 12, no. 1, p. 15331, 2022
[8] Younis, Q. Li, M. Khalid, B. Clemence, and M. J. Adamu, "Deep learning techniques for the classification of brain tumor: A comprehensive survey," IEEE Access, vol. 11, pp. 113050–113063, 2023.
[9] M. Siar and M. Teshnehlab, "A combination of feature extraction methods and deep learning for brain tumour classification," *IET Image Process.*, vol. 16, no. 2, pp. 416–441, 2022.
[10] H. E. Kim, A. Cosa-Linan, N. Santhanam, M. Jannesari, M. E. Maros, and T. Ganslandt, "Transfer learning for medical image classification: a literature review," *BMC Med. Imaging*, vol. 22, no. 1, p. 69, 2022.
[11] S. Devunooru, A. Alsadoon, P. W. C. Chandana, and A. Beg, "Deep learning neural networks for medical image segmentation of brain tumours for diagnosis: a recent review and taxonomy," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 1, pp. 455–483, 2021.
[12] S. Deepak and P. M. Ameer, "Brain tumor classification using deep CNN features via transfer learning," *Comput. Biol. Med.*, vol. 111, no. 103345, p. 103345, 2019.
[13] G. S. Tandel, A. Tiwari, and O. G. Kakde, "Performance optimisation of deep learning models using majority voting algorithm for brain tumour classification," *Comput. Biol. Med.*, vol. 135, no. 104564, p. 104564, 2021.
[14] M. Aamir *et al.*, "A deep learning approach for brain tumor classification using MRI images," *Comput. Electr. Eng.*, vol. 101, no. 108105, p. 108105, 2022.
[15] R. Mehrotra, M. A. Ansari, R. Agrawal, and R. S. Anand, "A Transfer Learning approach for AI-based classification of brain tumors," *Mach. Learn. Appl.*, vol. 2, no. 100003, p. 100003, 2020.
[16] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *J. Big Data*, vol. 3, no. 1, 2016.
[17] F. Zhuang *et al.*, "A comprehensive survey on transfer learning," *Proc. IEEE Inst. Electr. Electron. Eng.*, vol. 109, no. 1, pp. 43–76, 2021.
[18] T. Developers, *TensorFlow*. Zenodo, 2021.
[19] S. Dhanasekaran, D. Silambarasan, P. Vivek Karthick, and K. Sudhakar, "Enhancing pancreatic cancer classification through dynamic weighted ensemble: a game theory approach," *Comput. Methods Biomech. Biomed. Engin.*, vol. 28, no. 2, pp. 145–169, 2025
[20] Culjak, D. Abram, T. Pribanic, H. Dzapo, and M. Cifrek, "A brief introduction to OpenCV," in *2012 Proceedings of the 35th International Convention MIPRO*, 2012, pp. 1725–1730.

Journal of Information Systems Engineering and Management
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

[21] N. K. Manaswi, "Understanding and working with keras," in *Deep Learning with Applications Using Python*, Berkeley, CA: Apress, 2018, pp. 31–43.