JPPIPA 12(3) (2026)

Jurnal Penelitian Pendidikan IPA

http://jppipa.unram.ac.id/index.php/jppipa/index

Explainable Brain Tumor Classification Using EfficientNet-B2 and Grad-CAM on MRI Dataset

Tino Saputra[1]*, Wahyu Purnama Magribi[1], Vitri Tundjungsari[2]

[1] Magister of Computer, Department of Computer Science, Esa Unggul University, Indonesia.
[2] Department of Computer Science, Esa Unggul University, Indonesia.

Received: January 05, 2026
Revised: March 10, 2026
Accepted: March 25, 2026
Published: March 31, 2026

Corresponding Author:
Tino Saputra
saputra.tino85@gmail.com

DOI: 10.29303/jppipa.v12i3.14179

Abstract: Brain tumors are life-threatening central nervous system disorders requiring early and accurate diagnosis for effective clinical management. Although MRI is the standard modality for detection, manual interpretation remains prone to inconsistency, particularly for complex cases such as glioma. This study proposes an explainable deep learning framework integrating EfficientNet-B2 with a threshold-based two-stage classification scheme and Grad-CAM interpretability analysis. In the first stage, a one-versus-rest binary classifier with an optimized threshold ($\tau = 0.20$) performs glioma detection; the second stage classifies remaining cases into meningioma, pituitary tumor, or normal. The dataset comprises 7,023 MRI images across four classes from a public Kaggle repository. Preprocessing includes CLAHE contrast enhancement, normalization, and augmentation. EfficientNet-B0 serves as the baseline. EfficientNet-B2 achieves 97.9% overall accuracy, outperforming the baseline (96.7%), with a glioma F1-score of 0.988 at the optimal threshold. Grad-CAM visualizations confirm the model focuses on anatomically relevant regions, enhancing transparency and clinical trustworthiness. The proposed framework demonstrates that combining architectural capacity, threshold-based inference, and explainability yields a reliable system for computer-aided brain tumor diagnosis.

Keywords: Brain Tumor; Deep Learning; EfficientNet-B2; Grad-CAM; MRI.

Introduction

Brain tumors are serious diseases of the central nervous system that can disrupt cognitive, sensory, and motor functions. Delayed diagnosis can significantly increase mortality risk, making early detection and accurate classification essential for effective treatment planning. According to global cancer statistics, the number of brain and central nervous system tumor cases continues to increase worldwide, highlighting the importance of reliable diagnostic methods (World Health Organization, 2023).

Magnetic Resonance Imaging (MRI) is widely used as the primary modality for brain tumor detection because it provides detailed soft tissue visualization without ionizing radiation (Radiological Society of North America (RSNA) & American College of Radiology (ACR), 2024). However, MRI interpretation still relies heavily on radiologists' expertise. Variations in tumor morphology and heterogeneous intensity patterns often make manual diagnosis difficult and may lead to inconsistent interpretations (Lundervold & Lundervold, 2019).

Recent advances in artificial intelligence, particularly deep learning, have significantly improved automated medical image analysis. Convolutional Neural Networks (CNNs) have demonstrated strong performance in MRI-based brain tumor classification by automatically extracting discriminative features from imaging data (Ahmed et al., 2022; Arlis et al., 2024; Babu Vimala et al., 2023; Rachmawanto et al., 2024). Transfer learning approaches using pre-trained architectures such as EfficientNet have further enhanced classification performance while maintaining computational efficiency (Alanazi et al., 2023; Iqbal et al., 2024; Islam et al., 2024; Pacal et al., 2024). Recent studies have also

---

How to Cite:
Saputra, T., Magribi, W. P., & Tundjungsari, V. (2026). Explainable Brain Tumor Classification Using EfficientNet-B2 and Grad-CAM on MRI Dataset. *Jurnal Penelitian Pendidikan IPA*, *12*(3), 142–151. https://doi.org/10.29303/jppipa.v12i3.14179

explored ensemble learning and transformer-based architectures to improve diagnostic accuracy in medical imaging tasks (Benzorgat et al., 2024; Kar & Singh, 2025; Sharma & Ameta, 2025).

Despite these advancements, several challenges remain. Classification performance often varies across tumor classes due to differences in tumor morphology and intensity characteristics. In addition, many deep learning models require significant computational resources, which may limit their implementation in real clinical environments (Alanazi et al., 2023; Sharma & Ameta, 2025).

Another important challenge is the interpretability of deep learning models. Although these models can achieve high predictive accuracy, their decision-making process is often difficult to explain, which may reduce trust in AI-assisted medical diagnosis systems (Linda Bi et al., 2019). Explainable Artificial Intelligence (XAI) methods such as Gradient-weighted Class Activation Mapping (Grad-CAM) have been introduced to visualize image regions that contribute to model predictions (Selvaraju et al., 2017). Several recent studies have shown that Grad-CAM can improve the transparency of brain tumor classification models (Agrawal & Chaki, 2025; Ahmad et al., 2024; Haque et al., 2025; Yan et al., 2023).

Based on these challenges, this study proposes a brain tumor classification approach using the EfficientNet-B2 architecture combined with a two-stage classification scheme and Grad-CAM-based interpretability analysis. The proposed method aims to improve classification stability across tumor classes while providing visual explanations that enhance the transparency of model predictions.

The objective of this study is to evaluate the performance of EfficientNet-B2 for multi-class brain tumor classification using MRI images, compare it with the EfficientNet-B0 baseline model, and analyze the interpretability of the classification results using Grad-CAM visualization. The results are expected to contribute to the development of a more reliable and interpretable computer-aided brain tumor diagnosis system.

Method

Time and Place of the Research

This research was conducted from September 2025 to January 2026 as part of an experimental study in the field of medical image analysis using deep learning techniques. The research activities included dataset preparation, image preprocessing, model development, training, and evaluation of the classification models. All computational experiments were carried out in a GPU-supported computing environment using Python-based deep learning frameworks. The experimental process was implemented through a computer-based research environment designed to support image processing and neural network training efficiently. The use of a GPU-enabled system was necessary to accelerate the training process of convolutional neural network architectures and to handle the large number of MRI images used in this study.

Research Design

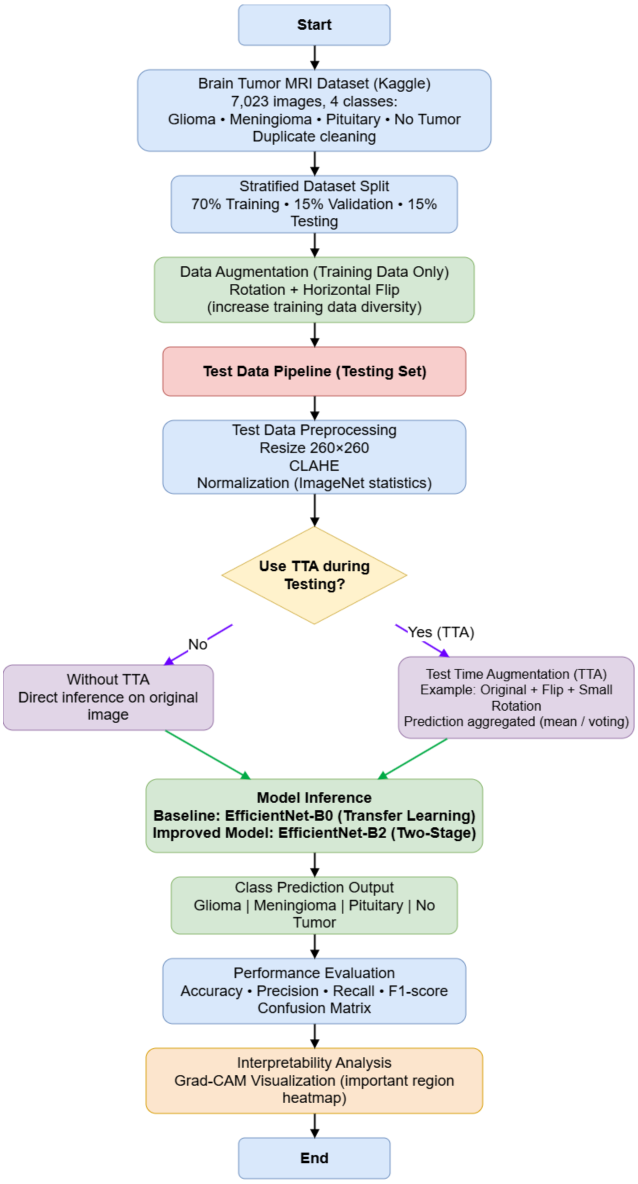

Figure 1. Flow Diagram Experiment

This study adopts a quantitative experimental approach aimed at developing and evaluating a deep learning based brain tumor classification model using Magnetic Resonance Imaging (MRI) images. This approach is selected to enable objective performance measurement through quantifiable evaluation metrics and direct comparison between the baseline model and the proposed model. Deep learning has become one of the most effective approaches in computer vision and medical image analysis due to its capability to automatically learn hierarchical feature representations from image data (Goodfellow et al., 2016; Litjens et al., 2017). The overall research workflow is illustrated in Figure 1.

This study utilizes data obtained from the Brain Tumor MRI Dataset, which is publicly available through the Kaggle platform (Table 1). Publicly available medical imaging datasets are widely used in deep learning research because they enable reproducibility and comparative evaluation across different models (Menze et al., 2015).
https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset.

Table 1. Data set Brain Tumor

Class	Image
Glioma Tumor	3.261
Meningioma Tumor	2.857
Pituitary Tumor	1.482
Normal (no_tumor)	427
4 Class	7.023 image

The dataset consists of brain MRI images categorized into four classes: glioma, meningioma, pituitary tumor, and normal (no tumor). Examples of MRI images from the dataset are illustrated in Figure 2.

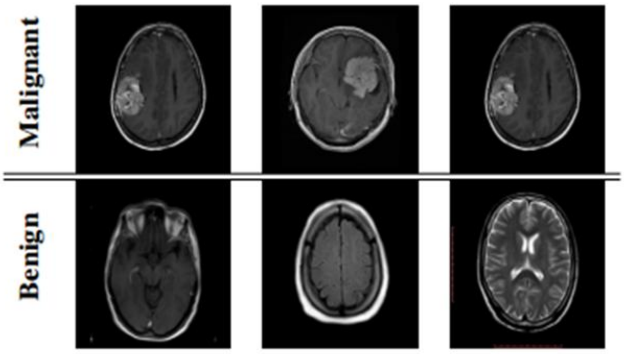

Figure 2. Example of MRI images from the Kaggle dataset.

The sampling technique applied is total sampling, where all available data are utilized as research samples because the dataset size is suitable for supervised deep learning experiments. The research variables include the input variable in the form of MRI brain images, the independent variable represented by the deep learning architectures used for classification, and the output variable which is the predicted tumor class.

Data collection is conducted using secondary data obtained from a publicly available medical imaging dataset on the Kaggle platform. The tools and materials used in this research include the Python programming language, TensorFlow/Keras deep learning framework, EfficientNet architectures for feature extraction and classification, and GPU-supported computing hardware for model training and evaluation. Convolutional neural networks have been widely applied in medical image analysis and tumor detection tasks due to their ability to capture spatial patterns and complex visual features in medical imaging data (Havaei et al., 2017; Ronneberger et al., 2015).

Research Procedure

The research procedure in this study consists of several sequential stages, starting from data preparation to model evaluation and interpretability analysis. The first stage is data collection. Brain MRI images are obtained from the Brain Tumor MRI Dataset available on the Kaggle platform. The dataset contains four classes of MRI images, namely glioma tumor, meningioma tumor, pituitary tumor, and normal brain images. These images serve as the primary input data for the classification experiment. Publicly available datasets are widely used in medical imaging research because they allow reproducibility and comparative evaluation across different deep learning models (Menze et al., 2015).

The second stage is image preprocessing. This step aims to improve data consistency and image quality before the training process. Preprocessing includes removing duplicate images, resizing all images to match the input size required by the neural network architecture, and applying intensity normalization to standardize pixel value distributions. In addition, contrast enhancement is performed using the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique to highlight structural differences and tumor boundaries in MRI images. Image preprocessing is an essential stage in medical image analysis because it helps improve the performance and stability of deep learning models when processing complex MRI data (Shen et al., 2017).

The third stage is data augmentation and dataset splitting. Data augmentation is applied to the training dataset using image transformations such as rotation, flipping, and scaling in order to increase data variability and reduce the risk of overfitting. Data augmentation techniques are commonly applied in deep learning-based medical imaging tasks to improve model generalization, especially when the dataset size is limited (Litjens et al., 2017). After preprocessing, the dataset is divided into training, validation, and testing sets while maintaining a balanced distribution across the four classes.

The fourth stage is model development and training. The classification model is developed using a transfer learning approach. EfficientNet-B0 is employed as the baseline model to obtain initial performance references, while EfficientNet-B2 is implemented as the proposed model due to its higher feature extraction capability while maintaining computational efficiency. EfficientNet introduces a compound scaling strategy that balances network depth, width, and input resolution to improve model performance and efficiency (Iqbal et al., 2024; Islam et al., 2024). Both models are initialized using ImageNet pre-trained weights and then fine-tuned to adapt to the characteristics of MRI images.

The fifth stage is the implementation of a two-stage classification scheme. In the first stage, a one-versus-rest classification approach is used to distinguish glioma from non-glioma images. Images predicted as non-glioma are then processed in the second stage to classify them into meningioma, pituitary tumor, or normal categories. This strategy is designed to improve sensitivity in detecting glioma tumors that often present complex morphological patterns. Several recent studies have explored multi-stage and hybrid deep learning architectures to improve classification performance in brain tumor detection tasks (Kar & Singh, 2025; Sharma & Ameta, 2025).

The final stage is model evaluation and interpretability analysis. Model performance is evaluated using several metrics, including accuracy, precision, recall, and F1-score. Confusion matrices are also used to analyze classification performance for each tumor class. In addition, model interpretability is examined using Gradient-weighted Class Activation Mapping (Grad-CAM), which visualizes the regions in MRI images that influence the model's classification decisions. Grad-CAM has been widely used to provide visual explanations for deep learning models in medical image analysis (Selvaraju et al., 2017). Recent studies have shown that explainable deep learning techniques can improve transparency and support clinical interpretation in brain tumor classification systems (Agrawal & Chaki, 2025; Ahmad et al., 2024).

Research Data Analysis

The analysis of research data in this study focuses on evaluating the performance of deep learning models for brain tumor classification using MRI images. After the model training process is completed, the classification results are evaluated using several quantitative performance metrics, including accuracy, precision, recall, and F1-score. These metrics are used to measure the model's ability to correctly classify MRI images into the four tumor categories, namely glioma, meningioma, pituitary tumor, and normal brain images (Babu Vimala et al., 2023). Accuracy represents the overall correctness of predictions, while precision and recall evaluate the model's ability to correctly identify positive cases. The F1-score is used to provide a balanced measurement between precision and recall.

In addition to these evaluation metrics, confusion matrix analysis is performed to provide a detailed understanding of the classification performance for each tumor class. The confusion matrix allows the identification of correct predictions and misclassifications across the four classes, which helps analyze the strengths and limitations of the classification models.

Furthermore, interpretability analysis is conducted using the Gradient-weighted Class Activation Mapping (Grad-CAM) method. Grad-CAM is applied to visualize the regions of MRI images that contribute to the classification decisions made by the model. This approach provides insights into whether the model focuses on clinically relevant anatomical areas when predicting tumor classes (Selvaraju et al., 2017; Yan et al., 2023). The interpretability analysis supports the reliability and transparency of the proposed deep learning model in medical image classification tasks.

Result and Discussion

Preparation Phase

Before the model performance evaluation is conducted, this study undergoes a preparation phase aimed at ensuring the readiness of both the data and the model for fair and representative testing. This phase is considered crucial, as the quality of classification results is strongly influenced by the initial condition of the data and the stability of the training process. Proper data preparation and preprocessing have been widely recognized as essential steps in improving the performance of deep learning models in medical image analysis tasks (Litjens et al., 2017; Shen et al., 2017).

At this stage, all MRI images obtained from the public dataset are first examined to eliminate duplicate data that could potentially introduce bias during training and evaluation. After data cleaning, the dataset is divided using a stratified split scheme to maintain a balanced distribution of each tumor class across the training, validation, and testing sets. This approach ensures that model performance is not affected by class imbalance.

The preprocessing steps applied include image resizing, intensity normalization, and contrast enhancement using Contrast Limited Adaptive Histogram Equalization (CLAHE), which aim to improve the readability of visual patterns in MRI images. Contrast enhancement has been shown to highlight tissue structure differences and tumor boundaries more clearly, particularly for the glioma class, which exhibits heterogeneous visual characteristics. In addition, limited data augmentation is applied to the training set to increase data variability and reduce the risk of overfitting, enabling the model to learn more generalized feature representations.

These visualizations demonstrate improved contrast and clearer tissue structures in MRI images, which are expected to support the model in learning relevant visual features during the training process.

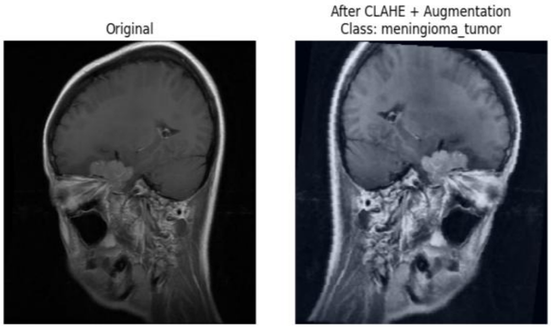

Figure 3. Example of before and after CLAHE

Model readiness is also evaluated through the initialization of pre trained weights and the gradual adjustment of training parameters. This process ensures that both the baseline model and the proposed model are in comparable conditions prior to performance evaluation. As a result, any differences observed in subsequent results can be directly attributed to variations in model architecture and classification strategy, rather than to data quality issues or differences in model configuration.

Evaluation Model (EfficientNet-B0)

The EfficientNet-B0 model is used as the baseline to obtain an initial overview of multi class brain tumor classification performance based on MRI images. The evaluation is conducted using test data that are not involved in the training process, so that the results reflect the model's generalization capability on unseen data.

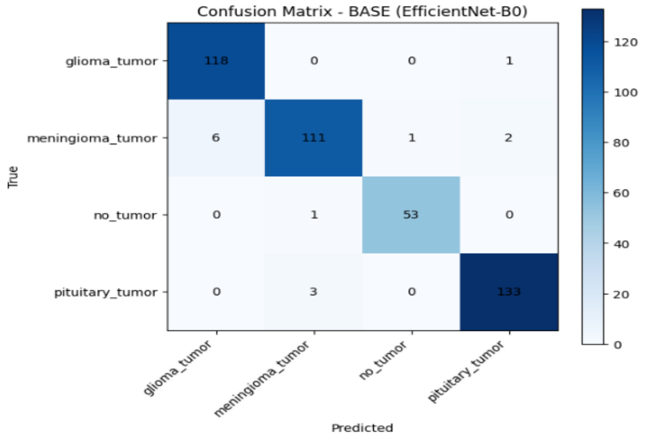

Figure 4. Confusion Matrix EffecienNet B0

Table 2. Evaluation ResultsEfficiennet B0

Class	Precision	Recall	F1-Score	Support
Glioma Tumor	0.952	0.992	0.971	119
Meningioma Tumor	0.965	0.925	0.945	120
No Tumor	0.982	0.982	0.982	54
Pituitary Tumor	0.978	0.978	0.978	136
Overall Accuracy	–	–	0.967	429

Based on the evaluation results of the EfficientNet B0 baseline model using the confusion matrix, the model demonstrates very strong classification performance across all brain tumor classes. Out of a total of 429 test images, the model correctly classifies 415 images, resulting in an overall accuracy of 96.7 percent. These results indicate that EfficientNet B0 has a strong generalization capability in distinguishing the four MRI image classes, namely glioma tumor, meningioma tumor, pituitary tumor, and no tumor.

At the class level, the model shows consistent performance with high precision, recall, and F1 score values. The glioma tumor class achieves a recall of 0.992, indicating that nearly all glioma images are correctly identified. Meanwhile, the meningioma tumor and pituitary tumor classes reach F1 scores of 0.945 and 0.978, respectively. The no tumor class also demonstrates stable performance with an F1 score of 0.982, confirming the model's ability to accurately differentiate normal brain images from tumor cases. The number of misclassifications is relatively small and does not indicate a dominant bias toward any particular class. These findings confirm that EfficientNet B0, although used as a baseline model, already provides highly competitive performance for MRI based brain tumor classification.

Evaluasi Model (EfficientNet-B2)

The EfficientNet-B2 model is evaluated as the proposed model to assess the effect of increased architectural capacity on brain tumor classification performance.

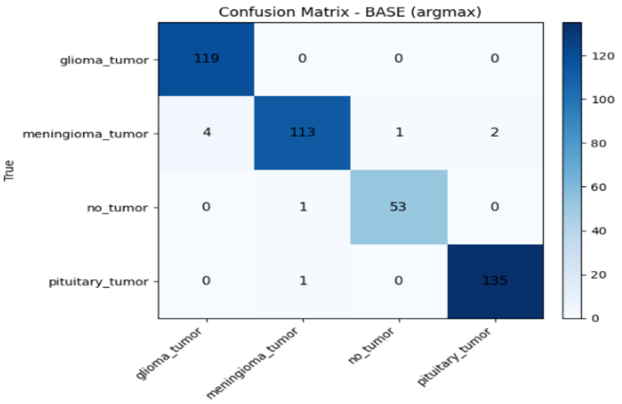

Figure 5. Confusion Matrix EffecienNet B2

Table 3. Evaluation Results Efficiennet B2

Class	Precision	Recall	F1-Score	Support
Glioma Tumor	0.967	1.000	0.983	119
Meningioma Tumor	0.979	0.942	0.960	120
No Tumor	0.981	0.982	0.982	54
Pituitary Tumor	0.985	0.993	0.989	136
Overall Accuracy	–	–	0.979	429

Based on the evaluation results using a confusion matrix with an argmax prediction scheme, the EfficientNet-B2 model demonstrates very high classification performance across all brain tumor classes. Out of a total of 429 test images, 420 images are correctly classified, resulting in an overall accuracy of 97.9 percent. These results indicate that the increased architectural capacity of EfficientNet-B2 has a positive impact on the model’s generalization capability compared to the baseline approach.

At the class level, the model exhibits consistent performance with high precision, recall, and F1-score values. The glioma tumor class achieves a recall of 1.000, indicating that all glioma images are correctly detected without errors. The meningioma tumor and pituitary tumor classes also show stable performance, with F1-scores of 0.960 and 0.989, respectively. Meanwhile, the no tumor class records an F1-score of 0.982, reflecting the model’s ability to accurately distinguish normal brain images from tumor cases. The number of misclassifications is relatively small and does not indicate any significant bias toward a particular class. These findings confirm that EfficientNet-B2 with a direct argmax prediction scheme provides highly competitive classification performance on the test dataset.

The training process employs cross-entropy loss and the Adam optimizer with a batch size of 32 for 10 epochs, while the learning rate is adjusted using the same experimental settings as those applied to the baseline model. By maintaining consistent training configurations, the observed performance differences can be directly attributed to variations in architectural capacity and the classification strategy employed.

EfficientNet-B2 dengan Threshold (τ) pada Two-Stage Classification

In the third scenario, the EfficientNet-B2 model is further developed by incorporating a threshold mechanism (τ) within a two-stage classification framework aimed at improving glioma tumor detection. In the first stage, EfficientNet-B2 functions as a one-vs-rest glioma detector that produces a binary probability for the glioma class. The classification decision is determined based on the threshold value (τ), where images with glioma probability ≥ τ are directly classified as glioma, while images with probability < τ proceed to the second stage for classification into meningioma tumor, pituitary tumor, or no tumor using a multiclass classification scheme. The threshold value is evaluated across several ranges and selected based on the best performance on the validation dataset. Experimental results indicate that the use of a threshold significantly improves glioma detection sensitivity while maintaining consistent performance across the remaining classes. Training parameters in this scenario are kept identical to those used for EfficientNet-B2 to ensure that the observed performance improvements arise from the threshold-based inference strategy rather than changes in training configuration.

Table 4. Evaluation ResultsEfficiennet B2 with tunning threshold (glioma)

Threshold (τ)	Precision	Recall	F1-Score
0.20	1.000	0.976	0.988
0.25	1.000	0.967	0.983
0.30–0.50	1.000	0.959	0.979
0.55	1.000	0.951	0.975

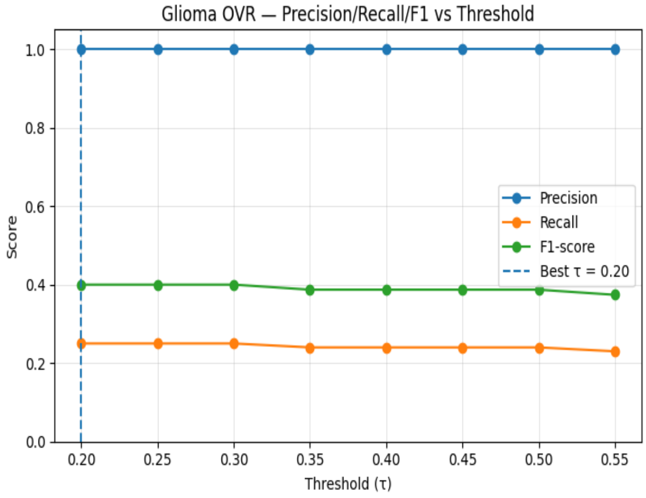

Figure 6. Evaluation Curve threshold

Based on the experimental results, an optimal threshold value of τ = 0.20 is obtained, achieving the highest F1-score of 0.988. Therefore, this threshold is selected as the decision parameter for the glioma detection stage in the two-stage classification scheme.

Performance Comparison between EfficientNet-B0 and EfficientNet-B2

A performance comparison between the EfficientNet-B0 and EfficientNet-B2 models is conducted to analyze the impact of increased architectural capacity on MRI-based brain tumor classification performance. Both models are evaluated using identical training configurations, allowing the observed performance differences to be directly attributed to variations in model architecture and the inference strategies employed (Findia Jiven & Rumini, 2025; Gottipati & Thumbur, 2024; Musa, 2024).

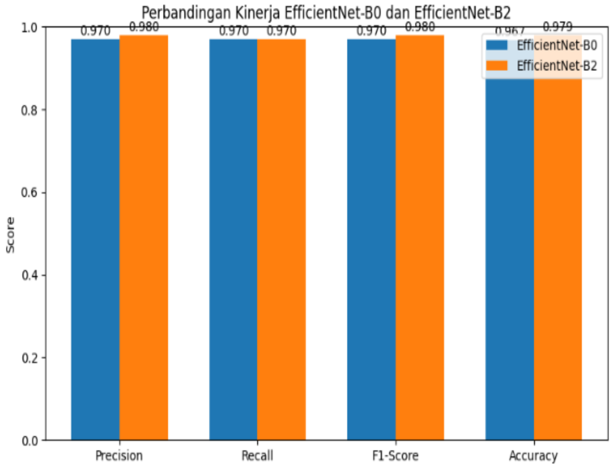

Figure 7. Comparison diagram of evaluation results

Based on the evaluation results, EfficientNet-B0, as the baseline model, demonstrates highly competitive performance with an overall accuracy of 96.7%. The model is able to classify all classes effectively, including the glioma tumor class, which achieves a high recall value. This indicates that EfficientNet-B0 already possesses strong feature representation capability for the dataset used. Nevertheless, in certain cases, the model still exhibits potential prediction ambiguity between tumor classes that share similar visual characteristics.

In contrast, EfficientNet-B2 shows a more consistent performance improvement, achieving an overall accuracy of 97.9% under the argmax prediction scheme. This improvement reflects the advantage of a deeper and more complex architecture in capturing visual patterns from MRI images. In particular, EfficientNet-B2 enhances prediction stability for the pituitary tumor class and reduces the overall number of misclassifications compared to the baseline model.

Beyond the improvement in global accuracy, the main advantage of EfficientNet-B2 becomes evident when it is combined with the threshold-based two-stage classification mechanism. By applying the optimal threshold value of τ = 0.20, EfficientNet-B2 significantly increases the sensitivity of glioma detection without compromising the overall system accuracy. This strategy is not applied to EfficientNet-B0, highlighting that EfficientNet-B2 offers greater flexibility for integration into advanced inference schemes that focus on critical classes.

Overall, the comparative results indicate that EfficientNet-B0 serves as a lightweight and efficient baseline model with strong performance, while EfficientNet-B2 provides additional advantages in terms of accuracy, prediction stability, and compatibility with threshold-based inference strategies. Therefore, EfficientNet-B2 is considered more suitable for clinical application scenarios that require high reliability, particularly for detecting high-risk tumor classes such as glioma.

Model Interpretability Analysis Using Grad-CAM

In addition to the performance improvements demonstrated by the evaluation metrics, this study also conducts an interpretability analysis to examine how the EfficientNet-B2 model makes decisions when classifying brain tumor MRI images (Hassan & Al-Zahrani, 2024; Iftikhar, 2025; Kakon et al., 2025). This analysis is essential to ensure that the model's decisions are based on medically relevant visual features rather than coincidental or non-informative patterns. For this purpose, the Gradient-weighted Class Activation Mapping (Grad-CAM) method is employed to visualize image regions that contribute most significantly to the model's predictions.

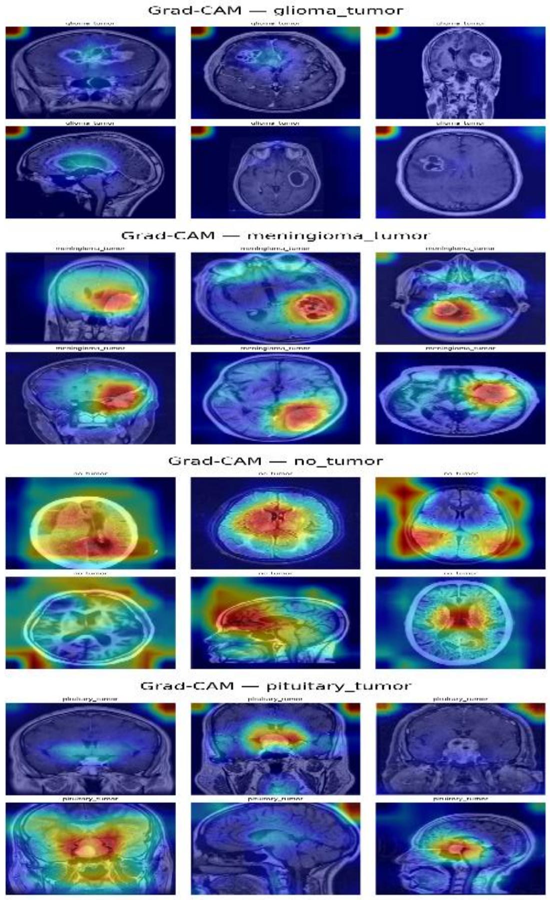

Figure 8. GradCam Visualization

Figure 8 presents the Grad-CAM visualizations for several MRI image examples from each class, namely glioma tumor, meningioma tumor, pituitary tumor, and no tumor. Regions with higher color intensity indicate areas that have the greatest influence on the model's decision-making process.

For the meningioma tumor class, Grad-CAM activations are clearly localized around the tumor mass. This indicates that the model is able to accurately identify the visual characteristics and boundaries of meningioma tumors. The focused activation pattern is consistent with the high precision and recall values obtained for this class, suggesting stable model performance.

In the pituitary tumor class, the dominant activation areas are concentrated in the central region of the brain, particularly in the sellar region. This activation focus corresponds well with the anatomical location of the pituitary gland, indicating that the model has successfully learned clinically relevant spatial representations for pituitary tumor detection.

In contrast, Grad-CAM activations for glioma tumor images tend to be more diffuse and less sharply localized. This pattern reflects the infiltrative nature of gliomas, which often have unclear boundaries and heterogeneous intensity characteristics. The dispersed activations help explain why glioma remains the most challenging class to classify and why its performance stability is generally lower compared to other tumor classes.

For the no tumor class, Grad-CAM results show relatively global and symmetrical activations across the brain area. This pattern suggests that the model bases its decision on the absence of specific pathological features rather than the presence of localized abnormal regions. This finding is consistent with the high recall achieved for the no tumor class, indicating reliable identification of normal images.

Overall, the interpretability analysis using Grad-CAM reinforces the quantitative evaluation results obtained earlier. The EfficientNet-B2 model not only demonstrates strong classification performance but also focuses on medically relevant regions of MRI images. Consequently, this approach enhances model transparency and reliability, supporting its potential application as an artificial intelligence–based decision support system for brain tumor diagnosis.

Conclusion

This study proposes an explainable deep learning framework for MRI-based brain tumor classification using EfficientNet architecture. Experimental results show that EfficientNet-B2 outperforms the EfficientNet-B0 baseline, achieving higher overall accuracy, precision, recall stability, and F1-score across most tumor classes. The findings indicate that increased architectural capacity improves the ability of the model to capture complex visual patterns present in brain MRI images. To address the issue of class performance imbalance, particularly in glioma detection, this study introduces a threshold-based two-stage classification strategy that improves detection sensitivity while maintaining stable performance for other tumor classes. In addition to performance improvements, interpretability analysis using Grad-CAM demonstrates that the proposed model focuses on anatomically relevant regions of MRI images, providing visual explanations that enhance the transparency of the classification process. Overall, the proposed framework combines improved classification accuracy, enhanced glioma detection stability, and explainable predictions within a single efficient architecture, contributing to the development of reliable and interpretable computer-aided diagnostic systems for brain tumor analysis.

Acknowledgments

Expression of Gratitude, throughout the completion of this research, the researcher received both moral and material support from various parties

Author Contributions

The primary author, Tino Saputra, contributed to research design, development and implementation of the data processing program, experimental execution, result analysis, and manuscript preparation. Wahyu Purnama Maghribi, as the second author and corresponding author, was responsible for application development as an implementation of the proposed system. Vitri Tunjdjungsari served as the research supervisor, providing scientific guidance, conducting substantive reviews, and editing the manuscript, as well as reviewing the initial draft and offering constructive feedback to improve the quality of the article.

Funding

No external funding

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper

References

Agrawal, A., & Chaki, J. (2025). CerebralNet meets Explainable AI: Brain tumor detection and classification with probabilistic augmentation and a deep learning approach. *Biomedical Signal Processing and Control*, *110*, 108210. https://doi.org/10.1016/j.bspc.2025.108210

Ahmad, P., Qamar, S., & Rizvi, S. S. (2024). Grad-CAM++ based interpretability for brain tumor detection in MRI using deep convolutional neural networks. *Journal of Digital Imaging*, *37*(1), 88–104. https://doi.org/10.1007/s10278-023-00912-x

Ahmed, M., Khalifa, F., Moustafa, H. E. D., Saleh, G. A.,

& AbdElhalim, E. (2022). A deep learning based system for accurate diagnosis of brain tumors using T1-w MRI. *Indonesian Journal of Electrical Engineering and Computer Science*, *28*(2), 1192–1202. https://doi.org/10.11591/ijeecs.v28.i2.pp1192-1202

Alanazi, M. F., Ali, M. U., Khan, S. J., & Zafar, A. (2023). EfficientNetV2-S for brain tumor classification using MRI: A transfer learning approach. *IEEE Access*, *11*, 24561–24575. https://doi.org/10.1109/ACCESS.2023.3254125

Arlis, S., Putra, M. R., & Yanto, M. (2024). Effectiveness of VGG19 in deep learning for brain tumor detection. *Indonesian Journal of Electrical Engineering and Computer Science*, *35*(2), 1210–1218. https://doi.org/10.11591/ijeecs.v35.i2.pp1210-1218

Babu Vimala, B., Srinivasan, S., Mathivanan, S. K., Mahalakshmi, Jayagopal, P., & Dalu, G. T. (2023). Detection and classification of brain tumor using hybrid deep learning models. *Scientific Reports*, *13*(1), 1–17. https://doi.org/10.1038/s41598-023-50505-6

Benzorgat, N., Xia, K., & Benzorgat, M. N. E. (2024). Enhancing brain tumor MRI classification with an ensemble of deep learning models and transformer integration. *PeerJ Computer Science*, *10*, e2425. https://doi.org/10.7717/peerj-cs.2425

Findia Jiven, F., & Rumini, R. (2025). MRI classification of brain tumors using EfficientNetB0 feature extraction and machine learning methods. *Journal of Applied Informatics and Computing (JAIC)*, *9*(1), 45–53. https://jurnal.polibatam.ac.id/index.php/JAIC/article/view/10363

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. https://www.deeplearningbook.org

Gottipati, S. B., & Thumbur, G. (2024). Multi-modal fusion deep transfer learning for accurate brain tumor classification using magnetic resonance imaging images. *Indonesian Journal of Electrical Engineering and Computer Science*, *34*(2), 825–834. https://doi.org/10.11591/ijeecs.v34.i2.pp825-834

Haque, R., Islam, M. S., & Islam, M. R. (2025). Explainable deep stacking ensemble model for accurate and transparent brain tumor diagnosis. *Computers in Biology and Medicine*, *191*, 110166. https://doi.org/10.1016/j.compbiomed.2025.110166

Hassan, M., & Al-Zahrani, A. (2024). Unfolding Explainable AI for brain tumor segmentation. *Neurocomputing*, *599*, 128058. https://doi.org/10.1016/j.neucom.2024.128058

Havaei, M., Davy, A., Varley, david wade, Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P. mac, & Larochelle, H. (2017). Brain Tumor Segmentation with Deep Neural Networks. *Medical Image Analysis*, *35*, 18–31. https://doi.org/10.1016/j.media.2016.05.004

Iftikhar, S. (2025). Explainable CNN for brain tumor detection and classification through XAI based key features identification. *Brain Informatics*, *12*(1), 10. https://doi.org/10.1186/s40708-025-00257-y

Iqbal, A., Jaffar, M. A., & Jahangir, R. (2024). Enhancing Brain Tumour Multi-Classification Using Efficient-Net B0-Based Intelligent Diagnosis for Internet of Medical Things (IoMT) Applications. *Information MDPI (Switzerland)*, *15*(8). https://doi.org/10.3390/info15080489

Islam, M. M., Talukder, M. A., & Hasan, M. K. (2024). BrainNet: Precision brain tumor classification with optimized EfficientNet architecture. *International Journal of Intelligent Systems*, *2024*(1), 3583612. https://doi.org/10.1155/2024/3583612

Kakon, S. C., Sazid, Z. Al, Begum, I. A., Samad, A., & Hosen, A. S. M. S. (2025). *Explainable Deep Ensemble Meta-Learning Framework for Brain Tumor Classification Using MRI Images*. 1–29. https://doi.org/10.3390/cancers17172853

Kar, S., & Singh, P. K. (2025). MBTC-Net: Multimodal brain tumor classification from CT and MRI scans using deep neural network with multi-head attention mechanism. *Medicine in Novel Technology and Devices*, *27*, 100382. https://doi.org/https://doi.org/10.1016/j.medntd.2025.100382

Linda Bi, W., Hosny, A., B, M. schabath, & L, M. (2019). Artificial Intelligence in Cancer Imaging: Clinical Challenges and Applications. *CA: A Cancer Journal for Clinicians*, *69*(2), 127–157. https://doi.org/10.3322/caac.21552

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., & Francesco Ciompi. (2017). A Survey on Deep Learning in Medical Image Analysis. *Medical Image Analysis*, *42*, 60–88. https://doi.org/10.1016/j.media.2017.07.005

Lundervold, A. S., & Lundervold, A. (2019). An Overview of Deep Learning in Medical Imaging Focusing on MRI. *Zeitschrift Für Medizinische Physik*, *29*(2), 102–127. https://doi.org/10.1016/j.zemedi.2018.11.002

Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., & Kirby, J. (2015). The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Transactions on Medical Imaging*, *34*(10), 1993–2024. https://doi.org/10.1109/TMI.2014.2377694

Musa, M. N. (2024). MRI-Based Brain Tumor Classification using ResNet-50 and Optimized

Softmax Regression. *Jurnal Infotel*, *16*(3), 598–614. https://doi.org/10.20895/infotel.v16i3.1175

Pacal, I., Celik, O., Bayram, B., & Cunha, A. (2024). Enhancing EfficientNetv2 with global and efficient channel attention mechanisms for accurate MRI-Based brain tumor classification. *Cluster Computing*, *27*(8), 11187–11212. https://doi.org/10.1007/s10586-024-04532-1

Rachmawanto, E. H., Sari, C. A., & Isinkaye, F. O. (2024). A good result of brain tumor classification based on simple convolutional neural network architecture. *Telkomnika (Telecommunication Computing Electronics and Control)*, *22*(3), 711–719. https://doi.org/10.12928/TELKOMNIKA.v22i3.25863

Radiological Society of North America (RSNA), & American College of Radiology (ACR). (2024). *Magnetic Resonance Imaging (MRI)*. https://www.radiologyinfo.org/en/info/mri

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015)*, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *Proceedings of the IEEE International Conference on Computer Vision*, 618–626. https://doi.org/10.1109/ICCV.2017.74

Sharma, V. K., & Ameta, G. K. (2025). Hybrid 3D CNN–transformer model for early brain tumor detection with multi-modal magnetic resonance imaging. *Bulletin of Electrical Engineering and Informatics*, *14*(5), 3923–3934. https://doi.org/10.11591/eei.v14i5.11082

Shen, D., Wu, G., & Suk, H.-I. (2017). Deep Learning in Medical Image Analysis. *Annual Review of Biomedical Engineering*, *19*, 221–248. https://doi.org/10.1146/annurev-bioeng-071516-044442

World Health Organization. (2023). *Global Cancer Observatory: Brain and Central Nervous System Cancer Statistics*. https://gco.iarc.who.int/media/globocan/factsheets/cancers/31-brain-central-nervous-system-fact-sheet.pdf%7D

Yan, F., Chen, Y., Xia, Y., Wang, Z., & Xiao, R. (2023). An Explainable Brain Tumor Detection Framework for MRI Analysis. *Applied Sciences MDPI (Switzerland)*, *13*(6), 3438. https://doi.org/10.3390/app13063438