JNEST EISSN: 2961-8738 ● PISSN: 2961-8916 ● DOI: 10.56741/jnest.v4i02.999
Article Received 05 May 2025, Revised 03 June 2025

Optimizing Brain Tumor Classification with Freeze-5 VGG16 and Dataset Fusion

Vicky*, Ronsen Purba

***Abstract*—Magnetic resonance imaging (MRI)-based brain tumor classification is pivotal for early diagnosis and treatment planning. This study enhances the VGG16 pretrained model through freeze-5 fine-tuning (i.e., freezing the first five convolutional layers) and dataset fusion of two public repositories, yielding 5,023 training and 1,311 testing images. Preprocessing includes normalization and grayscale-to-RGB conversion, followed by moderate augmentation (rotation ≤ 15°, shift ≤ 0.1, zoom ≤ 0.1, brightness [0.9–1.1]). The base VGG16 (without top layers) is extended with GlobalAveragePooling2D, Dense (1024, ReLU), Dropout (0.5), and Dense (4, softmax) layers. The model is compiled with the Adam optimizer (lr=1e-4), EarlyStopping, and ReduceLROnPlateau callbacks. On the test set, the proposed configuration achieves peak accuracy of 99.16 % and macro-F1 of 0.99, outperforming prior hybrid approaches. An ablation study confirms that the freeze-5 strategy combined with data augmentation significantly boosts generalization without overfitting. These results underscore the critical role of optimal layer-freezing and dataset fusion in brain tumor classification. Future work will explore ensemble architecture and real-time clinical deployment.**

***Index Terms*—Brain Tumor Classification, VGG16, Dataset Fusion, Fine-Tuning, Data Augmentation**

I. INTRODUCTION

BRAIN tumors are among the most serious diseases, presenting significant challenges in the medical field due to their symptoms varying according to the tumor's type, size, and location [1]. Brain tumors include several categories, such as glioma, meningioma, and pituitary tumors, each of which possesses unique characteristics that complicate the classification process [2]. Brain diseases may impact other body parts or organs. Benign brain tumors develop slowly, whereas malignant brain tumors spread quickly and damage the surrounding tissue. The type of brain tumor and the patient's age have an impact on the survival rate [3]. Additionally, early diagnosis can be substantially enhanced by combining medical experience with computer-based picture interpretation. Using a range of methods in image processing and computer vision, supervised and unsupervised feature extractions, picture classification, and segmentation have long been studied [4].

Gliomas, which arise from the supporting (glial) tissue within the brain, excluding neurons and blood vessels, are often malignant and aggressive. Meningiomas, originating from the protective membranes (meninges) that surround the brain and spinal cord, are mostly benign but can still produce compressive symptoms when large. Pituitary tumors, which develop in the pituitary gland within the skull, although typically benign, can cause hormonal disturbances and mass effects if they impinge on adjacent neural structures [5].

Early detection and classification of brain tumors are critical for planning appropriate and effective treatment, as delayed diagnosis often leads to a poor prognosis [6]. In clinical practice, medical imaging techniques such as Magnetic Resonance Imaging (MRI) have become the primary diagnostic tool for identifying brain tumors due to their high-resolution quality [7]. MRI scans use a magnetic field to provide fine-grained pictures of human organs [3]. Nevertheless, radiologists' manual interpretation of MRI scans is often subjective, time-consuming, and prone to human error [8], [9]. Therefore, accurate automated methods are required to process and classify MRI images to support rapid and precise diagnosis. Although classifying the many kinds of brain tumors is challenging, deep learning has already been applied in numerous studies [10].

Recent advances in deep learning technologies, particularly Convolutional Neural Networks (CNNs), have demonstrated significant potential in improving the accuracy and efficiency of MRI-based brain tumor classification [11]–[13]. One of the most widely used and effective CNN architectures is VGG16, renowned for its deep feature-extraction capabilities [14]. Several previous studies have successfully employed VGG16 for brain tumor classification with promising accuracy levels, thereby underscoring its considerable potential in medical applications [12], [14].

Although VGG16 has been widely adopted for brain tumor classification with reported accuracies exceeding 90% [15],[16], systematic studies examining the impact of freezing different depths of the network on classification's performance remain scarce. This study aim to address this gap by conducting an ablation analysis on freeze-15, freeze-10, and freeze-5 configurations, thereby providing practical insights into the optimal freezing strategy for improving accuracy and generalization in MRI-based brain tumor classification.

This research focuses on enhancing brain tumor classification by implementing a freeze-5 technique on the VGG16 model, where the weights of the initial five convolutional layers are kept fixed to prevent overfitting and improve the model's ability to generalize. Furthermore, we combined two publicly available datasets to create a more extensive dataset, aiming to boost the model's adaptability to various MRI image variations [6], [7], [11]. The findings from this study are anticipated to contribute substantially to the development of automated diagnostic tools for brain tumors and facilitate their broader clinical adoption in the future [17].

Vicky and **Ronsen Purba** are with the Information Technology Program, Universitas Mikroskil, Medan 20238, Indonesia. *Correspondence: vicky@students.mikroskil.ac.id

© 2025 IISTR

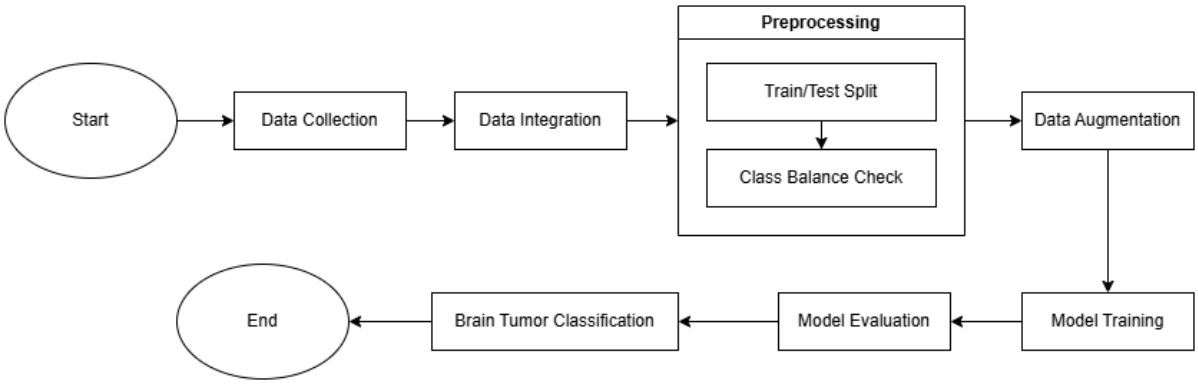


Fig. 1. Research methodology.

II. Materials and Methods

Figure 1 illustrates the research methodology which was carried out as follows. First, two publicly available MRI brain tumor datasets were collected and merged to form a unified dataset. Second, this combined dataset was split into training (80%) and testing (20%) subsets, followed by an assessment to ensure balanced class distribution. Third, during preprocessing, all images were resized to 224 × 224 pixels, normalized to a [0, 1] scale, and converted from grayscale to RGB format. Fourth, data augmentation techniques, including rotation, translation, zoom, and brightness adjustment, were applied to enhance the diversity of the training data through over-sampling and under-sampling. Fifth, model training was performed using a pretrained VGG16 network with the first five convolutional layers frozen (freeze-5) and the remaining layers fine-tuned, topped with a Dense classification layer. Sixth, the trained model was evaluated using a confusion matrix and classification report comprising precision, recall, and F1-score metrics. Finally, the model was deployed to classify new MRI images, completing the research workflow via automated tumor classification.

A. Data Collection

At this stage, brain tumor MRI images were sourced from two publicly accessible datasets: the Kaggle repository (kaggle.com/code/marianadeem755/brain-tumor-mri-image-classification-resnet50/input) [18] and the BraTS 2020 dataset (med.upenn.edu/cbica/brats2020) [19]. To maintain data quality and consistency, all images were manually reviewed to remove duplicates and validate labeling accuracy for the four tumor classes: glioma, meningioma, pituitary, and no tumor.

The two datasets were harmonized by standardizing image formats, normalizing pixel intensity distributions, and aligning labeling protocols under the supervision of expert radiologists, thereby minimizing domain shift effects caused by differences in imaging protocols and scanner types. After harmonization, the datasets were merged into a single integrated collection, which was stratified into 80% training (5,712 images) and 20% testing (1,311 images) subsets, preserving the proportional distribution of each class.

To prevent data leakage, dataset splitting was conducted at the subject (patient) level, ensuring that images from the same patient did not appear in both training and testing sets. This approach enhances the reliability and generalizability of the evaluation outcomes. The BraTS dataset was referenced according to its official publications [21]–[23].

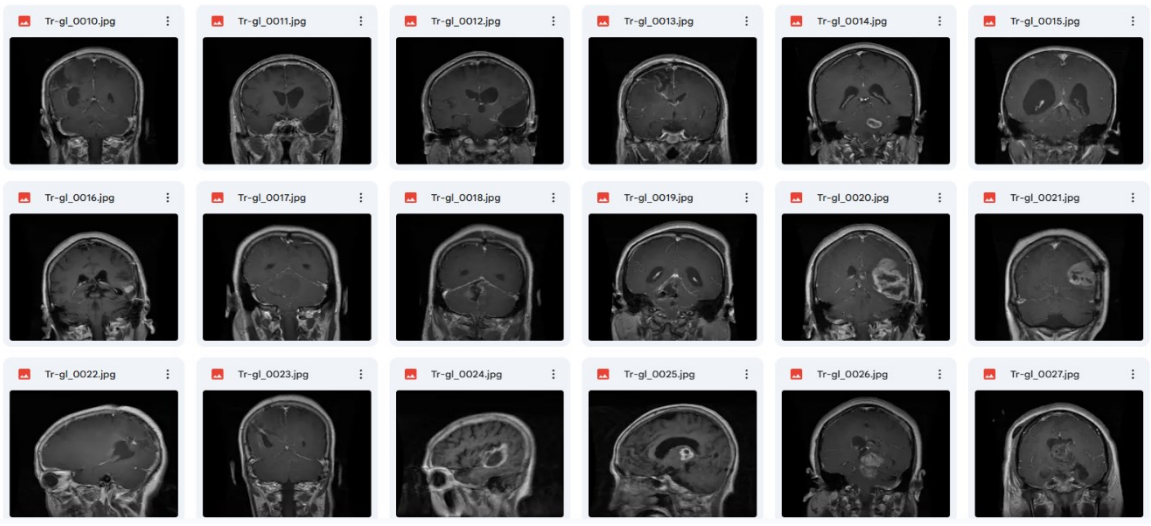


Fig. 2. MRI brain tumor glioma.

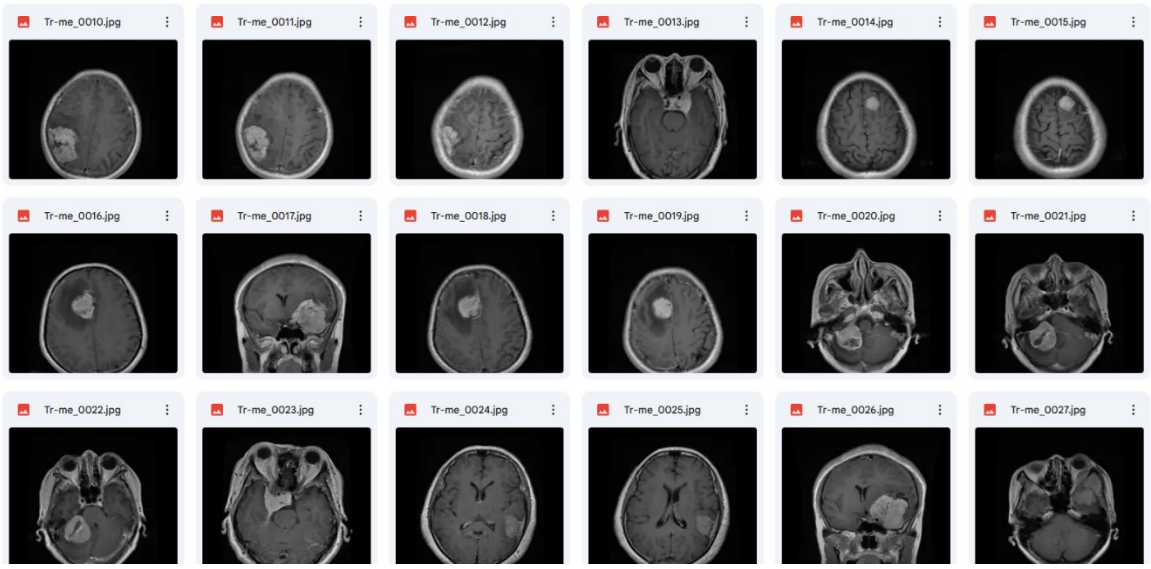


Fig. 3. MRI brain tumor meningioma.

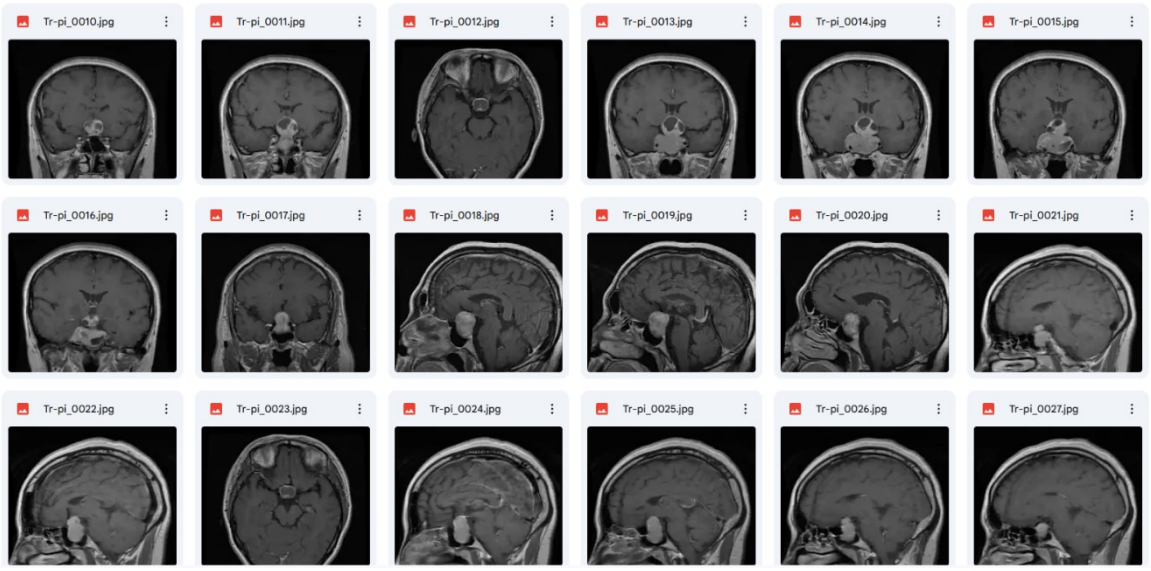


Fig. 4. MRI brain tumor pituitary.

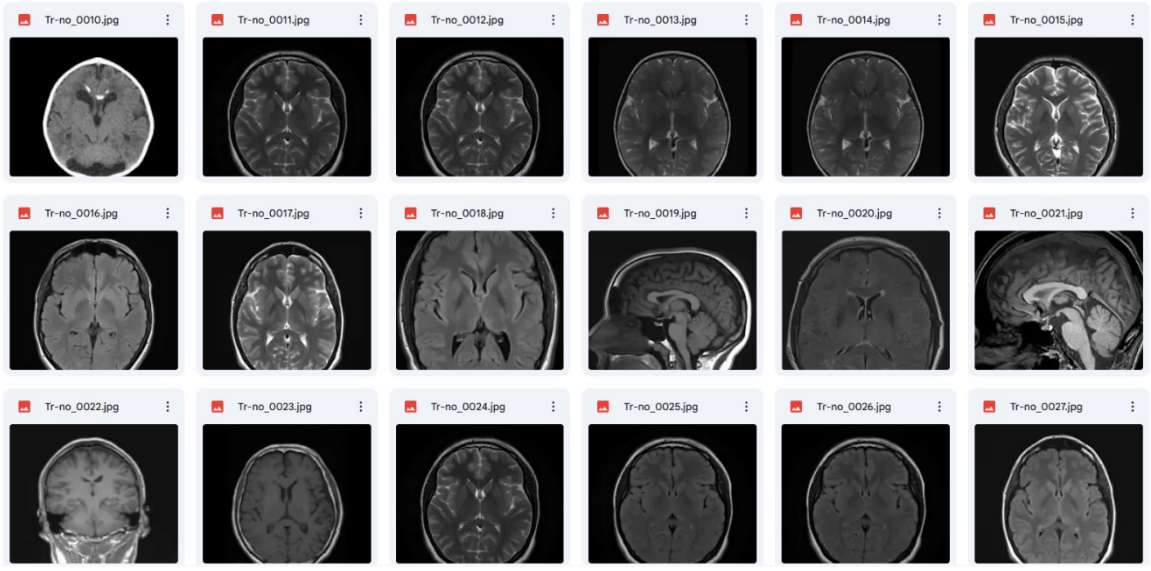


Fig. 5. MRI brain - no tumor.

TABLE I
MRI CLASS DISTRIBUTION OF BRAIN TUMORS

Class	Number of Images
Glioma	1,621
Meningioma	1,645
Pituitary	2,000
No Tumor	1,757
TOTAL	**7,023**

B. Data Preprocessing

All images in the integrated dataset underwent a preprocessing pipeline designed to ensure input uniformity and enhance model generalization. First, each image was resized to 224 × 224 pixels to match the VGG16 input specifications and minimize computational memory requirements. Second, pixel intensities were normalized to the [0, 1] range using a scaling factor of 1/255, ensuring consistent contrast values across images and accelerating convergence during training. Third, to accommodate grayscale data, a *preprocess_image* function was applied to replicate the single channel into three identical channels (i.e., RGB), guaranteeing that the model processes every input with the same channel structure. Before augmentation, we evaluated the sample distribution per class by counting label frequencies, then employed scikit-learn's *compute_class_weight* to generate class weights, thereby balancing the loss contribution among classes during optimization. This preprocessing pipeline not only standardizes image format and value range but also addresses data imbalance so that the model can learn brain-tumor patterns fairly and effectively.

C. Data Augmentation

This study employs real-time data augmentation via Keras's ImageDataGenerator class to enrich training data variability and mitigate overfitting. The augmentation scheme includes random rotations up to ±15° simulating patient head tilt horizontal and vertical shifts up to 10% to accommodate variations in scan positioning, and random zoom up to 10% to replicate differences in distance and focus. Brightness adjustments in the range [0.9–1.1] are also applied to mimic changes in contrast and MRI signal intensity. Horizontal flipping is omitted to preserve the brain's inherent asymmetry and avoid misleading artifacts. Each transformation is applied randomly before batches are fed to the network so that, in every epoch, the model "sees" a novel set of images, thereby enhancing generalization to unseen data. Visual validation of the augmented outputs is conducted via a *visualize_augmented_images* function to ensure tumor structures remain intact without excessive artifacts. Through this augmentation pipeline, the model learns not only from original images but also from representative synthetic variants, resulting in more robust classification performance against clinical MRI variability.

Six MRI brain images were randomly chosen as a sample and shown with their corresponding class labels so that the outcomes of the augmentation process could be visually examined. Each picture depicts a different tumor class, such as meningioma, glioma, pituitary tumor, or no tumor, as seen in Figure 6.

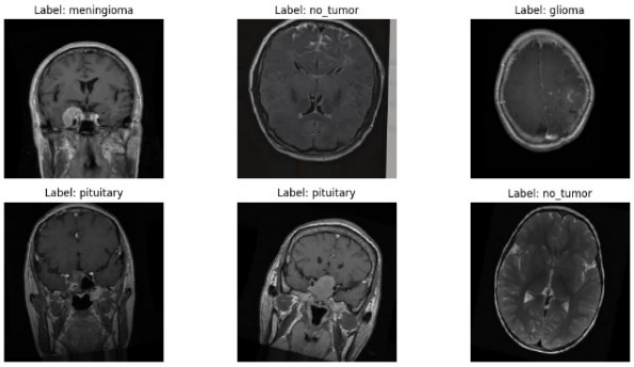


Fig. 6. Augmented MRI brain images with corresponding class labels.

Real-time augmentation modifications, such as random rotations, shifts, zooms, and brightness adjustments, have been applied to these photos. The visual examination verifies that following augmentation, anatomical features including tumor regions remain visible and undamaged. In order to make sure that the modifications being applied improve generalization without creating distortions that can deceive the classification model, this step is essential. Thus, a more robust and dependable learning process for brain tumor classification is made possible by the augmentation pipeline, which generates representative and varied training instances while maintaining clinical relevance.

D. Model Training

The base model employed is VGG16 pretrained on ImageNet, with the *include_top=False* option so that it functions solely as a feature extractor without its default classification head. In this transfer-learning implementation, the first five convolutional layers are frozen ("freeze-5") to preserve the pretrained weights that capture basic patterns such as edges, lines, and textures. From the sixth layer onward, weights are unfrozen to allow adaptation to the specific characteristics of brain-tumor MRI images. Following the feature-extractor output, a GlobalAveragePooling2D layer is added to reduce each feature map to a single scalar, which is then passed to a Dense layer with 1,024 units and ReLU activation to form nonlinear combinations of these features. To prevent overfitting, a Dropout layer with a rate of 0.5 is applied before the final layer. The output layer is a Dense layer with four units and softmax activation, producing class-probability scores for the four tumor categories. With this scheme, the early layers of VGG16 retain the generalization strength conferred by large-scale pretraining, while the fine-tuned layers learn MRI-specific tumor patterns, thereby achieving high accuracy with improved training efficiency.

E. Model Evaluation

Model evaluation in this study was conducted to assess the performance of brain-tumor classification using the confusion matrix and classification report. These two evaluation methods were chosen because they provide a

comprehensive and clear depiction of the classification model's performance.

The confusion matrix displays the model's predictions in a tabular form, where each row represents the true class and each column represents the predicted class. From this matrix, values of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) are obtained. These values serve as the basis for calculating key performance metrics, defined as follows:

Precision

Precision measures the accuracy of positive predictions made by the model. In other words, out of all instances predicted as positive (e.g., tumor), how many are actually positive. Precision helps to understand the rate of false positives (incorrect positive predictions).

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

Recall

Recall, also called sensitivity, measures the model's ability to identify all actual positive cases. That is, out of all real positive instances, how many did the model correctly detect. Recall reflects the rate of false negatives (missed positive cases).

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

F1-score

The *F1-score* is the harmonic mean of precision and recall, providing a balanced metric that considers both false positives and false negatives. It is especially useful when dealing with imbalanced class distributions.

$$F1 - score = 2 \times \frac{Precission \times Recall}{Precission + Recall} \quad (3)$$

The classification report provides a detailed summary of evaluation metrics for each brain tumor class, namely precision, recall, F1-score, and support. Precision measures the proportion of correct positive predictions (i.e., of all predictions made for a given class, how many truly belong to that class). Recall (sensitivity) evaluates the model's ability to capture all actual positive samples (i.e., of all images that belong to a certain class, how many are correctly detected). The F1-score, as the harmonic mean of precision and recall, offers a balanced measure of both metrics, which is especially valuable when class distributions are imbalanced. Finally, support indicates the number of actual observations in each class, helping to contextualize the other metrics (e.g., an F1-score for a class with low support should be interpreted with caution). By presenting the classification report alongside the confusion matrix, readers can assess not only the overall error rate but also the specific error characteristics for each class, which is crucial in clinical settings, where failing to detect a particular tumor type can have serious implications.

III. Result and Discussions

This ablation study compares three fine-tuning scenarios on the VGG16 architecture (i.e., freezing the first 15, 10, and 5 convolutional layers) using the combined dataset of 5,712 training images and 1,311 test images. The results demonstrate that the freeze-5 configuration, when paired with dataset fusion and moderate augmentation, achieves the highest performance.

Based on Table II, the model achieved a test accuracy of 99.16 % and a macro-F1 score of 0.99, outperforming the freeze-10 configuration (98.80 % / 0.99) and the freeze-15 configuration (97.80 % / 0.98). The narrow range of loss values between the training and validation sets further indicates a high degree of model generalization with no signs of overfitting.

TABLE II
Test Accuracy and Macro-F1 for Various Freeze-Layer Configurations

Freeze Layer	Test Accuracy (%)	Macro F1
15	97.80	0.98
10	98.80	0.99
5	99.16	0.99

The performance-curve analysis in Figure 7 indicates that all three models experienced rapid accuracy gains up to the fifth epoch; however, only the freeze-5 model maintained stability above 98 % through the end of training. Concurrently, validation loss decreased steadily to below 0.05 without any rebound spikes, reinforcing the conclusion that this model preserves its generalization capability despite strong fitting to the training data.

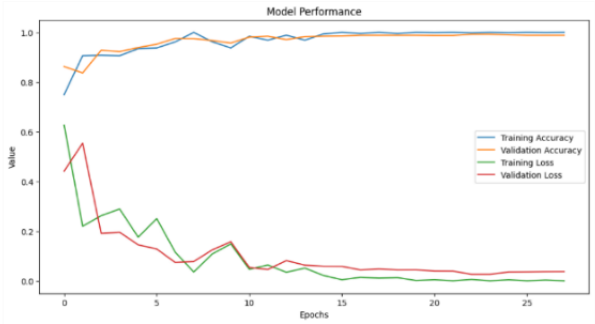


Fig. 7. Curve of training - freeze-5.

To further validate the robustness of our proposed model, we conducted a 5-fold cross-validation procedure. The results in Table III show consistently outstanding performance across all folds, with an average accuracy of 99.93 % and an extremely low standard deviation of 0.07 %, indicating stable and reliable predictions. Precision, recall, and F1-score metrics reached perfect scores of 1.00 for all tumor classes, demonstrating excellent model generalization and minimal variability.

TABLE III
FIVE-FOLD CROSS-VALIDATION RESULTS FOR THE FREEZE-5 VGG16 MODEL

Fold	Accuracy (%)	Precision	Recall	F1-score	Support
1	99.91	1.00	1.00	1.00	1,143
2	99.83	1.00	1.00	1.00	1,143
3	100.00	1.00	1.00	1.00	1,142
4	100.00	1.00	1.00	1.00	1,142
5	99.91	1.00	1.00	1.00	1,142
Avg	**99.93**	**1.00**	**1.00**	**1.00**	**5,692**
Std Dev	**0.07**	**0.00**	**0.00**	**0.00**	**-**

These statistical results reinforce confidence in the freeze-5 VGG16 fine-tuning strategy, supporting the reproducibility and robustness of our findings across multiple data splits.

Figure 8 shows the confusion matrix for the freeze-5 configuration, in which all tumor classes (glioma, meningioma, and pituitary) and the normal class are predicted with precision and recall scores above 0.97. The fewest misclassifications occur between the glioma and meningioma classes: two classes with visually similar characteristics, yet the error rate remains below 3 %.

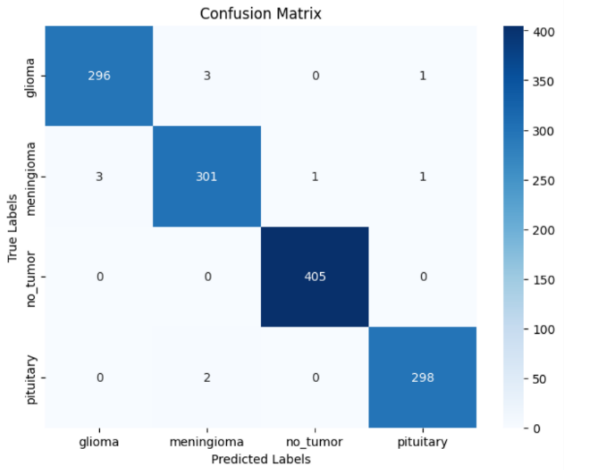


Fig. 8. Confusion matrix - freeze-5.

Figure 9. presents the full classification report, with precision, recall, and F1-score for each class approaching perfect values. For example, the F1-score for the glioma class is 0.98, indicating a strong balance between precision and recall.

```
Classification Report:
              precision    recall  f1-score   support

      glioma       0.99      0.99      0.99       300
  meningioma       0.98      0.98      0.98       306
    no_tumor       1.00      1.00      1.00       405
   pituitary       0.99      0.99      0.99       300

    accuracy                           0.99      1311
   macro avg       0.99      0.99      0.99      1311
weighted avg       0.99      0.99      0.99      1311
```

Fig. 9. Classification report - freeze-5.

TABLE IV
PERFORMANCE COMPARISON WITH STATE-OF-THE ART CNN MODELS

Authors	Dataset	Model	Classification Accuracy (%)
Çinar and Yıldırım [24]	Figshare	CNN	97.20
Chetana Srinivas et al. [26]	Kaggle	VGG16	98.14
Francisco Javier Díaz-Pernas [25]	Nanfang Hospital, Guangzhou, China, and General Hospital	Deep CNN	97.30
Alok Sarkar [27]	Kaggle	AlexNet CNN	98.15
Proposed Model	**Kaggle + BraTS 2020**	**VGG16 + Freeze-5**	**99.16**

Table IV demonstrates that our proposed method (VGG16 architecture with a freeze-5 strategy trained on the combined Kaggle and BraTS 2020 datasets) achieves superior accuracy compared to other CNN-based approaches reported on publicly available brain-tumor MRI datasets. These results empirically reinforce our justification for employing the freeze-5 configuration, underscoring its practical benefit and accuracy improvements over other state-of-the-art methods.

This finding is consistent with previous literature, which states that a fine-tuning strategy that freezes a portion of the early layers can combine the power of general pretrained features with the need for adaptation to domain-specific data, thus maximizing classification performance on limited medical datasets [20]. This method is also more computationally efficient than complex hybrid approaches while maintaining near-perfect accuracy.

The limitations of this study lie in its evaluation scope, which remains confined to public datasets; therefore, extensive validation using multi-center clinical datasets is highly recommended to ensure the model's robustness in real-world environments. Domain shifts arising from differences in data acquisition protocols, equipment variability, and patient demographics could significantly

impact model generalizability. Thus, future studies must incorporate external clinical validation from diverse medical centers to rigorously evaluate the robustness and clinical applicability of our proposed model.

Additionally, ethical considerations are paramount when applying AI models in clinical settings. Potential biases in training data could lead to disparities in diagnostic accuracy across different patient groups. To address this, transparency and interpretability of the model are crucial. Techniques such as Gradient-weighted Class Activation Mapping (Grad-CAM) should be explored to provide visual explanations of model decisions, thereby enhancing clinician trust and facilitating integration into routine workflows. Data privacy and security must also be prioritized to comply with healthcare regulations and protect patient information. Incorporating these considerations will be essential for responsible and safe deployment of AI-based diagnostic tools in clinical practice.

Our findings reinforce and extend previous literature that supports freezing a subset of early convolutional layers to leverage pretrained generic features while allowing adaptation to domain-specific MRI data [20]. Unlike prior studies that apply freeze-5 heuristically, our systematic ablation across multiple freezing depths clarifies the specific advantage of the freeze-5 configuration for brain tumor classification. This empirical evidence contributes novel practical guidance for transfer learning configurations in medical imaging, balancing model complexity, training efficiency, and accuracy gains. Consequently, this work provides meaningful advancement over existing approaches and offers actionable insights for future research and clinical applications.

IV. Conclusion

In conclusion, the systematic series of experiments conducted in this study demonstrate that a VGG16 fine-tuning strategy (i.e., freezing the first five convolutional layers (freeze-5), combined with moderate augmentation and fusion of two public datasets) yields a substantial improvement in MRI-based brain-tumor classification performance, achieving 99.16% accuracy and a macro-F1 score of 0.99. These findings confirm that freezing through the fifth layer effectively merges the power of pretrained low-level features (such as edge and texture detection) with the capacity of the middle and higher layers to learn highly specific tumor patterns, resulting in strong generalization without traces of overfitting. Moreover, the anatomically preserving augmentation design, including rotation, translation, zoom, and brightness variation, proved capable of generating representative training-data diversity, enhancing the model's robustness against clinical MRI heterogeneity. The depth of the ablation study further contributes to a technical recommendation regarding the optimal number of frozen layers for medical-classification tasks with limited data sources. Although the results are highly promising, further validation on multi-center clinical datasets is recommended to ensure the model's robustness across different MRI equipment and scanning protocols. Additionally, practical considerations for deployment, such as computational efficiency, inference time, and compatibility with various clinical hardware including mobile and edge devices, should be explicitly addressed in future implementations. Exploring explainability methods (e.g., Grad-CAM or LIME) is also recommended to improve transparency, allowing clinicians to better interpret and trust AI-driven decisions. Furthermore, exploring ensemble architecture and integrating the model into real-time diagnostic systems represent strategic next steps toward realizing a clinically reliable automated diagnostic solution. To support reproducibility and encourage further research, the source code and processed datasets used in this study are publicly available through a GitHub repository [28]. Consequently, this research not only enriches the understanding of fine-tuning in transfer learning but also paves a practical pathway for broader AI adoption in supporting clinical decision making.

Acknowledgment

The authors thank Erwin Panjaitan for his insightful feedback throughout the *S2 program*. The author is also grateful to Hardy, Rector of Universitas Mikroskil, for his continuous support and encouragement. Appreciation is extended to the faculty and staff of the Information Technology Program at Universitas Mikroskil for providing a stimulating research environment. Finally, the author acknowledges the wider academic community for access to literature and tools that have greatly facilitated this work.

References

[1] R. Rudiansyah and A. Husein, "Klasifikasi Tumor Otak pada gambar Magnetic Resonance Images (MRI) dengan Pendekatan Pembelajaran Mendalam," *Data Sci. Indones.*, vol. 4, no. 1, pp. 62–68, Jul. 2024, doi: 10.47709/dsi.v4i1.4265. [*Eng Trans: Classification of Brain Tumors on Magnetic Resonance Images (MRI) with Deep Learning Approach*].

[2] Y. Chen, "Comparative Analysis of Brain Tumor Classification and Models Based on VGG16," in *Proceedings of the 1st International Conference on Engineering Management, Information Technology and Intelligence*, *SCITEPRESS - Science and Technology Publications*, pp. 652–658, 2024. doi: 10.5220/0012961300004508.

[3] A. Muis, S. Sunardi, and A. Yudhana, "Medical Image Classification of Brain Tumor using Convolutional Neural Network Algorithm," *J. Infotel*, vol. 15, no. 3, pp. 227–232, 2023, doi: 10.20895/infotel.v15i3.964.

[4] C. B. Muddaraju, Shrinivasa, S. Narasimhamurthy, and V. Sontakke, "Automated Diagnosis of Brain Tumor Classification and Segmentation of Magnetic Resonance Imaging Images," *IAES Int. J. Artif. Intell.*, vol. 13, no. 4, pp. 4833–4842, 2024, doi: 10.11591/ijai.v13.i4.pp4833-4842.

[5] M. N. Musa, "MRI-Based Brain Tumor Classification using ResNet-50 and Optimized Softmax Regression," *J. Infotel*, vol. 16, no. 3, Sep. 2024, doi: 10.20895/infotel.v16i3.1175.

[6] S. Y. Prasetyo and G. Z. Nabiilah, "Perbandingan Model Machine Learning pada Klasifikasi Tumor Otak Menggunakan Fitur Discrete Cosine Transform," *J. Teknol. Terpadu*, vol. 9, no. 1, pp. 29–34, 2023. [*Eng Trans: Comparison of Machine Learning Models for Brain Tumor Classification Using Discrete Cosine Transform Features*].

[7] F. Citra R, F. Indriyani, and I. R. Rahadjeng, "Klasifikasi Tumor Otak Berbasis Magnetic Resonance Imaging menggunakan Algoritma Convolutional Neural Network," *Digit. Transform. Technol.*, vol. 3, no. 2, pp. 918–924, Jan. 2024, doi: 10.47709/digitech.v3i2.3469. [*Eng Trans: Magnetic Resonance Imag-ing Based Brain Tumor Classification Using Convolutional Neural Network Algorithm*].

[8] A. Nurhidayat, W. A. Arrosyid, and R. Samsinar, "Prediksi Tumor Otak Menggunakan Metode Convolutional Neural Network (CNN) dan Algoritma Decision Tree," *Seminar Nasional Teknologi & Sains*. Vol. 4. No. 1. 2025. 2025. [*Eng Trans: Brain Tumor*

Prediction Using Convolutional Neural Network (CNN) Method and Decision Tree Algorithm].

[9] B. Ulu, "Comparison of Deep Learning Methods in Brain Tumor Diagnosis: High-Performance Classification with MRI Data," *Commun. Fac. Sci. Univ. Ankara Ser. A2-A3 Phys. Sci. Eng.*, vol. 67, no. 1, pp. 59–73, Feb. 2025, doi: 10.33769/aupse.1619837.

[10] A. Biswas and M. S. Islam, "A Hybrid Deep CNN-SVM Approach for Brain Tumor Classification," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 9, no. 1, pp. 1–15, 2023, doi: 10.20473/jisebi.9.1.1-15.

[11] E. U. Haq, H. Jianjun, X. Huarong, K. Li, and L. Weng, "A Hybrid Approach Based on Deep CNN and Machine Learning Classifiers for the Tumor Segmentation and Classification in Brain MRI," *Comput. Math. Methods Med.*, vol. 2022, 2022, doi: 10.1155/2022/6446680.

[12] M. A. Purnama Wibowo, Muhammad Bima Al Fayyadl, Yufis Azhar, and Zamah Sari, "Classification of Brain Tumors on MRI Images Using Convolutional Neural Network Model EfficientNet," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 4, pp. 538–547, Aug. 2022, doi: 10.29207/resti.v6i4.4119.

[13] S. Aksoy, "Brain Tumor Detection and Classification via VGG16-Based Deep Learning on MRI Imaging," *Int. J. Multidiscip. Res.*, vol. 7, no. 2, Mar. 2025, doi: 10.36948/ijfmr.2025.v07i02.39435.

[14] O. A. Supriadi, E. Utami, and D. Ariatmanto, "Deteksi Tumor Otak Melalui Gambar MRI Berdasarkan Vision Transformers dengan Tensorflow dan Keras," *J. Inform. Univ. Pamulang*, vol. 8, no. 3, pp. 385–392, Sep. 2023, doi: 10.32493/informatika.v8i3.32707. [*Eng Trans: Brain Tumor Detection Through MRI Images Based on Vi-sion Transformers with Tensorflow and Keras*]

[15] B. Babu Vimala, S. Srinivasan, S. K. Mathivanan, Mahalakshmi, P. Jayagopal, and G. T. Dalu, "Detection and Classification of Brain Tumor using Hybrid Deep Learning Models," *Sci. Rep.*, vol. 13, no. 1, pp. 1–18, 2023, doi: 10.1038/s41598-023-50505-6.

[16] D. S. Wankhede, C. J.shelke, V. K. Shrivastava, R. Achary, and S. N. Mohanty, "Brain Tumor Detection and Classification Using Adjusted InceptionV3, AlexNet, VGG16, VGG19 with ResNet50-152 CNN Model," *EAI Endorsed Trans. Pervasive Heal. Technol.*, vol. 10, 2024, doi: 10.4108/eetpht.10.6377.

[17] S. Li, "Classification of Brain Tumors Based on Magnetic Resonance Imaging (MRI) using Deep Learning Models," *Appl. Comput. Eng.*, vol. 40, no. 1, pp. 247–254, Feb. 2024, doi: 10.54254/2755-2721/40/20230660.

[18] Maria Nadeem, "Brain Tumor MRI Image Classification:ResNet50," Kaggle. [Online]. Available: https://www.kaggle.com/code/marianadeem755/brain-tumor-mri-image-classification-resnet50/input

[19] Awsaf, "Brain Tumor Segmentation (BraTS2020)," Kaggle. [Online]. Available: https://www.kaggle.com/datasets/awsaf49/brats2020-training-data

[20] Z. Li and X. Zhou, "A Global-Local Parallel Dual-Branch Deep Learning Model with Attention-Enhanced Feature Fusion for Brain Tumor MRI Classification," *Comput. Mater. Contin.*, vol. 83, no. 1, pp. 739–760, 2025, doi: 10.32604/cmc.2025.059807.

[21] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al. "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)", IEEE Transactions on Medical Imaging 34(10), 1993-2024 (2015) DOI: 10.1109/TMI.2014.2377694

[22] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J.S. Kirby, et al., "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features", *Nature Scientific Data*, 4:170117, 2017, DOI: 10.1038/sdata.2017.117

[23] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, et al., "Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge", arXiv preprint arXiv:1811.02629, 2018.

[24] A. Çinar and M. Yıldırım, "Detection of Tumors on Brain MRI Images Using the Hybrid Convolutional Neural Network Architecture," *Medical Hypotheses*, vol. 139, article 109684, 2020.

[25] Díaz-Pernas, F. J., Martínez-Zarzuela, M., Antón-Rodríguez, M. & González-Ortega, D. A Deep Learning Approach for Brain Tumor Classifcation and Segmentation Using a Multiscale Convolutional Neural Network. *Feature Papers Artif. Intell*. Med. Vol 9, No.2, pp. 1–14, 2021. doi: 10.3390/healthcare9020153.

[26] M. S. Majib, M. D. M. Rahman, T. M. S. Sazzad, N. I. Khan, and S. K. Dey, "VGG-SCNet: A VGG Net-Based Deep Learning Framework for Brain Tumor Detection on MRI Images," IEEE Access, vol. 9, pp. 116942–116952, 2021, doi: 10.1109/ACCESS.2021.3105874.

[27] A. Sarkar, M. A. Maniruzzaman, and M. A. Alahe, "An Efective and Novel Approach for Brain Tumor Classifcation using Alexnet CNN Feature Extractor and Multiple Eminent Machine Learning Classifers in MRIs," *J. Sensors Hindawi,* 1224619, pp. 1–19, 2023.

[28] Vicky and R. Purba, "Brain Tumor Detection using Freeze-5 VGG16 and Dataset Fusion," Github. [Online]. Available: github.com/LuckyDevelop/Brain-Tumor-Detection-Using-Freeze-5-VGG16-and-Dataset-Fusion.git.



Vicky is a Master's candidate in the Information Technology program at Universitas Mikroskil, Medan, Indonesia, where he is researching deep-learning approaches for brain tumor classification (email: vicky@students.mikroskil.ac.id).



Ronsen Purba is a Lecturer in Master of Information Technology (MTI) Program at Universitas Mikroskil, Medan, Indonesia (email: ronsen@mikroskil.ac.id).