# Ensemble Combination of CNN for MRI-Based Brain Tumor Classification

Akbar Sidqi
Department of Informatics Engineering, Universitas Islam Negeri Maulana Malik Ibrahim Malang, Malang 65144, Indonesia

Irwan Budi Santoso
Department of Informatics Engineering, Universitas Islam Negeri Maulana Malik Ibrahim Malang, Malang 65144, Indonesia

Sri Harini
Department of Informatics Engineering, Universitas Islam Negeri Maulana Malik Ibrahim Malang, Malang 65144, Indonesia

**Abstract:** Classifying 17 types of brain tumors remains a major challenge in the medical field, especially in improving diagnostic accuracy and accelerating patient care. This study proposes a CNN-based model with an ensemble combination approach to improve accuracy by integrating multiple architectures through Majority Voting and Weighted Average for more reliable predictions. The models are evaluated using accuracy, precision, recall, and F1-score metrics. The results show that CNN3 with Nadam achieves the best performance (accuracy: 0.90–0.91), outperforming CNN1 (0.87–0.89) and CNN2 (0.82–0.87). The ensemble combination improves accuracy across all models, with CNN3 achieving the highest accuracy (0.96), followed by CNN1 (0.94–0.95) and CNN2 (0.91–0.92). This study demonstrates that the ensemble combination approach can improve the performance of brain tumor classification using deep learning, contributing to faster and more accurate medical diagnosis. Furthermore, these findings open up opportunities for further research in advancing brain tumor detection systems.

**Keywords**: Ensemble Combination, Convolutional Neural Network, Brain Tumor Classification, MRI Image.

Correspondents Author:
Akbar Sidqi, Department of Informatics Engineering, Universitas Islam Negeri Maulana Malik Ibrahim Malang, Malang 65144, Indonesia.
Email: akbarsidqi@gmail.com

Received January 20, 2025; Revised February 23, 2025; Accepted March 25, 2025; Published March 26, 2025

## Introduction

Brain as the primary regulator of body functions, it is a very complex organ. Changes in the brain, such as uncontrolled cell growth, can lead to brain tumors, which can be either benign or malignant (I. B. Santoso et al., 2024). Delayed diagnosis of brain tumors often leads to patient death, making it necessary to have effective methods for accurately classifying brain tumors from the initial diagnosis. Magnetic Resonance Imaging (MRI) technology produces detailed 3D images without invasive procedures, making it an essential tool for radiologists in accurately diagnosing and classifying brain tumors (R. Zahid et al., 2024).

Convolutional Neural Networks (CNN) are widely used in MRI image analysis to identify and classify brain tumors by capturing complex patterns. Previous studies have shown that CNN can enhance classification performance, even with just feature extraction. (I. B. Santoso et al., 2024). For example, (Ruqsar Zaitoon et al., 2023) used two CNN architectures with 5 and 6 layers, achieving an accuracy of 99.87%. However, this method has limitations when applied to large datasets due to limited computing resources. Other studies, such as by (Hassan Ali Khan et al., 2020); (Saeed Mohsen et al. 2023), also shows the limitations of CNN in maintaining consistent accuracy with limited datasets. (Alwas Muis et al., 2023). Showed an accuracy of 84% in brain tumor image classification, with the main challenge being the variation in image complexity and feature overlap between tumor types ( I. B. Santoso et al., 2024; Eko Hari Rachmawanto *et al*., 2024). The English translation is: Therefore, to improve the classification performance, it is necessary to apply multiple CNN models.

Several studies have developed advanced approaches. For instance, Naveen Mukkapati et al. (2022) combined CNN with U-Net, RefineNet, and SegNet, achieving an accuracy of 96.85%. In segmentation and classification, Eko Hari Rachmawanto et al. (2024) employed a simple CNN on the BRATS 2018-2019 dataset, reaching an accuracy of 94.14%. Mohamed Amine Mahjoubi et al. (2023) applied grid search on CNN and compared it with AlexNet and ResNet-50, obtaining an accuracy of 95.44%. Nassar, S.E. utilized five pre-trained CNN models with the majority voting technique, enhancing accuracy up to 99.31%.

Hybrid models have also proven effective. Nazik Alturki et al. (2023) combined CNN features with a voting classifier, achieving an accuracy of 99.9%, outperforming NGBoost (98.5%) and EfficientNet-B0 (98.8%). Gergo Bogacsovics et al. (2024) implemented a CNN-based ensemble using AlexNet, MobileNetv2, and EfficientNet, attaining an accuracy above 92%. Anand Vatsala et al. (2023) applied transfer learning and weighted average ensemble, achieving a sensitivity of 96%, precision of 99%, and an F1-score of 97%.

This study proposes brain tumor classification by implementing multiple CNN models with different convolutional filter sizes. Each filter is designed to extract features from MRI images, while the classification results are combined using an ensemble combination approach to enhance model performance. This approach aims to maximize the model's ability to distinguish between 17 types of brain tumors, including glioma and meningioma, with higher accuracy.

The ensemble combination is evaluated using accuracy, precision, recall, specificity, and confusion matrix metrics. Optimizing the combination of CNN models in this study not only improves the accuracy of brain tumor classification but also addresses dataset limitations and tumor feature complexity. Furthermore, this approach has the potential to serve as a foundation for developing more reliable medical diagnostics and more effective treatment planning.

# Research Method

This section describes the method used to classify brain tumor types with deep learning, namely Convolutional Neural Network (CNN), which is combined with Ensemble Combination to achieve the best performance in the proposed model.

## Dataset of Experiment

The MRI image dataset uses axial perspective on T1, T1C+, and T2 sensitivity, which is sourced from (Fernando Feltrin) Downloaded from Kaggle. This dataset consists of 4415 images divided into 17 classes, namely: glioma (T1), glioma (T1C+), glioma (T2), meningioma (T1), meningioma (T1C+), meningioma (T2), neurocitoma (T1), neurocitoma (T1C+), neurocitoma (T2), normal (T1), normal (T2), outros (T1), outros (T1C+), outros (T2), schwannoma (T1), schwannoma (T1C+), and schwannoma (T2). The differences in brain tumor types in this dataset can be seen in Figure 1.

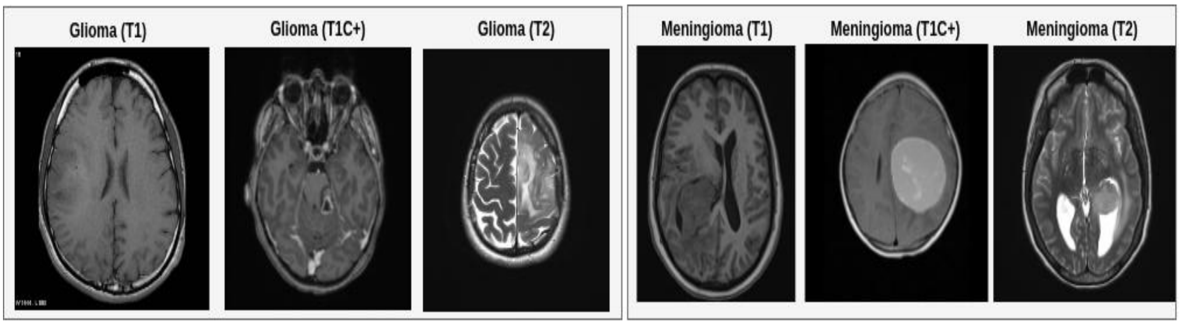



**Figure 1 Normal brain and brain tumor T1, T1C+, and T2**

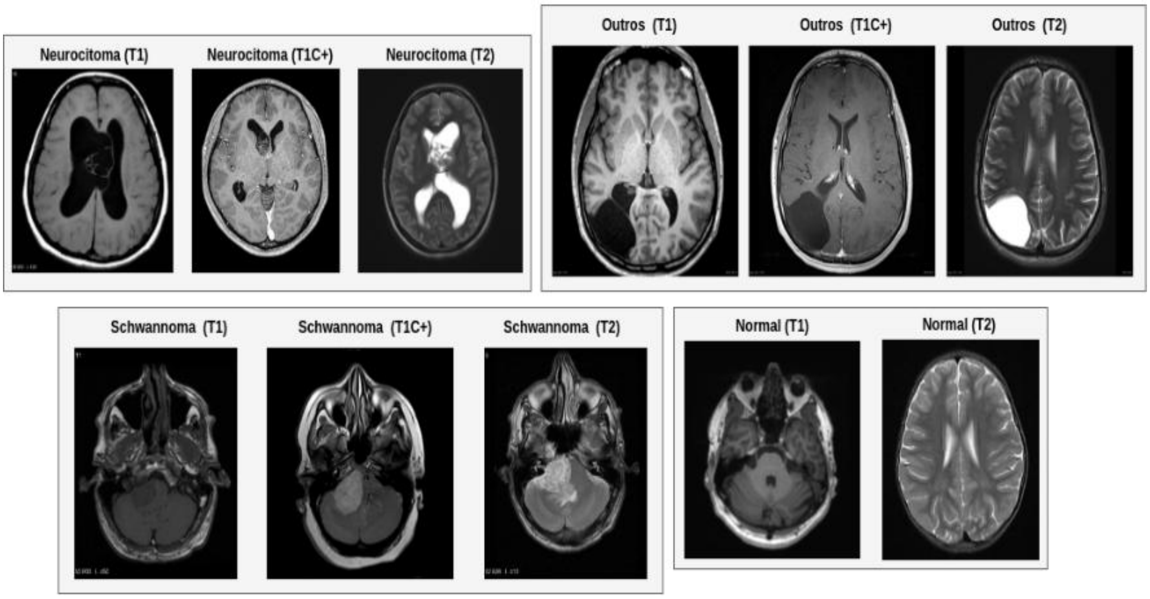



**Figure 1 Normal brain and brain tumor T1, T1C+, and T2 (*continue*)**

## Data pre-processing

Pre-processing is performed to prepare the MRI images as input for the model, including resizing the images to 224 × 224 pixels (Mohamed Amine Mahjoubi et al., 2023). This size is chosen to adjust the image dimensions without altering the content or important proportions of the MRI images, ensuring that the images have a uniform size before being input into the CNN model.

### Augmentation

Image data augmentation is used to enhance the quality of the dataset, allowing the model to learn more diverse patterns (X. Xiao). The augmentation techniques applied to the CNN model. (YD Zhang et al., 2018). Include rotation_range, zoom_range, horizontal_flip, and fill_mode.

### Design System

The following diagram shows the stages of applying the CNN model to the brain tumor dataset, starting from preprocessing, model design, Ensemble Combination, to accuracy evaluation, as shown in Figure 2.

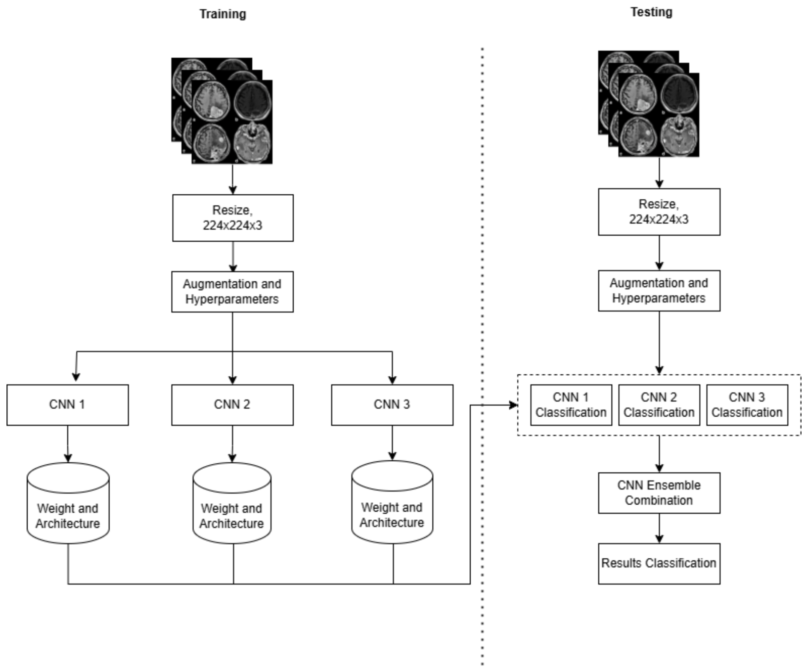



**Figure 2 Design System**

## Convolutional Neural Network

By applying the CNN model architecture (Naveen Mukkapati et al., 2022; A. Muis et. Al., 2024. Effectively to classify complex images and process grid-structured data to analyze pixel values and reduce noise (Abdullah A et al., 2024). This CNN model is used on the brain tumor dataset to classify 17 types of tumors, involving convolutional layers, activation layers, pooling layers, flattening, fully connected layers, and output layers. ( I. B. Santoso et al., 2024; Nihal Remzan et al., 2022).

### Convolutional Layer

Each brain MRI image input into neurons is processed through convolution with various filters, producing a feature map to identify the characteristics of the MRI image. (I. B. Santoso et al., 2024). This process can be mathematically expressed as follows.

$$Z_i = f(W_i X + b_i), i = 1,2,3 \qquad (1)$$

### Activation Layer

The activation layer is an essential component for enhancing the non-linear properties of the decision function, which is achieved by applying a stable and undistorted activation function. The Rectified Linear Unit (ReLU) activation function is applied to each convolution process. (I. B. Santoso et al., 2024 ). Mathematicall, it can be expressed as follows.

$$\breve{Z}(Z_i) = \begin{cases} Z_i Jika Z_i \geq 0 \\ 0 Jika Z_i < 0 \end{cases}, i = 1,2,3 \tag{2}$$

## Pooling Layer

The pooling layer is used to reduce the spatial representation size of the convolution results, which helps reduce computational load and prevent overfitting. Max-pooling with a 2x2 patch size is applied. (Mohamed Amine Mahjoubi et al., 2023; I. B. Santoso et al., 2024), Which can be explained mathematically as follows:

$$g_i(\tilde{Z}_i) = Max\{\tilde{Z}_{ij}\}j, i = 1,2,3 \tag{3}$$

## Flattening

The flattening process is performed after max-pooling on each CNN network path to convert the obtained features into a vector. Each path undergoes this process before the features are combined and fed into the fully connected (FC) layer. Flattening aims to convert multidimensional data, such as images, into a vector format so that the model can perform classification based on the extracted features (I. B. Santoso et al., 2024). With the mathematical equation as follows :

$$h_i = flattening(g_i), i = 1,2,3 \tag{4}$$

## Output (Classification) Layer

Before entering the classification layer, the fully connected feed-forward stage connects each neuron with neurons in the previous layer to combine features and classify the tumor type in the MRI image. (Mohamed Amine Mahjoubi et al., 2023). The number of output classes in this layer is adjusted according to the number of classes in the training dataset. The output from the fully connected layer is passed to the output layer, where the softmax activation function is used to determine the probability of the presence of 17 types of tumors based on the output from the previous layer. (I. B. Santoso et al., 2024). With the mathematical equation as follows:

$$y_k(\hat{h}) = \frac{exp(\hat{h}_k)}{\sum_{j=1}^{4} exp(\hat{h}_j)}, k = 17 \tag{5}$$

## Ensemble Combination

### Majorty Voting

Majority voting is used to select the tumor class based on three CNN models. Each model provides one prediction for one of the 17 tumor classes. A vote is cast where each model gives a prediction $T_k$ with $i \in \{1,2,3\}$ for one of these classes $k \in \{1,...,17\}$. Each model contributes one vote for its predicted class. The class receiving the highest number of votes is selected as the final classification outcome (I. B. Santoso et al., 2024; Nazik Alturki et al., 2023). With the mathematical equation as follows:

$$V_k = \sum_{i=1}^{3} V_{ik} \tag{6}$$

The selected class as the final classification result.

$$h = argmax_k(V_k) \;\; h \in \{1,2,3\} \;\; k \in \{17\} \tag{7}$$

### Weighted Average

Prediction combination with weighted average. (Anand Vatsala et al., 2023). Calculated by taking the average of the softmax values from each model. (I. B. Santoso et al., 2024; Gergo Bogacsovics et al., 2024). The class with the highest average softmax value is then selected as the classification result. Mathematically, this can be expressed as follows :

$$h = argmax_k \sum_{i=1}^{3} V_{ik}/3 \; i \in \{1,2,3\} \; k \in \{17\} \tag{8}$$

### Perfomance Evaluation

The evaluation of the proposed method uses metrics such as accuracy, precision, sensitivity (recall), specificity, and F1-score. (I. B. Santoso et al., 2024). Used to assess the model's performance in classifying brain tumors from MRI images. Mathematically, this can be expressed as follows (9-13).

$$Akurasi = \frac{(tp+tn)}{(tp+fp+tn+fn)} \tag{9}$$

$$Presisi = \frac{tp}{(tp+fp)} \tag{10}$$

$$Recall\ (Sensitivitas) = \frac{tp}{(tp+fn)} \tag{11}$$

$$Spesifitas = \frac{tn}{(tn+fp)} \tag{12}$$

$$F - Score = \frac{2x\ svt\ x\ prs}{(svt+prs)} \tag{13}$$

## Experiments

Three base CNN models are used to classify 17 types of brain tumors using Python libraries. The MRI dataset is divided with a 60:20:20 ratio for training, validation, and testing. Model evaluation is performed using the Adam, AdamW, and Nadam optimizers. (Anindya Nag et al., 2024). Below is the architecture of the Base CNN model shown in Figure 3.

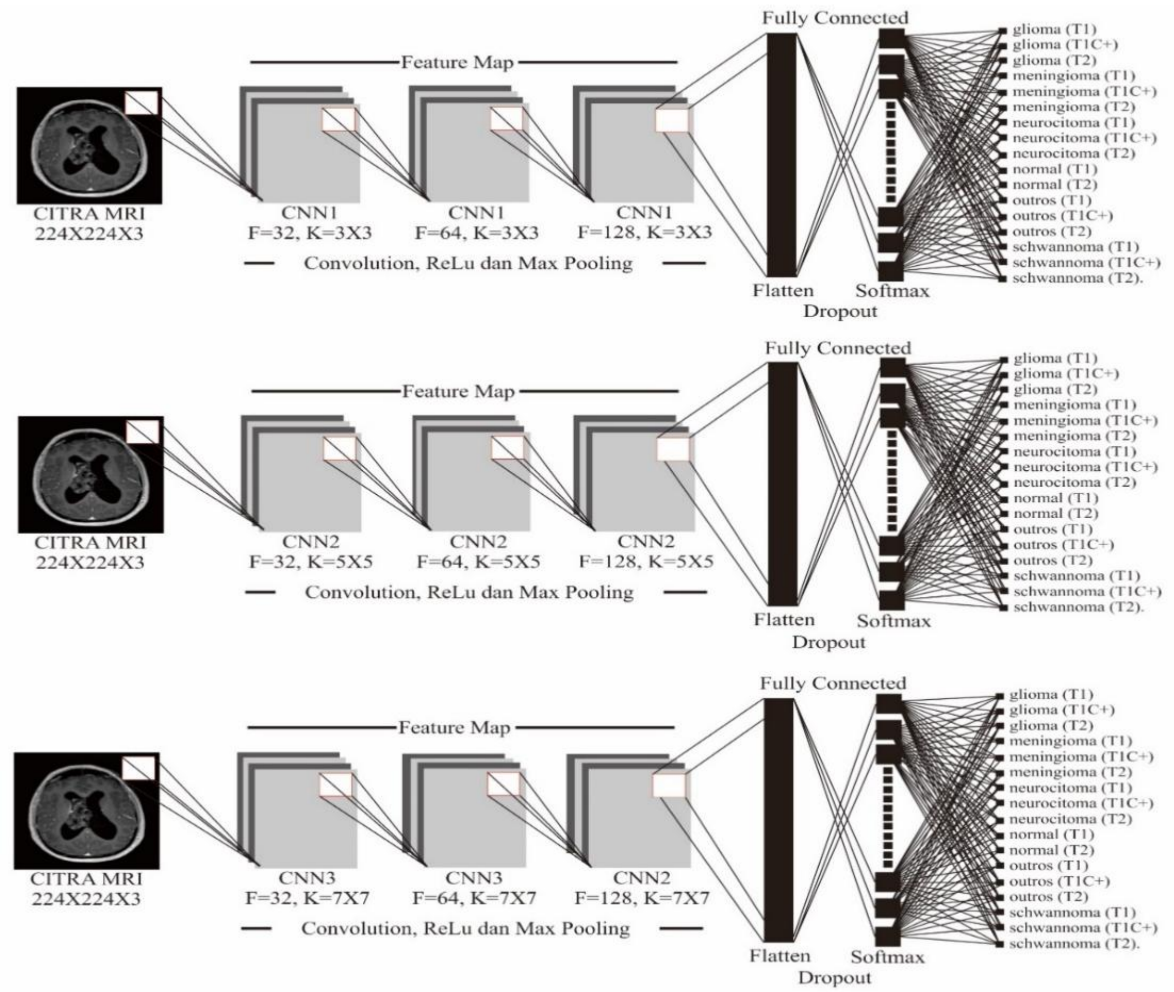



**Figure 3 Base Model CNN Architecture**

The proposed Ensemble Combination model uses three base CNN models, namely CNN1, CNN2, and CNN3. The prediction results from the three models are combined using two ensemble methods, namely Majority Voting and Weighted Average, to improve the brain tumor classification accuracy.

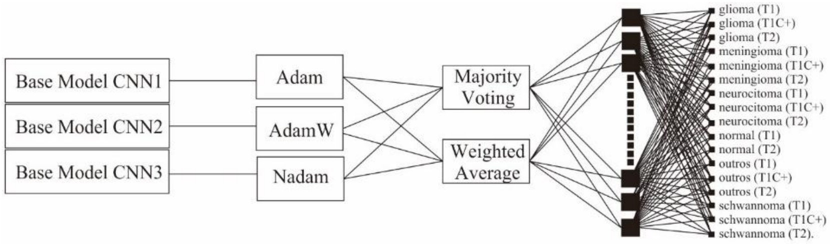



**Figure 4 Majority Voting and Weighted Average Scheme (Proposed)**

# Result and Discussion

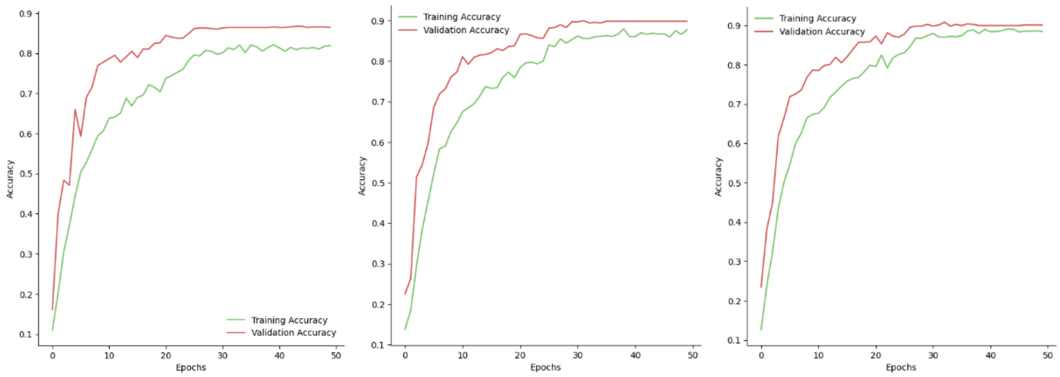
The testing of the base CNN model is conducted with different filter sizes to assess the impact of the optimizer on the developed CNN model. Table 1 shows the testing of the base CNN model along with the description of each scenario.

**Table 1 Description of Base CNN Model Scenario**

| Model | Batch Size | Learning Rate | Epoch | Kernel Size | Optimizer | Filter Size |
|---|---|---|---|---|---|---|
| CNN1 | 32 | 0.001 | 50 | 3x3 | Adam | 32, 64, 128 |
| | | | | 5x5 | Adam | |
| | | | | 7x7 | Adam | |
| CNN2 | 32 | 0.001 | 50 | 3x3 | AdamW | 32, 64, 128 |
| | | | | 5x5 | AdamW | |
| | | | | 7x7 | AdamW | |
| CNN3 | 32 | 0.001 | 50 | 3x3 | Nadam | 32, 64, 128 |
| | | | | 5x5 | Nadam | |
| | | | | 7x7 | Nadam | |

## Base Model CNN1

The testing of base CNN model CNN1 is carried out as listed in Table 1. In this process, training and validation are performed using the Adam optimizer with a learning rate of 0.001, for 50 epochs, using categorical classification. Figure 5 shows the accuracy, validation, and loss plot generated for the base CNN model CNN1.

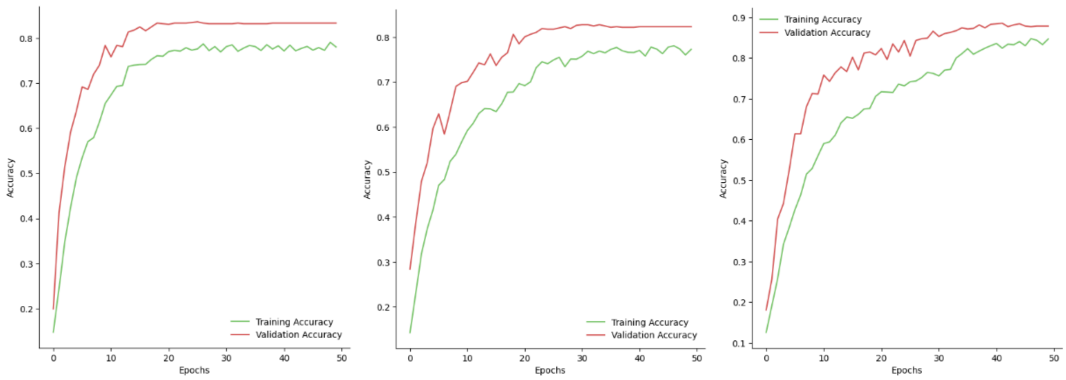



Kernel Size 3x3 Kernel Size 5x5 Kernel Size 7x7

**Figure 5 Results of Base CNN Model CNN1**

## Base Model CNN2

The testing of base CNN model CNN2 is carried out as listed in Table 1. In this process, training and validation are performed using the AdamW optimizer with a learning rate of 0.001, for 50 epochs, using categorical classification. Figure 6 shows the accuracy, validation, and loss plot generated for the base CNN model CNN2.

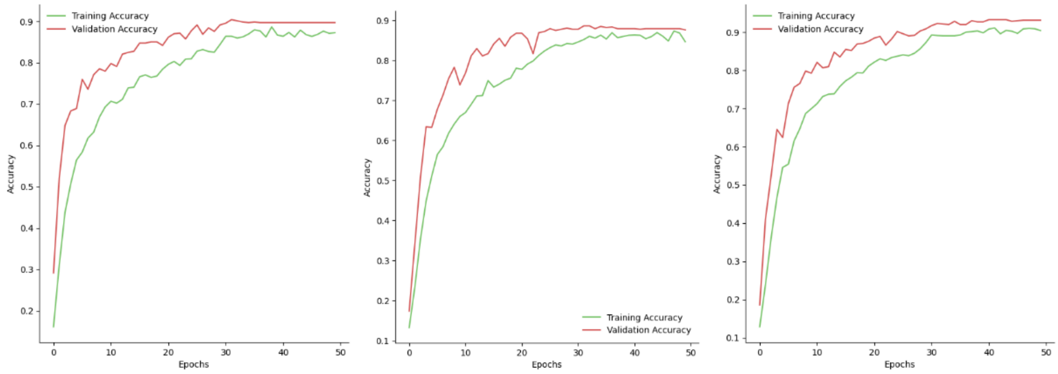



Kernel Size 3x3 Kernel Size 5x5 Kernel Size 7x7

**Figure 6 Results of Base CNN Model CNN2**

## Base Model CNN3

The testing of base CNN model CNN3 is carried out as listed in Table 1. In this process, training and validation are performed using the Nadam optimizer with a learning rate of 0.001, for 50 epochs, using categorical classification. Figure 7 shows the accuracy, validation, and loss plot generated for the base CNN model CNN3.




Kernel Size 3x3 Kernel Size 5x5 Kernel Size 7x7

**Figure 7 Results of Base CNN Model CNN3**

## Result Base Model CNN Evaluation

The test results from each base CNN model show that the application of the optimizer has a significant impact on determining the performance of the proposed model, especially in terms of accuracy. The accuracy results of the testing for the base CNN models can be seen in the following Table 2.

**Table 2 Evaluation Results of Each Base CNN Model**

| Base Model CNN1 Optimizer Adam | | | | | |
|---|---|---|---|---|---|
| **Kernel Size** | **Accuracy** | **Precision** | **Recall** | **Specificity** | **F1-Score** |
| 3x3 | 0.87 | 0.87 | 0.87 | 0.99 | 0.86 |
| 5x5 | 0.89 | 0.89 | 0.89 | 0.99 | 0.89 |
| 7x7 | 0.89 | 0.89 | 0.89 | 0.99 | 0.89 |
| **Base Model CNN2 Optimizer AdamW** | | | | | |
| **Kernel Size** | **Accuracy** | **Precision** | **Recall** | **Specificity** | **F1-Score** |
| 3x3 | 0.85 | 0.85 | 0.85 | 0.99 | 0.84 |
| 5x5 | 0.82 | 0.83 | 0.82 | 0.99 | 0.82 |
| 7x7 | 0.87 | 0.87 | 0.87 | 0.99 | 0.87 |
| **Base Model CNN3 Optimizer Nadam** | | | | | |
| **Kernel Size** | **Accuracy** | **Precision** | **Recall** | **Specificity** | **F1-Score** |
| 3x3 | 0.90 | 0.90 | 0.90 | 0.99 | 0.90 |
| 5x5 | 0.90 | 0.91 | 0.90 | 0.99 | 0.91 |
| 7x7 | 0.91 | 0.91 | 0.91 | 0.99 | 0.91 |

## Result Ensembel Combination (Proposed)

The results of the proposed Ensemble Combination aim to measure the effectiveness of the model in improving accuracy and stability of classification by combining the prediction results from various filter sizes in each base CNN model. The use of ensemble methods, such as Majority Voting and Weighted Average, allows the model to reduce potential errors that may occur in each base CNN model. Through this combination, it is expected that the model can provide more accurate and consistent results in classifying brain tumor types. The test results from the three model scenarios applied can be seen in Table 3, which summarizes the accuracy comparison obtained from each model combination and performance evaluation.

**Table 3 Result Ensembel Combination (Proposed)**

| Ensemble Combination | Accuracy | Precision | Recall | Specificity | F1-Score |
|---|---|---|---|---|---|
| Majority Voting (1) | 0.94 | 0.95 | 0.94 | 0.99 | 0.94 |
| Weighted Average (1) | 0.95 | 0.95 | 0.95 | 0.99 | 0.95 |
| Majority Voting (2) | 0.91 | 0.92 | 0.91 | 0.99 | 0.91 |
| Weighted Average (2) | 0.92 | 0.93 | 0.92 | 0.99 | 0.92 |
| Majority Voting (3) | 0.96 | 0.96 | 0.96 | 0.99 | 0.96 |
| Weighted Average (3) | 0.96 | 0.96 | 0.96 | 0.99 | 0.96 |

# Conclusions

Performance analysis shows that CNN with the Nadam optimizer (CNN3) has the highest accuracy (0.90–0.91) compared to Adam (CNN1) with 0.87–0.89 and AdamW (CNN2) with 0.82–0.87. The application of ensemble combination improves accuracy, with CNN1 reaching 0.94–0.95, CNN2 at 0.91–0.92, and CNN3 achieving the best accuracy (0.96) in both Majority Voting and Weighted Average. Overall, the ensemble combination enhances the effectiveness of classifying 17 types of brain tumors. Further research is recommended to explore various optimizers and combinations of optimizers to improve model performance. The use of larger and more diverse datasets, along with the application of image augmentation techniques, is expected to improve the model's generalization ability. Additionally, experiments with newer model architectures and hybrid optimizer approaches may contribute to improving brain tumor classification accuracy and provide further insights into the effectiveness of each method in various contexts.

# References


I. B. Santoso, Shoffin Nahwa Utama, Supriyono. (2024). A New Voting of Convolutional Neural Networks for Brain Tumor Detection Based on MRI Images. *International Journal of Intelligent Engineering and System*, 17(1), 212-227. Available online at: https://doi.org/10.22266/ijies2024.0229.21

R. Zahid, Y. Ma, U. Inam, & ... (2024). Integrating Convolutional Neural Networks with Attention Mechanisms for Magnetic Resonance Imaging-Based Classification of Brain

Tumors. *Bioengineering*, 11(7): 701–701, Available online at: https://doi.org/10.3390/bioengineering11070701.

I. B. Santoso, Supriyono, Shoffin Nahwa Utama. (2024). Multi-Model of Convolutional Neural Networks for Brain Tumor Classification in Magnetic Resonance Imaging Images. *International Journal of Intelligent Engineering and Systems*, 17(5): 741-758. Available online at: https://doi:10.22266/ijies2024.1031.56.

Ruqsar Zaitoon, Hussain Syed. (2023). Enhanced Brain Tumor Detection and Classification in MRI Scans using Convolutional Neural Networks. *IJACSA*, 14(9), Available online at: https://10.14569/IJACSA.2023.0140929.

Hassan Ali Khan, Wu Jue, Muhammad Umer Mushtaq & ... (2020). Brain tumor classification in MRI image using convolutional neural network. *MBE*, 17(5): 6203–6216. Available online at: https://doi:10.3934/mbe.2020328.

Saeed Mohsen, Wael Mohamed Fawaz, Abdel-Rehim & ... (2023). Convolutional Neural Network for Automatic Brain Tumor Detection. *Proceedings of Engineering and Technology Innovation*, (24):15-21. Available online at: https://doi.org/10.46604/peti.2023.10307.

Alwas Muis, Sunardi, Anton Yudhana. (2023). Medical image classification of brain tumor using convolutional neural network algorithm. *JURNAL INFOTEL*. 15(3): 227-232. Available online at: https://10.20895/infotel.v15i3.964.

Eko Hari Rachmawanto, Christy Atika Sari, Folasade Olubusola Isinkaye. (2024). A good result of brain tumor classification based on simple convolutional neural network architecture. *Telecommunication Computing Electronics and Control*. 22(3): 711-719. Available online at: https://doi.org/10.12928/TELKOMNIKA.v22i3.25863

Fernando Feltrin. *Brain Tumor MRI Images 17 Classes*. Available online at: https://www.kaggle.com/datasets/fernando2rad/brain-tumor-mri-images-17-classes.

Mohamed Amine Mahjoubi, Soufiane Hamida, Oussama El Gannour, & ... (2023). Improved Multiclass Brain Tumor Detection using Convolutional Neural Networks and Magnetic Resonance Imaging. *International Journal of Advanced Computer Science and Applications*. 14(3): 406-4014. Available online at: https://doi.org/10.14569/ijacsa.2023.0140346

X. Xiao, M. Yan, S. Basodi, C. Ji, and Y. Pan. (2020). Efficient Hyperparameter Optimization in Deep Learning Using a Variable Length Genetic Algorithm, *arXiv*.

YD Zhang, C. Pan, J. Sun, and C. Tang. (2018). Multiple sclerosis identification by convolutional neural network with dropout and parametric ReLU. *J. Comput. Sci.* vol. 28, pp. 1–10, Available online at: https://doi.org/10.1016/j.jocs.2018.07.003

Abdullah A. Asiri, Toufique Ahmed Soomro, Ahmed Ali Shah, & ... (2024). Optimized Brain Tumor Detection: A Dual-Module Approach for MRI Image Enhancement and Tumor Classification. IEEE Access, (12): 42868-42887. Available online at: https://doi.org/10.1109/ACCESS.2024.3379136.

Naveen Mukkapati, M.S. Anbarasi. (2022). Brain Tumor Classification Based on Enhanced CNN Model. *Revue d'Intelligence Artificielle*. 36(1):125-130. Available online at: https://doi.org/10.18280/ria.360114.

Nihal Remzan, Karim Tahiry, Abdelmajid Farchi. (2022). Brain tumor classification in magnetic resonance imaging images using convolutional neural network. *International Journal of Electrical and Computer Engineering (IJECE)*. Vol. 12, No. 6, 6664-6674. Available online at: https://doi.org/10.11591/ijece.v12i6.pp6664-6674

A. Muis, S. Sunardi, A. Yudhana. (2024). CNN-based Approach for Enhancing Brain Tumor Image Classification Accuracy. *International Journal of Engineering (IJE)*. Vol. 37, No. 05, 984-996. Available online at: https://doi.org/10.5829/ije.2024.37.05b.15.

Nazik Alturki, Muhammad Ume, Abid Ishaq, & ... (2023). Combining CNN Features with Voting Classifiers for Optimizing Performance of Brain Tumor Classification. *Cancers*. 15, 1767. Available online at: https://doi.org/10.3390/cancers15061767.

Gergo Bogacsovics, Balazs Harangi, Andras Hajdu (2024). Developing diverse ensemble architectures for automatic brain tumor classification. *Multimedia Tools and Applications*. Available online at: https://doi.org/10.1007/s11042-024-19657-z

Anand Vatsala, Sheifali Gupta, Deepali Gupta, & ... (2023). Weighted Average Ensemble Deep Learning Model for Stratification of Brain Tumor in MRI Images, Diagnostics. Available online at: https:/10.3390/diagnostics13071320.

Anindya Nag, Hirak Mondal, Md Mehedi Hassan, & ... (2024). TumorGANet: A Transfer Learning and Generative Adversarial Network-Based Data Augmentation Model for Brain Tumor Classification. in IEEE Access, vol. 12, pp. 103060-103081, Available online at: https://doi.org/ 10.1109/ACCESS.2024.3429633.

Nassar, S.E., Yasser, I., Amer, HM, Mohamed, MA. (2023). A robust MRI based brain tumor classification via a hybrid deep learning technique. Jurnal Superkomputing. 80: 2403–2427. Available online at: https://doi.org/10.1007/s11227-023-05549-w