

## MRI-Based Brain Tumor Classification Using Deep Learning Models

### MRI Tabanlı Beyin Tümörü Görüntülerinin Derin Öğrenme ile Sınıflandırılması

Rawan ABOALQARAYA<sup>1</sup>, Fatih ÖZYURT<sup>2</sup>, Canan Koç<sup>3</sup>, Esra YÜZGEÇ ÖZDEMİR<sup>4</sup>

<sup>1,2,3,4</sup> Department of Software Engineering, Engineering Faculty, Firat University, Elazığ, Turkey.

<sup>1</sup> [rawan.qaraya@gmail.com](mailto:rawan.qaraya@gmail.com), <sup>2</sup> [fatihozyurt@firat.edu.tr](mailto:fatihozyurt@firat.edu.tr), <sup>3</sup> [canan.koc@firat.edu.tr](mailto:canan.koc@firat.edu.tr), <sup>4</sup> [eyuzgec@firat.edu.tr](mailto:eyuzgec@firat.edu.tr)

<sup>1</sup>(ORCID: 0009-0008-0995-2629), <sup>2</sup>(ORCID: 0000-0002-8154-6691), <sup>3</sup>(ORCID: 0000-0002-2651-9471)

#### Abstract

This study proposes an automatic classification approach for brain tumors using magnetic resonance imaging (MRI). The dataset comprises four classes: glioma, meningioma, pituitary tumor, and no tumor. We evaluated a custom-designed convolutional neural network (CNN) alongside transfer learning models such as VGG16, VGG19, MobileNetV2, and ResNet50. To enhance model generalization, data augmentation techniques were employed during training. Model performance was measured using accuracy, classification reports, and confusion matrices. The custom CNN model achieved the highest test accuracy of 90.08%, followed closely by MobileNetV2. These findings suggest that transfer learning models do not always outperform custom architectures, and dataset-specific, optimized models can yield superior results.

**Keywords:** brain tumor, deep learning, convolutional neural network, transfer learning, MRI classification, mobilenetv2

#### 1. INTRODUCTION

Brain tumors are complex neoplasms that affect the central nervous system and are often associated with severe morbidity. These tumors are classified into various subtypes, including gliomas, meningiomas, and pituitary adenomas, each of which exhibits distinct biological behaviors and clinical outcomes. Early and accurate diagnosis is crucial for improving patient survival and quality of life, particularly in malignant tumors such as gliomas. However, conventional diagnostic methods are often based on invasive procedures and radiologic expert opinion and are prone to human error and interpreter dependency.

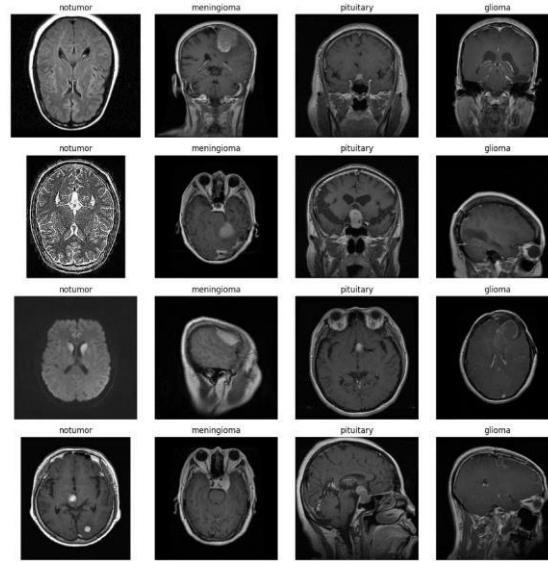
Magnetic Resonance Imaging (MRI) is the most widely used imaging modality for the non-invasive diagnosis of brain tumors. MRI enables high-resolution examination of tumor structure, localization, and its relationship with surrounding tissues, thanks to its soft tissue contrast. However, since manual analysis of MRI data is both time-consuming and subjective, there is an increasing need for fast and accurate automated diagnostic systems. Especially in resource-limited clinical settings, such decision support systems reduce the workload and standardize the diagnostic process. In recent years, the success of deep learning techniques in image processing has enabled the development of automated systems for brain tumor diagnosis. In this study, we propose a classification system that can distinguish between four different classes of brain tumors based on MRI images. A specially designed convolutional neural network (CNN) model is compared with common transfer learning architectures (VGG16, VGG19, MobileNetV2, and ResNet50), and the generalizability of the model is improved by data augmentation methods. The results show that architectures specifically optimized for specific datasets can be more successful than pre-trained deep networks.

Brain tumors are serious health problems affecting the central nervous system, and early diagnosis is crucial for effective treatment. Medical imaging techniques such as Magnetic Resonance Imaging (MRI) are commonly used for the detection

<sup>1</sup>Rawan ABOALQARAYA

and classification of tumors. However, manual analysis of these images is time-consuming and subject to interpretation. Therefore, the need for artificial intelligence-based automatic systems has increased [1]. The use of deep learning methods in medical imaging has gained significant momentum in recent years. In particular, convolutional neural networks (CNNs) and more advanced architectures have achieved great success in brain tumor classification [2].

In this study, we propose a deep learning-based method to classify brain tumors from MRI images. Our goal is to develop a model that can accurately analyze brain tumor images and distinguish among four classes: glioma, meningioma, pituitary tumor, and no tumor [3].



**Figure 1.** Sample MRI images from the dataset

Brain tumor classification has made significant progress in recent years with the widespread use of deep learning methods in the field of medical imaging. Below are some important studies conducted in this area:

Afshar et al. (2019) used Capsule Networks (CapsNet) for the classification of brain tumor types from MRI images. CapsNet aims to overcome the limitations of traditional convolutional neural networks (CNNs). In their study, the proposed CapsNet architecture was tested on the Figshare dataset and achieved an accuracy of 86.56%. This result was attributed to CapsNet's ability to work effectively with limited data and to model spatial relationships more efficiently. Moreover, this architecture is more robust to tumor orientation and location variations and can learn effectively even with less data [1].

Yang et al. (2023) compared deep learning and traditional machine learning algorithms in brain tumor classification. Using 2D Convolutional Neural Networks (CNN) and convolutional autoencoder models, they achieved an accuracy of 93.45%, demonstrating the effectiveness of deep learning methods. The study showed that the 2D CNN model achieved the best performance with an AUC score of 0.99, highlighting the significant advantage of deep learning methods over traditional approaches [4].

Kang et al. (2021) proposed a hybrid method combining deep learning-based feature extraction with traditional machine learning classifiers for brain tumor classification. The study adopted a transfer learning approach, using 13 different pre-trained CNN models to extract deep features from MRI images. These features were then evaluated using nine different machine learning classifiers, including Support Vector Machines (SVM), Random Forests (RF), k-Nearest Neighbors (k-NN), Gaussian Naive Bayes, and Extreme Learning Machines (ELM). For each dataset, features obtained from the top

three performing CNN models were selected and combined to create a “feature-level ensemble.” This combined feature set was again classified using various machine learning algorithms. Experimental results showed that, especially for large datasets, the SVM classifier with a radial basis function (RBF) kernel achieved the highest accuracy rates (98.67% in binary classification and 93.72% in multi-class classification). This study demonstrates that integrating deep learning and traditional machine learning techniques is an effective approach to achieve high accuracy in brain tumor classification tasks [5].

Haq et al. (2021) demonstrated how a deep learning-based paradigm achieved high success rates in classifying glioma, meningioma, and pituitary tumors using the Figshare and BraTS 2018 datasets. The model achieved 96.5% accuracy and a Dice Similarity Coefficient (DSC) of 94.3% on the BraTS 2018 dataset. Furthermore, performance improvements were made using techniques such as data augmentation, intensity normalization, and conditional random fields (CRF) [6]. Nassar et al. (2023) developed a hybrid deep learning model and achieved an accuracy rate of 99.31% on 3,064 T1-weighted contrast-enhanced MRI images from 233 patients. Pre-trained models such as GoogleNet, AlexNet, and SqueezeNet were employed for brain tumor classification, and their outputs were combined using majority voting to obtain more reliable results [7].

Dal, Eliaçık, and Işık (2023) proposed a deep learning-based approach for brain tumor classification using convolutional neural networks (CNN). Their study aimed to accurately classify different tumor types (glioma, meningioma, and pituitary tumors) from brain MR images. By leveraging the automatic feature extraction capability offered by deep learning, the classification process was performed without the need for traditional image processing steps. A large dataset was used for training and testing the developed model; the reported accuracy at the end of training was 98%. Additionally, the proposed CNN architecture provided significant advantages over traditional methods in terms of training time and classification accuracy. This work highlights that deep learning-based approaches have the potential to provide highly accurate diagnoses in medical image analysis [8]. Aslan (2022) focused on the automatic detection of brain tumors by developing a deep learning-based model. Using convolutional neural networks (CNN), brain tumors were accurately classified from MRI images. The author reported high accuracy for this deep learning model and indicated that it outperformed traditional methods. The model's accuracy was reported as 96%, demonstrating how powerful deep learning can be in brain tumor diagnosis systems. The study also discussed the impact of pre-processing techniques used to improve the model's accuracy. The results show that deep learning techniques can play an important role in medical image analysis and potentially be useful in clinical applications [9].

**Table 1.** Approaches used for brain tumor detection in the related literature

Author(s)	Year	Method	Dataset	Performance / Results
Afshar Et Al.	2019	Capsule Networks (CapsNet)	Figshare dataset	86.56% accuracy
Yang Et Al.	2023	2D CNN, Convolutional Autoencoders	MRI images	93.45% accuracy, 0.99 AUC score
Kang Et Al.	2021	Deep features + Machine Learning classifiers	MRI images	98.67% accuracy (RBF-SVM, binary class), 93.72% accuracy (RBF-SVM, multi-class)
Haq Et Al.	2021	Deep learning-based model (BraTS 2018)	BraTS 2018, Figshare	96.5% accuracy, 94.3% DSC
Nassar Et Al.	2023	Hybrid deep learning, majority voting	T1-weighted MRI images	99.31% accuracy
Dal, Eliaçık, And Işık	2023	Deep learning, convolutional neural networks (CNN)	MRI images	98% accuracy

Aslan	2022	Deep learning	MRI images	96% accuracy
-------	------	---------------	------------	--------------

## 1.1. MATERIAL AND METHOD

### 1.2. Dataset

The brain MRI images used in this study were obtained from the Kaggle platform and consist of a combined dataset from Figshare, SARTAJ, and Br35H datasets. The dataset contains a total of 7,023 images, divided into four classes: glioma, meningioma, pituitary, and no tumor. Some mislabeled images in the dataset were corrected, and the images underwent preprocessing steps, including resizing and edge cleaning, before training. This dataset was used for training deep learning models for brain tumor classification.

Since the images in the dataset have varying dimensions, all images were resized to 240x240 pixels prior to model training. Additionally, pixel values were normalized from the range 0–255 to 0–1. These operations enable the model to learn more quickly and efficiently. Due to the limited number of images in medical imaging, data augmentation techniques were applied to reduce the risk of overfitting and improve the model’s generalization ability. Random rotations ( $\pm 20$  degrees), zooming (10%), and horizontal flipping were applied to the training images, creating different variations of each image. This approach made the model more robust against the variability within the dataset.

### 1.3. Models

In this study, both a simple CNN model designed from scratch and pre-trained models, including VGG16, VGG19, MobileNetV2, and ResNet50, were used. For the pre-trained models, transfer learning was applied by freezing the base model weights and adding new classification layers at the end. The models were trained for 20 epochs using the Adam optimizer and categorical cross-entropy loss function. Additionally, EarlyStopping and ReduceLROnPlateau callbacks were employed to prevent overfitting. This model is a custom-built neural network created from scratch. It consists of three convolutional (Conv2D) layers, followed by max-pooling layers to increase depth. Finally, a flatten, fully connected (Dense), and dropout layer lead to the output. A dropout rate of 50% was applied to reduce the overfitting tendency. This architecture was chosen as a basic starting point due to its low computational cost and ability to deliver results quickly.

VGG16 is a 16-layer deep convolutional neural network pre-trained on the ImageNet dataset [10]. In this study, transfer learning was applied by freezing the pre-trained layers and retraining only the last layers. Due to its deep architecture and small filter sizes, VGG16 is effective in extracting detailed features. It is a widely preferred powerful architecture, especially for image classification problems. VGG19 is an improved version of the VGG16 model, consisting of 19 layers [10]. Although deeper, it uses a similar structure with small filter sizes and repeated blocks. This model was also adapted via transfer learning with only the last layers retrained. While its depth increases the capacity to learn more complex features, it requires longer training times and has a higher risk of overfitting.

MobileNetV2 is a lightweight CNN architecture optimized for mobile and embedded systems [11]. It has lower depth but high efficiency, enabling faster operation with fewer parameters. In this study, the base weights were frozen and only the classification layers were trained. MobileNetV2 provided balanced performance with low resource usage and high accuracy.

ResNet50 is a 50-layer deep convolutional network characterized by the use of “residual blocks” [12]. This structure was developed to mitigate the vanishing gradient problem commonly encountered in deep networks. In this study, the model was adapted with transfer learning; however, its deep architecture made training more challenging compared to other models. Especially when trained with limited data and augmented samples, its performance was lower.

The performances of the models were evaluated on the test dataset using various metrics. These included accuracy, confusion matrix, and classification reports comprising precision, recall, and F1-score. Accuracy graphs for each model’s training process were plotted to analyze learning trends. Additionally, class-wise correct and incorrect classifications were visualized through confusion matrices. The CNN model achieved the highest test accuracy of 90.08%, followed by

MobileNetV2 with 89.09%. The VGG16 and VGG19 models attained accuracies of 80.7% and 79.32%, respectively, while ResNet50 showed the lowest performance with 67.42%. These results reveal the discrimination power of the models across classes and their fit to the dataset.

To assess the success of the classification models, the relationship between the predicted class labels and the true class values was examined. This evaluation was conducted based on true positive (TP), false positive (FP), true negative (TN), and false negative (FN) values obtained from the confusion matrix. Several evaluation criteria were used to measure model performance, including accuracy, which indicates the overall success rate; recall (sensitivity), which measures the proportion of true positive samples correctly identified; and specificity, which reflects the proportion of true negatives correctly classified. Additionally, precision, which measures the accuracy of positive predictions; F1-score, the harmonic mean of precision and recall; and Matthews Correlation Coefficient (MCC), which provides more reliable results in cases of class imbalance, were included. These metrics not only reflect overall success but also reveal strengths and weaknesses of the model at the class level, allowing for detailed performance analysis. The formulas for these evaluation criteria used in this study are presented in Equations (1) - (6) [8-13].

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100 \quad (1)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100 \quad (2)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100 \quad (3)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100 \quad (4)$$

$$\text{F1-Score} = 2 \times \frac{\text{Sensitivity} \times \text{Specificity}}{\text{Duy} + \text{Specificity}} \times 100 \quad (5)$$

$$\text{MCC} = 2 \times \frac{\text{TP} \times \text{TN} - \text{FP} \times \text{FN}}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}} \quad (6)$$

## 2. EXPERIMENTAL RESULTS

The experimental studies were carried out to evaluate the effectiveness of the proposed model and to present a comparative performance analysis with other common deep learning architectures. In this context, a specially developed convolutional neural network (CNN) model and common transfer learning-based models such as VGG16, VGG19, MobileNetV2, and ResNet50 were tested on a four-class dataset of MRI images. Model performances are analyzed in detail through accuracy, classification reports (precision, recall, F1-score), and complexity matrices. In addition, the effect of data augmentation techniques applied during training on generalization performance was observed. The findings presented below highlight the primary contributions of the study, encompassing both model performance and architectural choices.

Table 2 presents in detail the classification performance of the proposed CNN model on a four-class brain tumor dataset. The model achieved a high success rate of 90.08% in terms of the overall accuracy metric, demonstrating the reliability of the model. In particular, the highest performance was observed in the “No Tumor” class with 100% recall and an F1 score of 0.93%, indicating that the model has a high sensitivity in recognizing healthy samples. However, for the “Pituitary” tumor class, a very successful classification was also achieved with an F1 score of 0.96% and a recall of 0.99%.

On the other hand, the “Meningioma” class shows a relatively lower performance, with a recall of 0.76% and an F1 score of 0.80%, indicating that the model has a more cautious success in discriminating this tumor type. This is likely due to the fact that meningioma images are more visually similar to those of the other classes. In terms of macro and weighted averages (macro avg / weighted avg), 0.90% was achieved in all three metrics, suggesting that the model maintains its

generalizability despite the imbalance between classes. The results show that a specially designed CNN architecture can offer a competitive alternative to transfer learning-based models, emphasizing the importance of dataset-specific, optimized architectures.

**Table 2.** CNN Classification Results

Class	Precision	Recall	F1-Score	Support
<b>Glioma</b>	0.95	0.83	0.88	300
<b>Meningioma</b>	0.86	0.76	0.80	306
<b>No Tumor</b>	0.87	1.00	0.93	405
<b>Pituitary</b>	0.94	0.99	0.96	300
<b>Accuracy</b>			<b>0.90</b>	1311
<b>Macro Avg</b>	0.90	0.89	0.90	1311
<b>Weighted Avg</b>	0.90	0.90	0.90	1311

Table 3 details the performance of the transfer learning-based VGG16 model on the four-class brain tumor classification task. The model achieves an overall accuracy of 79%, which is lower than that of the custom-designed CNN architecture. Especially for the “No Tumor” class, a very successful classification was achieved with an F1 score of 0.91% and a recall of 0.93%. This shows that the VGG16 model can distinguish healthy samples with high accuracy. In addition, the 0.97% recall and 0.84% F1 score for “Pituitary” tumor shows that the model can effectively recognize this class as well.

In contrast, the model performs poorly for the “Meningioma” class, with a recall of 0.50% and an F1 score of 0.58%, indicating that this tumor type has a high misclassification rate. For the “Glioma” class, the F1 score of 0.76% indicates that the model performs at an average level in this category. The macro and weighted average metrics are 0.77% and 0.78% respectively, indicating that the imbalance between classes affects the model performance. These results suggest that transfer learning-based architectures may not always achieve the performance of custom models optimized for the dataset.

**Table 3.** VGG16 Classification Results

Class	Precision	Recall	F1-Score	Support
<b>Glioma</b>	0.80	0.73	0.76	300
<b>Meningioma</b>	0.69	0.50	0.58	306
<b>No Tumor</b>	0.89	0.93	0.91	405
<b>Pituitary</b>	0.75	0.97	0.84	300
<b>Accuracy</b>			<b>0.79</b>	1311
<b>Macro Avg</b>	0.78	0.78	0.77	1311
<b>Weighted Avg</b>	0.79	0.79	0.78	1311

Table 4 shows the performance of the VGG19 transfer learning architecture on the four-class brain tumor classification task. The model produced an overall accuracy of 79%, which is almost on par with VGG16. The 0.91% F1 score and 0.93% recall value obtained for the “No Tumor” class show that the model can recognize healthy

individuals with high accuracy. Similarly, the “Pituitary” tumor class performed well with 0.97% recall and 0.84% F1 score. These results show that VGG19 offers stable prediction capability in some classes.

On the other hand, a very poor performance was observed for the “Meningioma” class with 0.50% recall and 0.58% F1 score. This indicates that the VGG19 model has difficulty in distinguishing meningioma tumors from other classes. With an F1 score of 0.76% for the “Glioma” class, it can be said that the model recognizes this class at a moderate level. The macro and weighted averages are 0.77%-0.78% respectively, indicating that the overall success of the model is limited due to the imbalance between classes. As a result, although the VGG19 architecture shows acceptable performance in certain classes, it lags behind the custom CNN architecture in terms of overall success.

**Table 4.** VGG19 classification report

Class	Precision	Recall	F1-Score	Support
<b>Glioma</b>	0.80	0.73	0.76	300
<b>Meningioma</b>	0.69	0.50	0.58	306
<b>No Tumor</b>	0.89	0.93	0.91	405
<b>Pituitary</b>	0.75	0.97	0.84	300
<b>Accuracy</b>			<b>0.79</b>	1311
<b>Macro Avg</b>	0.78	0.78	0.77	1311
<b>Weighted Avg</b>	0.79	0.79	0.78	1311

Table 5 presents the performance of MobileNetV2, a transfer learning architecture, in a four-class brain tumor classification task. The model achieved an overall accuracy of 89%, which is very close to that of the CNN architecture developed specifically for this task. Especially in the “No Tumor” class, the model's ability to recognize healthy individuals with high accuracy is remarkable, with an F1 score of 0.96 and a recall value of 0.97. Similarly, the “Pituitary” tumor class showed a very strong classification performance with 1.00% recall and 0.91% F1 score. These results demonstrate that MobileNetV2 exhibits strong generalization performance despite its parametric efficiency.

The model produced an F1 score of 0.88% for the “Glioma” class and an F1 score of 0.78% for the “Meningioma” class, indicating a very good performance in these two tumor types. In particular, the performance in the “Meningioma” class reveals that MobileNetV2 offers a more balanced approach in this class, where larger models such as VGG16 and VGG19 lag behind. The macro and weighted averages were each in the range of 0.88% to 0.89%, indicating a relatively balanced distribution across classes. In this context, the MobileNetV2 model can be considered a viable and competitive solution in terms of both performance and efficiency.

**Table 5.** MobileNetV2 Classification Results

Class	Precision	Recall	F1-Score	Support
<b>Glioma</b>	0.94	0.83	0.88	300
<b>Meningioma</b>	0.83	0.74	0.78	306
<b>No Tumor</b>	0.94	0.97	0.96	405
<b>Pituitary</b>	0.84	1.00	0.91	300
<b>Accuracy</b>			<b>0.89</b>	1311

<b>Macro Avg</b>	0.89	0.88	0.88	1311
<b>Weighted Avg</b>	0.89	0.89	0.89	1311

Table 6 shows the performance of ResNet50, a deep transfer learning model, on the four-class brain tumor classification problem. The overall accuracy of the model is 67%, which is quite low compared to both the specially designed CNN architecture and other transfer learning based models. For the “No Tumor” class, an F1 score of 0.80% and a recall value of 0.86% were obtained, and this class was the most successfully recognized category by the model. The “Pituitary” tumor class also performed relatively well with an F1 score of 0.72%, but these levels were not enough to improve the overall success.

The model had serious classification problems, especially in the “Meningioma” class. The 0.32% recall and 0.41% F1 score obtained for this class indicate that the model frequently misclassifies meningioma samples. Similarly, only a 0.65% F1 score was obtained for the “Glioma” class. The macro and weighted averages are in the range of 0.65%-0.66% for all metrics, indicating that the model fails to provide a balanced and reliable classification across classes. These results suggest that the ResNet50 model exhibits low generalization capability on the current dataset and that more optimized or lightweight architectures may be needed for such medical imaging tasks.

**Table 6.** ResNet50 Classification Results

<b>Class</b>	<b>Precision</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Support</b>
<b>Glioma</b>	0.64	0.67	0.65	300
<b>Meningioma</b>	0.57	0.32	0.41	306
<b>No Tumor</b>	0.74	0.86	0.80	405
<b>Pituitary</b>	0.66	0.80	0.72	300
<b>Accuracy</b>			<b>0.67</b>	1311
<b>Macro Avg</b>	0.65	0.66	0.65	1311
<b>Weighted Avg</b>	0.66	0.67	0.66	1311

Table 7 compares the test accuracies of the proposed custom CNN model and four different transfer learning architectures (MobileNetV2, VGG16, VGG19, and ResNet50). The results show that the custom CNN architecture shows the highest performance with an accuracy of 90.08%. This was followed by MobileNetV2, which achieved an accuracy of 89.09%, demonstrating a highly competitive performance despite its parametric efficiency and low computational requirements. VGG16 and VGG19 produced very close results with 80.70% and 79.32% accuracy, respectively, but lagged behind the specialized CNN and MobileNetV2 architectures. The lowest accuracy was 67.42% for ResNet-50, indicating that the model is not well-adapted to the current dataset. This comparison strongly suggests that custom architectures optimized for specific datasets can outperform transfer learning-based approaches.

**Table 7.** Comparison of The Models

<b>Model</b>	<b>Test Doğruluğu</b>
<b>CNN</b>	% 90.08
<b>MobileNetV2</b>	% 89.09
<b>VGG16</b>	% 80.70
<b>VGG19</b>	% 79.32

### 3. CONCLUSION

In this study, a deep learning approach for automatic classification of brain tumors from MRI images is presented. The specially developed convolutional neural network (CNN) model was able to discriminate four different tumor classes with high accuracy and outperformed all transfer learning models with a test accuracy of 90.08%. The high F1 scores, especially in the “No Tumor” and “Pituitary” classes, showed that the model can effectively classify both healthy and pathological samples. MobileNetV2 architecture was the closest competitor with an accuracy of 89.09%, showing a strong performance despite the low number of parameters.

The results of the comparative analysis show that transfer learning-based models do not always guarantee the best performance; on the contrary, custom models optimized for the data set may produce better results. The findings demonstrate the value of customized artificial intelligence solutions in the development of clinical decision support systems and pave the way for future studies with different imaging types, tumor subclasses, and multiple modalities. Furthermore, it is suggested that data augmentation and weighted learning strategies be integrated more effectively to overcome important problems, such as class imbalance.

#### 3.1. REFERENCES

- [1] Afshar P, Mohammadi A, Plataniotis KN. “Brain tumor type classification via capsule networks”. 2019 IEEE International Conference on Image Processing, Taipei, Taiwan, 22–25 September 2019.
- [2] Chakrabarty A, Bhunia AK. “A survey on brain tumor detection using image processing techniques”. International Journal of Scientific & Technology Research, 9(1), 1727–1731, 2020.
- [3] Nickparvar M. “Brain Tumor MRI Dataset”. <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>.
- [4] Yang W, Al-Khassaweneh M, Alzghool M, Fraiwan L, Aljarrah I, Al-Fahoum A. “MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques”. BMC Medical Informatics and Decision Making, 23(1), 1–19, 2023.
- [5] Kang J, Ullah Z, Gwak J. “MRI-based brain tumor classification using ensemble of deep features and machine learning classifiers”. Sensors, 21(6), 2222, 2021.
- [6] Haq EU, Azam MA, Khan MA, Saba T, Rehman A, Tariq U, Miceli M. “A hybrid deep learning-based paradigm to detect brain tumor from MRIs”. Journal of Ambient Intelligence and Humanized Computing, 13, 225–241, 2021.
- [7] Nassar M, Almotairi S, Alshmrany S, Alsheikhy A. “A hybrid deep learning model for brain tumor classification based on majority voting technique”. The Journal of Supercomputing, 2023. <https://doi.org/10.1007/s11227-023-05549-w>
- [8] Dal A, Eliaçık B, Işık AH. “Beyin tümörü MR görüntülerinin evrimsel sinir ağları ile sınıflandırılması”. Uluborlu Mesleki Bilimler Dergisi, 6(1), 71–91, 2023.
- [9] Aslan M. “Derin öğrenme tabanlı otomatik beyin tümör tespiti”. Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 34(1), 399–407, 2022.
- [10] Simonyan K, Zisserman A. “Very deep convolutional networks for large-scale image recognition”. International Conference on Learning Representations (ICLR), San Diego, USA, 7–9 May 2015.
- [11] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. “MobileNetV2: Inverted residuals and linear bottlenecks”. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 18–23 June 2018.
- [12] He K, Zhang X, Ren S, Sun J. “Deep residual learning for image recognition”. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 27–30 June 2016.

- [13] Japkowicz N, Shah M. Evaluating Learning Algorithms: A Classification Perspective. Cambridge, UK, Cambridge University Press, 2011.