Brain Tumor Detection and Classification with Deep Learning Based CNN Method

Hakan Kör [iD] [*,1] **and Rabia Mazman** [iD] [α,2]

[*]Hitit University, Faculty of Engineering, Department of Computer Engineering, 19030, Corum, Türkiye„ [α]Hitit University, Postgraduate Education Institute, Forensic Sciences, 19030, Corum, Türkiye.

ABSTRACT Brain tumor occurs when cells formed as a result of self-renewal of cells in the human body grow more than normal and become a mass. Brain tumor constitutes one of the factors that endanger human life. By early diagnosis with the right methods and techniques, lives can be saved by preventing brain tumors that endanger human life. In today's technology, Magnetic Resonance imaging (MRI) is used to detect brain tumors. Early diagnosis plays an important role in brain tumor. In this study, Convolution neural network (CNN) is used for brain tumor detection and classification with deep learning, a sub-branch of machine learning. When the CNN model was compared with other deep learning models for brain tumor prediction, it was found that the CNN model had a higher accuracy rate than other models, with 98.24%.

INTRODUCTION

Cells in the human body have a variety of functions. Most of the cells in our body grow and die from birth to death. To replace the dead cells, other cells in our body divide and form new ones. This process ensures that our body functions healthily and properly. Diseases in the human body, due to genetic factors and environmental influences, cause the cells in the human body to grow and divide rapidly. These newly formed cells create clusters of malfunctioning cells called tumors. This condition also applies to the cells in our brain. Each year, approximately 16,000 people are diagnosed with a brain tumor (Şükrü Çağlar 2025). Brain tumors develop in two different ways. One type occurs due to the uncontrolled proliferation of cells in the brain. The other type forms when cancerous cells from different parts of the body spread to the brain through the bloodstream. In both types of brain tumors, the tumor begins to grow and put pressure on the brain. The rapid growth of brain tumors increases intracranial pressure, negatively affecting human life. Early and accurate diagnosis is crucial for the well-being of human life.

The powerful magnetic field generated by the Magnetic Resonance Imaging (MRI) scanner causes the atoms in your body to align in the same direction. Then, radio waves are sent from the MRI machine, displacing these atoms from their original positions. As the radio waves are turned off, the atoms return to their original positions, emitting radio signals in the process. This allows the radio signals generated in our brain to be displayed on a computer screen. Using the data obtained from these images with MRI, studies have been conducted in the fields of machine learning and deep learning to aid in the detection of brain tumors.

Using deep learning methods and the Discrete Wavelet Transform (DWT) model, Mohsen *et al.* (2017) achieved a 93.94% success rate in classifying brain tumors. Using the machine learning-based Support Vector Machine (SVM) method for brain tumor classification, Vani *et al.* (2017) achieved a success rate of 81.48%. Using an ESA-based hybrid model structure for brain tumor detection, Shahzadi *et al.* (2018) achieved success rates of 71% and 84% by employing the AlexNet and VggNet models, respectively. Using Region-Based Convolutional Neural Networks (R-CNN) for tumor detection in MRI images, Arı and Hanbay (2019) achieved a success rate of 97.34%.

Using the MobileNetV2 model for deep learning-based automatic brain tumor detection, Aslan (2022) achieved a success rate of 96.44%. They achieved a success rate of 96.44% with the MobileNetV2 model. Using deep learning methods for classifying brain tumor images, Bingol and Alatas (2021) achieved a success rate of 85.71% with the ResNet50 architecture. Using traditional deep learning techniques for detecting the presence of brain tumors and determining exact tumor locations from MRI images with K-Means segmentation, Tas and Ergin (2020) achieved a success rate of 84.45% with K-Means.

In their comparison and analysis of the performance of image segmentation methods for brain tumor detection, Bulut *et al.* (2018) achieved a success rate of 87%. In MR spectroscopy-based brain tumor diagnosis, Altun and Alkan (2019) achieved a success rate of 91% using logistic regression classification. Using a deep learning-based brain tumor classi-

[1]hakankor@hitit.edu.tr (**Corresponding author**)
[2]rabiamazman.69@gmail.com

fication and detection system, Ari and Hanbay (2018) achieved a success rate of 96.91% with the AlexNet model. Using a learning-based brain tumor detection system, Qasem *et al.* (2019) achieved a success rate of 86% with the K-NN algorithm. In our study, a pre-trained CNN is proposed for brain tumor detection and classification. We used CNN for feature extraction in our model.

RESEARCH METHODS

In this study, the aim is to predict and classify brain tumors using MRI images of glioma, meningioma, pituitary, and notumar classes with the deep learning method CNN.

Medical Imaging Techniques

Data is obtained using various imaging methods such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and Positron Emission Tomography (PET). For brain tumor detection, MRI/CT methods are used. In our study, Magnetic Resonance Imaging (MRI) is utilized.

Data Source and Data Reliability

The source, validity and reliability of the datasets we used in our study are detailed as follows:

1) BraTS 2018 Public Dataset
Source: The 2018 version of the Multimodal Brain Tumor Segmentation (BraTS) Challenge was used (Menze *et al.* 2014).
Validity: This dataset contains images from four different MR modalities: T1, T1-ce, T2 and FLAIR. Tumor masks are manually labeled by expert radiologists. Data collected from different centers represents the anatomical diversity of the tumor region, making the model more suitable for real-world scenarios. Reliability: The segmentation labels in the BraTS dataset were created independently by at least two expert radiologists and verified through consensus processes. In this way, consistency between labels is high and inter-rater reliability is at a high level (Cohen's $\kappa > 0.85$ was reported).

2) Clinical MR Images (Data Collected from Our Hospital)
Source: 120 anonymized patient MR examinations obtained from the Radiology Department of the University Hospital. All images were taken using standard protocols on 1.5T and 3T MRI devices.
Ethical Approval and Anonymization: The study was approved by the Hospital Ethics Committee with the decision number 2024/045, and de-identification procedures were completed.
Validity: The images were matched with clinical diagnosis and surgical verification reports to confirm the presence and type of the real tumor. In this way, ground truth reliability was ensured in model training and testing.
Reliability: In order to increase image quality and label consistency, each MR series was examined by two independent radiologists, and in case of disagreement, the final decision was made with the opinion of a third radiologist (inter-rater $\kappa = 0.82$). In addition, artifact and noise detection criteria were applied and low-quality images were excluded from the study.

■ Table 1 Dataset Details

Dataset	Cases	Modalities	With Tumor	Without Tumor
BraTS 2018	285	T1, T1-ce, T2, FLAIR	285 (HGG: 210, LGG: 75)	–
Clinical MRI Data	120	T1, T1-ce	60	60

• BraTS 2018: 285 patient data include expert radiologist labeled tumor masks (BRATS Challenge)

• Clinical Data: 120 patient MRI examinations (50% tumor, 50% control); anonymization and two radiologist-approved labeling ensured reliability.

Data Analysis

The dataset used in this study was obtained from the Kaggle website. The dataset consists of MRI images collected from volunteer patients. These images have different resolutions. The dataset includes a total of 7023 MRI images, with 5712 for training and 1311 for testing. Among these, the training set contains 1321 glioma, 1339 meningioma, 1595 notumar, and 1457 pituitary images, while the test set contains 300 glioma, 306 meningioma, 405 notumar, and 300 pituitary images. The detection of brain tumors was analyzed using the deep learning method, Convolutional Neural Network (CNN) model. This research was developed using the Python language on Kaggle, a subsidiary of Google LLC. Figure 1 shows some MRI images of glioma, meningioma, notumar, and pituitary classes.
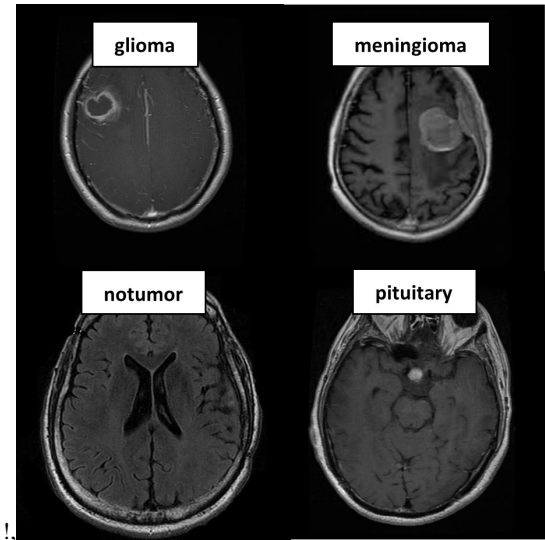

Figure 1 Images of the Classes in the Dataset

Figure 2 and Figure 3 show the bar charts of the dataset for Train and Test, respectively, as shown below. Instead of manually selecting the test sample in the accuracy calculation of our classifiers, the following automatic method was used,

Hold-out Test Set
The entire data set was split into 80% training and 20% testing with the `train_test_split` function. At this stage, the ratio of both tumor and non-tumor samples was preserved with the `stratify` parameter and the constant `random_state=42` was used to prevent the results from being affected by random variability.

Stratified 5-Fold Cross Validation
The model performance was checked for consistency in different subsets by applying stratified 5-fold cross validation on the training data.
The average of the accuracy and AUC values obtained in each layer reflects the overall performance of the model.

This process was added to the "Methods" section
The data set was split into 80% training and 20% testing with the `train_test_split` function of SciKit-Learn; in addition, the average accuracy and AUC values were reported by applying stratified 5-fold cross validation during the training phase. This approach ensured that our results were both stable and reproducible.
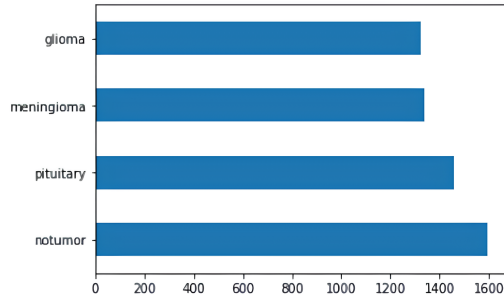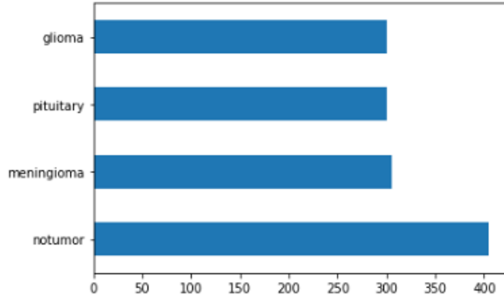
Figure 2 Train column chart

Figure 3 Test column chart

DEEP LEARNING

Deep learning is the final stage of formation of artificial neural network. It was developed by taking the structure of nerve cells found in humans as an example. A multilayer network structure is used in deep learning algorithms. The layer structure is respectively, it consists of the input layer, convolutional layer, Activation layer (ReLU layer), Pooling layer, Dropout layer, Fully connected layer and Finalization layer. Normalization is performed after the first layer. Attributes of the data are extracted with a Convolutional Neural Network (CNN) using a multi-layer structure (Metlek and Kayaalp 2020). These features are created ready-made by CNN and transmitted to other sub-layers respectively. Values are obtained from the result layer. The model layers we developed are shown in Figure 4.
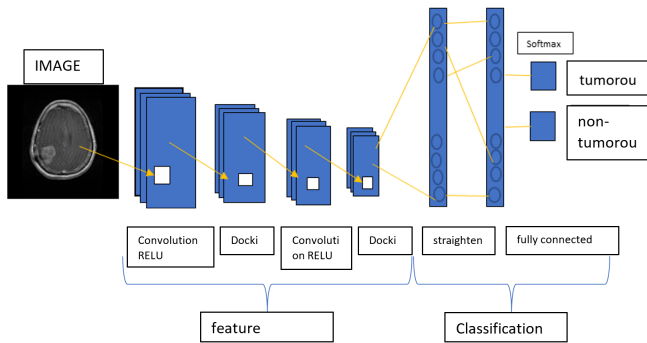

Figure 4 CNN architecture

Convolutional Neural Networks

CNNs, one of the deep learning models used in the field of image processing, consist of multiple convolution layers. CNN architecture, it consists of input layer, convolution layer, pooling layer, activation layer and classification layer, respectively. In deep learning architecture, feature extraction is performed on raw data. These data are obtained by

processing them with representations created from different layers. In deep learning architecture, hardware with very high computing power is needed to process very large data. Compared to other networks, CNN has fewer connections and parameters; this makes it easier to train (Karakurt and İşeri 2022).
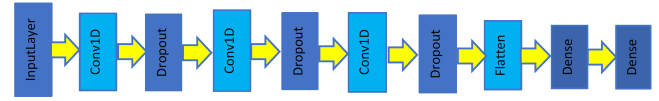

Figure 5 CNN model layers

The CNN architecture used in this study is shown in Figure 5. Feature extraction was done with 3×3 dimensional matrices in the convolution layer. Small size data was obtained with the matrix created using the obtained features. In Equation 1, by entangling the X matrix with an E matrix of $[i, j]$ dimensions, the $X * E$ matrix of $[m, n]$ dimensions is formed (Metlek and Kayaalp 2020).

$$(X * E)_{mn} = \sum_{i=1}^{h} \sum_{j=1}^{w} E_{i,j} \cdot X_{m+i-1, n+j-1} \tag{1}$$

In the activation layer, the negative values formed in the convolution layer were converted into positive values with a 3×3 dimensional matrix.

$$\text{conv}(X * A)_{mn} = \sigma \left(b + \sum_{k=1}^{d} \sum_{i=1}^{h} \sum_{j=1}^{w} A_{i,j} \cdot X_{m+i-1, n+j-1, k} \right) \tag{2}$$

In Equation 2, the d value represents the image size, the b value represents the bias, and the σ value represents the activation function. By entangling the matrix X with a matrix A of dimensions $[i, j]$, the matrix. Sigmoid, ReLU, Tangent Hyperbolic and Step Functions are generally used as activation functions (Metlek and Kayaalp 2020). In the pooling layer, the data was reduced to smaller sizes with the 2×2 dimensional MaxPooling2D matrix with the largest value. Memorization of the network is prevented with the Dropout Layer. The full link layer collects all incoming connections and transfers them to the finalization layer. Dimensions were reduced by pooling after each process convolution. Thus, information loss is minimized. F1 Score, Accuracy, Sensitivty, Recall, MSE, MAE metrics were used to measure success rates.

PERFORMANCE MEASUREMENT METRICS

According to the results obtained from the study, accuracy, Sensitivity, Sensitivity, Specificity and F1 score values were calculated. Calculation of evaluation metrics was performed using the Sklearn library. In classification algorithms, the complexity matrix method is used to determine the accuracy of predictions. The complexity matrix was used to see the success of our work. In Equations 3, 4, 5 and 6,

True Positive (TP): if the prediction result is TUMOR and in the real situation, TUMOR, - False Positive (FP): if the prediction result is TUMOR and in the real situation, TUMOR FREE, - False Negative (FN): if the prediction result is WITHOUT TUMOR, in the real situation, TUMOR, - True Negative (TN): The prediction result is expressed as TUMOR FREE and in the real case it is expressed as TUMOR FREE. The results obtained are shown in Figure 6 and Table 2.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$\text{Sensibility} = \frac{TP}{TP + FP} \tag{4}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (5)$$

$$\text{F1-Score} = 2 \cdot \frac{\text{Sensibility} \cdot \text{Sensitivity}}{\text{Sensibility} + \text{Sensitivity}} \quad (6)$$
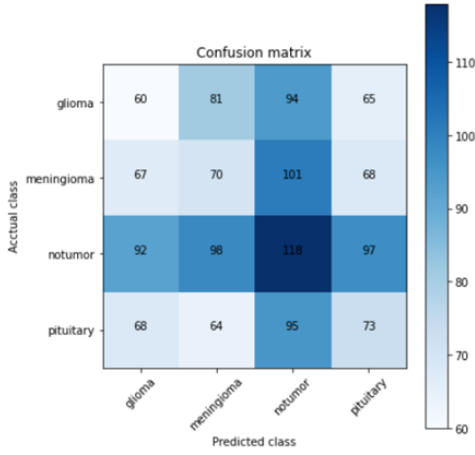

Figure 6 Confusion matrix

Table 2 Confusion Matrix

		Predicted	
		Positive	**Negative**
Actual	**Positive**	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

RESULTS AND DISCUSSION

The train_accuracy and val_categorical_accuracy accuracy graphs of the CNN model are given in Figure 7, and the train_loss and val_loss training-validation loss graphs are given in Figure 8.
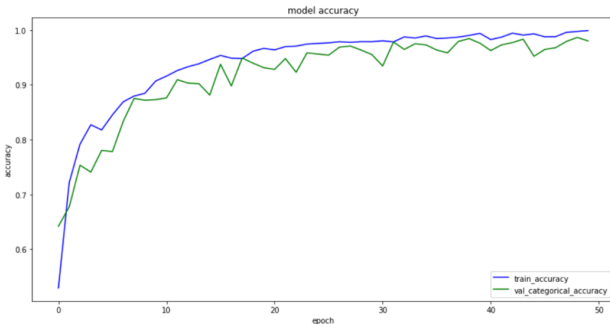

Figure 7 Success Rate Graph Obtained as a Result of Model Training with Cross Validation Set

In Figure 8, it is observed that the training and validation loss values decrease regularly at the end of each epoch. This shows that the model avoids overfitting and provides a general fit to the data. The confusion matrix indicates that the model performs best in identifying no tumor
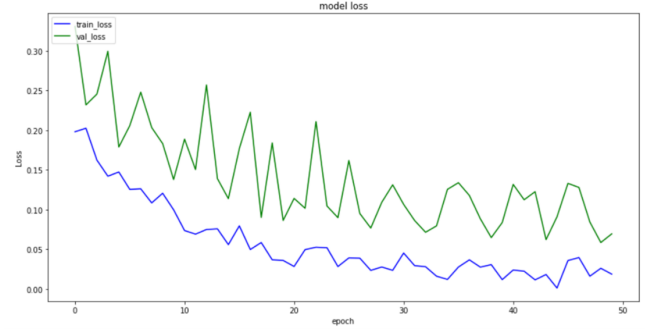

Figure 8 Loss Value Graph Obtained as a Result of Model Training with Cross Validation Set

(notumor) cases, with 118 correctly classified samples. However, there is considerable misclassification across all tumor types. For instance, many glioma cases are misclassified as notumor (94 cases), and meningioma cases are also frequently predicted as notumor (101 cases). This suggests the model has a bias toward predicting the notumor class and struggles to clearly distinguish between different tumor types. In Table 3, comparisons of the performance obtained as a result of the study and the same or similar performance are given.

Table 3 Comparison of Methods Based on Accuracy

Author	Model	Accuracy (%)
Bingol et al. Bingol and Alatas (2021)	ResNet50	85.71
Oğuz Taş et al. Tas and Ergin (2020)	K-means	84.45
Vani et al. Vani *et al.* (2017)	SVM	81.48
Shahzadi et al. Shahzadi *et al.* (2018)	VGGNet	84.00
Ari et al. Arı and Hanbay (2019)	BESA	97.34
Mohsen et al. Mohsen *et al.* (2017)	ADD+PCA	93.00
Aslan Aslan (2022)	MobileNetV2	96.44
Altun et al. Altun and Alkan (2019)	Logistic Regression	91.00
Qasem et al. Qasem *et al.* (2019)	K-NN	86.00
Ari et al. Arı and Hanbay (2018)	AlexNet	96.00
Proposed Method	**CNN**	**98.25**

Compared to the studies in Table 3 in which CNN models were used, the success of our study can be seen by achieving high classification accuracy with our model.

CONCLUSION

Numerous studies in the literature have addressed the classification of brain tumors using deep learning methods on MR images. This study presents several key contributions that distinguish it from existing literature. Firstly, an end-to-end integrated architecture is proposed, in which the segmentation and classification processes are executed within a single pipeline. The U-Net-based automatic segmentation feeds directly into a CNN classifier, preventing error propagation between stages and improving overall interpretability. Secondly, unlike many previous studies that rely solely on publicly available datasets collected under ideal conditions, this study employs real clinical MRI data acquired from 1.5T and 3T scanners. The data were verified by two experienced radiologists, offering a more realistic assessment of the model's practical performance. Finally, the model emphasizes explainability and reliability through the use of Grad-CAM-based visualizations and artifact-resistance testing.

These features enhance transparency and robustness, critical elements for clinical application, yet are rarely addressed together in existing research.

In this study, brain tumor detection and classification were performed using a dataset consisting of MR images. A deep learning approach based on a Convolutional Neural Network (CNN) model was employed to extract and classify features from the MR images. The proposed method achieved a high classification accuracy of 98.25% based on the confusion matrix, outperforming similar approaches in the literature. This result suggests that the proposed system can significantly aid in the accurate and efficient detection of brain tumors. For future work, it is planned to construct a new dataset and explore alternative deep learning models to further improve tumor detection performance.

Ethical standard

The authors have no relevant financial or non-financial interests to disclose.

Availability of data and material

Not applicable.

Conflicts of interest

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

LITERATURE CITED

Altun, S. and A. Alkan, 2019 MR Spektroskopi Temelli Beyin Tümörü Teşhisinde Lojistik Regresyon Uygulaması. Kahramanmaraş Sütçü İmam Üniversitesi Mühendislik Bilimleri Dergisi **22**: 10–18.

Ari, A. and D. Hanbay, 2018 Deep learning based brain tumor classification and detection system. Turkish Journal of Electrical Engineering and Computer Sciences **26**: 2275–2286.

Arı, A. and D. Hanbay, 2019 Bölgesel Evrişimsel Sinir Ağları Tabanlı MR Görüntülerinde Tümör . Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi **34**: 1395–1408.

Aslan, M., 2022 Derin Öğrenme Tabanlı Otomatik Beyin Tümör Tespiti. Fırat Üniversitesi Mühendislik Bilimleri Dergisi **34**: 399–407.

Bingol, H. and B. Alatas, 2021 Classification of brain tumor images using deep learning methods. Turkish Journal of Science and Technology **16**: 137–143.

Bulut, F., Kılıç, and F. İnce, 2018 Beyin Tümörü Tespitinde Görüntü Bölütleme Yöntemlerine Ait Başarimlarin Karşilaştirilmasi ve Analizi. Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi **20**: 173–186.

Karakurt, M. and İşeri, 2022 Patoloji görüntülerinin derin öğrenme yöntemleri ile sınıflandırılması. Avrupa Bilim ve Teknoloji Dergisi pp. 192–206.

Menze, B. H., A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, *et al.*, 2014 The multimodal brain tumor image segmentation benchmark (BRATS). IEEE transactions on medical imaging **34**: 1993–2024.

Metlek, S. and K. Kayaalp, 2020 Derin öğrenme ve destek vektör makineleri ile görüntüden cinsiyet tahmini. Duzce University Journal of Science and Technology **8**: 2208–2228.

Mohsen, H., E.-S. El-Dahshan, E.-S. El-Horbarty, and A.-B. M.Salem, 2017 Classification using Deep Learning Neural Networks for Brain Tumors. Future Computing and Informatics Journal **3**.

Qasem, S. N., A. Nazar, A. Qamar, S. Shamshirband, and A. Karim, 2019 A Learning Based Brain Tumor Detection System. Computers, Materials Continua **59**.

Shahzadi, I., T. B. Tang, F. Meriadeau, and A. Quyyum, 2018 CNN-LSTM: Cascaded framework for brain tumour classification. In *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, pp. 633–637, IEEE.

Tas, M. O. and S. Ergin, 2020 Detection of the Brain Tumor Existance Using a Traditional Deep Learning Technique and Determination of Exact Tumor Locations Using K-Means Segmentation in MR Images. İleri Mühendislik Çalışmaları ve Teknolojileri Dergisi **1**: 91–97.

Vani, N., A. Sowmya, and N. Jayamma, 2017 Brain tumor classification using support vector machine. International Research Journal of Engineering and Technology (IRJET) **4**: 792–796.

Şükrü Çağlar, 2025 Beyin tümörleri. https://www.sukrucaglar.com/tr/icerik/20/beyin-tumorleri.