Computers and Electronics in Medicine
Cutting-Edge Scientific Solutions

e-ISSN: 3023-8609
RESEARCH ARTICLE
Vol.3 / No.1 / 2026 / pp.77-85
https://doi.org/10.69882/adba.cem.2026019

Classification of Brain MRI Images using Deep Learning: The DeiT3 Model and the Use of Feature Fusion Methods

Sena Kahraman [*,1] and **Mesut Toğaçar** [α,2]
[*]Technology and Information Management, Institute of Social Sciences, Fırat Unıversity, Elazığ 23119, Türkiye, [α]Department of Management Information Systems, Faculty of Economics and Administrative Sciences, Fırat Unıversity, Elazığ 23119, Türkiye.

ABSTRACT Brain tumors are among the diseases that can seriously threaten human life and can be fatal. Early diagnosis of brain tumors plays a crucial role in the treatment process of the disease. However, accurately and quickly diagnosing this disease remains one of the significant challenges of modern medical technologies. Currently, advanced imaging techniques such as magnetic resonance imaging (MRI) are generally used for detecting brain tumors. This study proposes an artificial intelligence-based diagnostic approach using MRI images that include brain tumor types and consist of four classes. The proposed approach includes preprocessing, model training, feature fusion, and selection as final steps. In the preprocessing step, Grad-CAM and LBP techniques are applied to the original dataset, resulting in a total of three datasets, including the original one. These datasets are then trained with the DeiT3 model to obtain three separate feature sets (original, Grad-CAM-based, LBP-based). The feature sets are fused using a feature fusion technique, and the performance of the combined sets is evaluated using SVM methods. Feature selection methods (Chi2, Relief) are applied to the best-performing Grad-CAM & LBP-based feature set to highlight the most efficient features. Experimental analysis results show that a success rate of 99.5% was achieved using the SVM method.

KEYWORDS
Brain tumor
Transformer model
Image processing
Feature fusion
Feature selection

INTRODUCTION

A brain tumor is a mass formed by the abnormal growth of cells located in the brain. The brain is the control center of the body, enabling humans to perform their basic life functions. Therefore, the detection and treatment of brain tumors are also very important for human life. It has been observed that brain tumor disease significantly affects mortality rates (Yılmaz 2023). It has been reported that the number of people dying from brain tumor disease in China is between 50,000 and 100,000 per year, and that approximately 80,000 people are diagnosed with brain tumors in the United States each year (Taşdemir and Barışçı 2024). When looking at causes of death, brain tumors rank tenth worldwide and eighth in Türkiye (Erçelik and Hanbay 2023b).

Medical imaging techniques, which are also used in the diagnosis of many diseases, are used to detect brain tumors. Medical imaging techniques include techniques such as computed tomography, magnetic resonance imaging (MRI), mammography, etc. The most preferred imaging technique for detecting brain tumors is MRI. Erçelik et al. stated that it is difficult to detect brain tumors using only MRI images and that this is due to the complex structure of the brain. The difficulty doctors face in diagnosing diseases using such traditional methods and the advancement of technology have led to the emergence of the concept of "Artificial Intelligence in Health." Today, approaches such as image processing and deep learning are used for disease detection (Serttaş and Deniz 2023). There are multiple types of brain tumors. Some studies extract features for classification, while others incorporate deep learning methods (Aslan 2022).

Numerous artificial intelligence-based studies have been conducted on brain tumors in the literature. Examining some recent studies, Aslan (2024) presents the LSTM-ESA model, which combines LSTM (Long Short-Term Memory) and ESA (Convolutional Neural Network Numerous artificial intelligence-based studies have been conducted on brain tumors in the literature. E s). He states that this model achieved a score of 98.1% in brain tumor detection. He notes that this score is higher than the score achieved by the ESA model. Demirel and Soylu (2024) aimed to compare the performance of MobileNet, DenseNet-121, and DenseNet-201 models in terms of brain tumor detection in their study. It was stated that all models achieved excellent accuracy rates during the training phase, but the DenseNet-121 model showed the best performance in terms of generalization. It was stated that an accuracy rate of 98.81% was achieved with the DenseNet-121 model. Erçelik and Hanbay (2023a) aimed to compare the Gaussian Filtering and ResNet50 models and identify the model with the best performance. Noise components in the images were cleaned us-

Manuscript received: 11 November 2025,
Revised: 15 January 2026,
Accepted: 15 January 2026.

[1]sena_kahraman2002@hotmail.com (**Corresponding author**)
[2]mtogacar@firat.edu.tr

ing the Gaussian function. It was determined that cleaning the noise yielded better performance compared to ResNet50. Yenikaya and Oktaysoy (2023) aimed to test the success of ResNet101 and GoogLeNet models in brain tumor detection in their study. At the end of model training, it was found that the ResNet101 model achieved better success than the GoogLeNet model with 91.5%. The ResNet model achieved 87.9% success. Das *et al.* (2025) developed a method using a VGG-16-based deep learning model to accurately classify brain tumors in FLAIR MR images. The model classified these images, which had three different tumor classes, with 99% accuracy. In their study, Yang *et al.* (2025) developed an artificial intelligence model called GMDNet to accurately classify brain tumor MR images. With this model, they successfully classified the relationship between different MR images. The GMDNet model also provided high accuracy in the presence of missing data. To achieve this, a special method called reuse modality was developed.

In this study, data diversity was increased and the model's learning capacity was enhanced by applying Grad-CAM and LBP techniques to the original data set. The model was trained using next-generation technology-based transformers. The features extracted from the model training offer an innovative approach that is more powerful and contributes to performance thanks to the feature fusion technique. Feature selection was applied to the dataset that yielded the best performance as a result of feature fusion, and the most efficient features were obtained. Consequently, time and cost savings were also achieved. This developed approach produces reliable outputs in the brain tumor detection process and makes an important contribution to the early diagnosis of the disease. This will help in rapid detection and determining the correct treatment methods. A brief summary of the other sections of the article is as follows: Information about the data set and model training is provided in the second section. Detailed information about the analysis results is provided in the third section. The fourth section contains the discussion. The final section, the fifth section, contains the conclusions.

MATERIALS AND METHODS

In this study, a hybrid approach incorporating the DeiT3 model and feature fusion is proposed for the automatic classification of brain MRI images according to brain tumor types. The proposed hybrid model consists of data set preparation, preprocessing steps, model training, feature fusion, and feature selection steps.

Dataset

In this study, a new data set was used, created from a combination of the Figshare (Cheng 2024), Sartaj dataset (Bhuvaji 2025), and Br35H (Hamada 2020) data sets, which consist of brain MRI images. This data set was obtained from the open-access Kaggle website (Nickparvar 2021). The dataset contains 7,023 brain MRI images in JPG format, divided into four classes. The class types consist of glioma, meningioma, pituitary diseases, and non-tumor MRI images. The non-tumor class of the Br35H dataset was not included in the combined new dataset. By class type, there are 1,621 glioma, 1,645 meningioma, 1,757 pituitary, and 2,000 non-tumor images. The class types are evenly distributed. The images in the dataset have variable resolution. A sample subset of images from the dataset is shown in Figure 1.

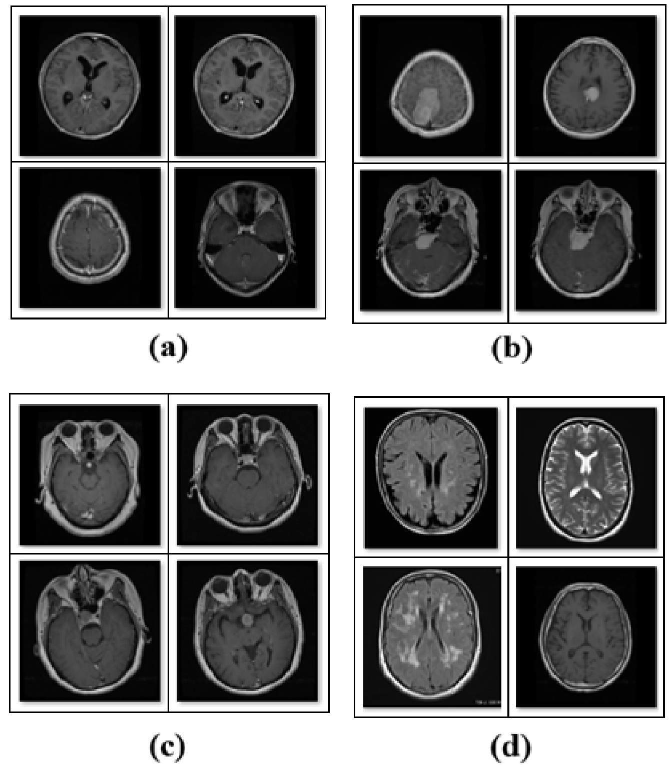

Figure 1 Sample images of the class types in the dataset: (a) glioma, (b) meningioma, (c) pituitary, and (d) non-tumor.

Image processing methods

Gradient-weighted class activation mapping (Grad-CAM) is a technique used to visualize the regions that a convolutional neural network (CNN) model considers most important when making predictions in the form of a heatmap (Senjoba *et al.* 2024). Grad-CAM helps understand which regions influence an image processing model to select a specific class during classification (Livieris *et al.* 2023). This method is also widely used in different fields such as object detection and medical image analysis. It is a general and flexible method that can be applied to different CNN architectures (Ennab and Mcheick 2025). In this study, Grad-CAM was trained using the DarkNet-19 model architecture. This facilitated the detection of diseased regions in images and allowed Grad-CAM to focus on these regions in the images. Sample images from the new dataset obtained by applying the Grad-CAM technique to the original dataset are shown in Figure 2.

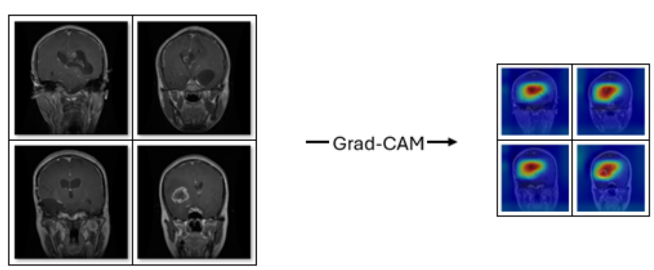

Figure 2 Images from the original dataset and the dataset created using Grad-CAM.

Local Binary Pattern (LBP) is a widely used method in image processing and computer vision for extracting texture features from images (Almohamade *et al.* 2025). Its computational efficiency and adaptability to different scenarios demonstrate its versatility. It produces a binary code consisting of 0s and 1s by comparing the intensity of each pixel in an image with its neighboring pixels. The resulting binary codes describe the local textural structure of the image. These codes are converted into a histogram, which is a numerical representation summarizing the texture in the image (Attallah 2025). This histogram provides features (Aydemir 2022). The steps of the LBP method are shown in Figure 3. Sample images

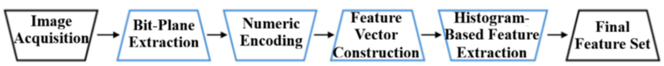

Figure 3 Stages of the LBP method.

from the new dataset obtained by applying the LBP technique to the original dataset are shown in Figure 4.

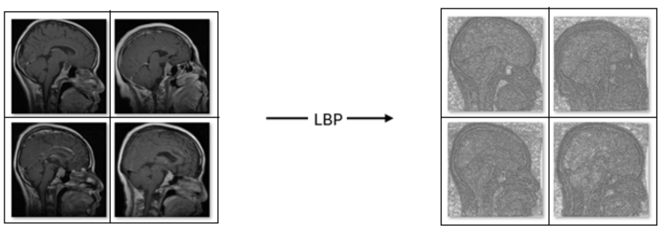

Figure 4 Images from the original dataset and the dataset created by applying LBP.

New generation transformer model: DeiT3

Transformer models typically process data in six stages. Sometimes, additional layers may be included in this basic process due to structural differences in these models. The first of these stages involves dividing the received image into smaller pieces. The divided image pieces are converted into vectors by passing through an embedding layer. Location information is added to each piece to ensure that the model preserves the spatial order. Subsequently, information is processed in layers through encoder blocks, and deep features are extracted. The Multi-Head Self-Attention mechanism identifies relationships between image fragments and enables the model to generate contextual meanings. The information obtained in the final layer is processed using neural networks and transmitted to the output layer for classification (Toğaçar 2025).

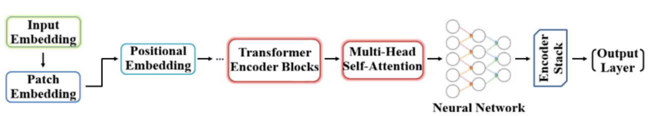

Figure 5 Stages of the LBP method.

Data-efficient image transformers (DeiT) are a transformer-based deep learning architecture that aims for high performance with minimal data. DeiT is derived from the vision transformer (ViT) architecture. It reduces data requirements by utilizing the knowledge distillation method. In this architecture, the input image is typically resized to 224x224 pixels. The image is then divided into small 16x16 pixel patches. Location information and distillation tokens are added to each patch to preserve the structural features of the data. These patches are then converted into a vector format that the transformer can process. The relationships between the pieces are analyzed through the attention mechanism and feedforward layers. The encoder block consists of three recurrent self-attention and feedforward neural network layers. In the final stage, the image labels are determined through the classification layer (Sevinc *et al.* 2025).

Feature selection methods

The chi-square (Chi2) test is a statistical method used for feature selection. This test helps determine how one variable affects another. Specifically, it examines the relationship between the target variable and the features. Features that show a stronger relationship with the target variable are selected, while independent feature variables are removed. This ensures that only important features are included in the model (Rahman *et al.* 2023). The Chi2 test works by calculating the difference between expected and observed frequencies. If there is independence between two features, the Chi2 value will be low. A high Chi2 value indicates a stronger relationship with the target variable. This feature selection method is particularly effective in high-dimensional and complex data sets. However, it can lead to erroneous results in low-frequency cells and cause poor performance in data imbalance (Devi *et al.* 2023).

The RelieF method is an effective feature selection method commonly used in classification. This method aims to improve model performance by analyzing the importance level of different features in the data. For each data point, the nearest neighbors are analyzed; the weight score for each feature is determined using both neighbors from the same class and neighbors from different classes. In improving model performance, features with high weight values have a more significant impact than those with low weights. Features with low weights may contain unnecessary or misleading information. The greatest advantage of the RelieF method is its ability to deliver successful results in high-dimensional, complex, and large datasets (Gür *et al.* 2025).

Support vector machines method

Support vector machines (SVM) effectively perform both regression and classification operations on high-dimensional and complex data sets. Data is transformed into a higher-dimensional space using kernel functions. This makes it easier to capture non-linear relationships between features. The model's performance and confusion are adjusted through the regularization parameter and kernel coefficient (gamma) (Haldar *et al.* 2025). The SVM algorithm is a powerful machine learning technique that works successfully with both linear and non-linear data structures, providing high accuracy by making clear distinctions between classes. It achieves high accuracy by separating classes with the maximum margin (Toğaçar 2025). It can classify with high accuracy despite complex and noisy data (Tubog *et al.* 2025). The parameter information and selected values of the SVM method used in the experimental analyses of this study are given in Table 1.

Proposed Approach

The proposed approach consists of artificial intelligence-based models that analyze CT images to detect brain tumors. In this study, the DeiT3 model from the ViT family forms the basis of this system to successfully detect the disease. This approach combines the DeiT3 model with approaches such as Grad-CAM and LBP to achieve more successful disease detection. The proposed approach consists of preprocessing steps, model training, feature fusion, and post-processing steps such as feature selection.

In the preprocessing step, image processing methods are applied. CT images collected in a hospital environment may have

■ **Table 1** Parameters of the SVM method used

Parameter	Selection/Value
Function	Cubic
Box Restriction Level	1
Core Scale Mode	Automatic
Multi-Class Method	One-on-One

different resolutions. In this step, all images are cropped to a fixed resolution of 224 × 224 pixels. This resolution is chosen because ViT-based models process images of this size as input. Then, new image sets are created using image processing methods (Grad-CAM and LBP) on the original data set. DeiT3, which stands out among the new generation image transformers, is used in the model training step. Feature extraction is applied to the data sets trained with the DeiT3 model. In this step, 768 features are obtained from each of the Grad-CAM (A), original (B), and LBP (C) datasets. Feature fusion is then applied to the datasets, and the best fusion set is selected from the resulting fusion sets. The goal of this step is to determine the most efficient feature set for feature selection in the next step.

In the final processing step, the most significant features are selected from the 1536 features obtained from the best combination set (for this study, the combination of B and C is the best combination set) using Chi2 and Relief techniques, respectively. This approach assigns a score to each of the 1536 features according to its own statistical methods and ranks them. The features are ranked from highest to lowest according to their importance. The most meaningful 1000 or 500 features are selected from this ranking. They are then classified using the SVM method.

The design of the proposed model is shown in Figure 6.

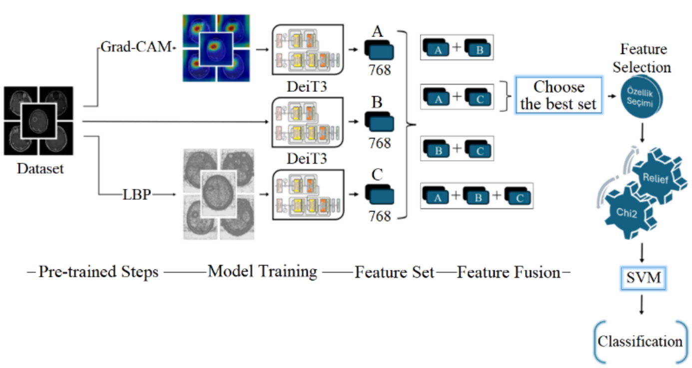

Figure 6 Design of the proposed model.

EXPERIMENTAL RESULTS

During the experimental analyses, MATLAB 2024 software was used to perform feature fusion, feature selection, and classification stages (SVM method). The Python programming language was used for the LBP and Grad-CAM methods, which are preprocessing steps, and for training the DeiT3 model, and these codes were run via the Jupyter Notebook interface. A computer with a 3.40 GHz Intel Core i7 processor, 32 GB RAM, and a 10 GB graphics card was used to perform the experimental analyses. A confusion matrix, which provides the number of correctly and incorrectly classified data in the dataset, was used to evaluate the analysis results (Çelik and Koç 2021). The metrics and formulas for the confusion matrix are given below. There are four metrics: accuracy, f-score, precision, and sensitivity (Kuştaşı and Yağanoğlu 2024). The equations contained in the metrics have four basic classification elements: positive (P), negative (N), true (T), and false (F).

$$\text{Sensitivity (Se)} = \frac{TP}{TP + FN} \tag{1}$$

$$\text{Precision (Pre)} = \frac{TP}{TP + FP} \tag{2}$$

$$\text{F-score (f-scr)} = \frac{TP}{TP + FN} \tag{3}$$

$$\text{Accuracy (Acc)} = \frac{TP + TN}{FP + FN + TP + TN} \tag{4}$$

The preferred parameters in the recommended approach are given in Table 2. Default values have been selected for parameters other than those listed.

■ **Table 2** Parameters used in the proposed approach and selected values

Model / Method	Parameter	Preference / Value
DeiT3	Classifier	Linear
DeiT3	Epoch	11
DeiT3	Learning rate	10-4
DeiT3	Loss function	CrossEntropyLoss
DeiT3	Mini-batch	32
DeiT3	Optimization	SGD
DeiT3	Training & Test rate	%80 – %20
SVM	Preset	Cubic
SVM	Core scale	Auto
SVM	Box restriction level	1
SVM	Multi-class method	One-vs-One

The experimental analysis consists of three stages. The first stage consists of the preprocessing process. In this step, Grad-CAM and LBP methods were applied to the original data set, resulting in three different image sets. The images belonging to these three data sets are shown in Figure 7.

The second stage involves evaluating the performance of the three datasets obtained in the first stage using a model. In this stage, the new-generation DeiT3 model, developed to increase the efficiency of transformation models in image processing, was used. After the three data sets were trained by the model, the training-test success graphs shown in Figure 8 were obtained. The confusion matrices obtained at the end of the model training are given in Figure 9. The overall success of the metric results obtained from the analyses is shown in Table 3. According to the metric values obtained as a result of model training, the original dataset showed 89.75% overall accuracy, the dataset obtained with the LBP method showed 83.27%, and the dataset obtained with the Grad-CAM method showed 98.22%.

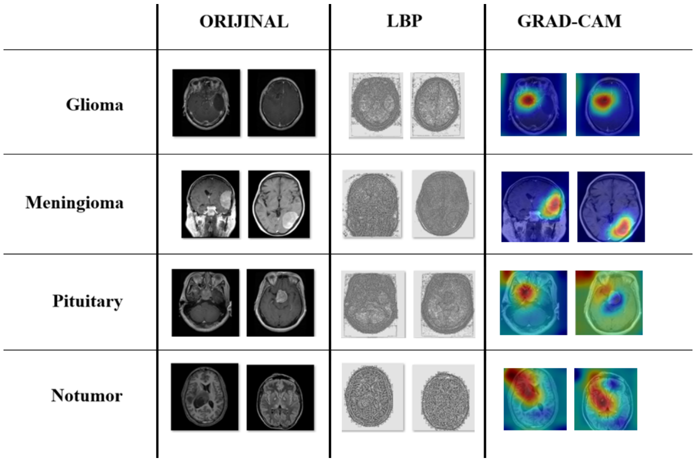

Figure 7 Sample images obtained by applying LBP and Grad-CAM methods to the original dataset.

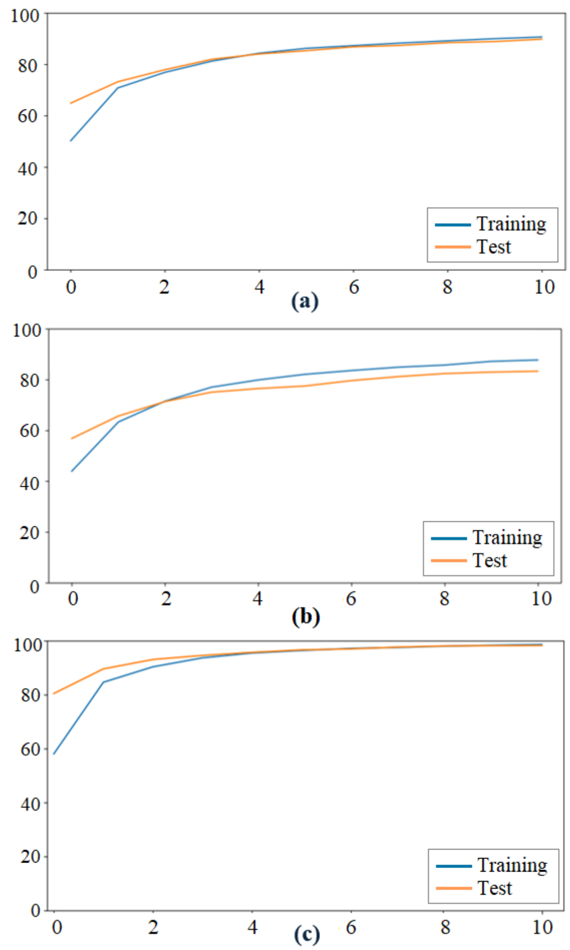

Figure 8 Training-test success graphs of the DeiT3 model; a) Original dataset, b) LBP dataset, c) Grad-CAM dataset.

The third stage involves applying feature fusion to the feature sets extracted from each image set in the previous stage. The goal of this stage is to determine the most efficient combination obtained by applying feature fusion. A feature set of size [image count x 768] was extracted from the final layer of the DeiT3 model. This

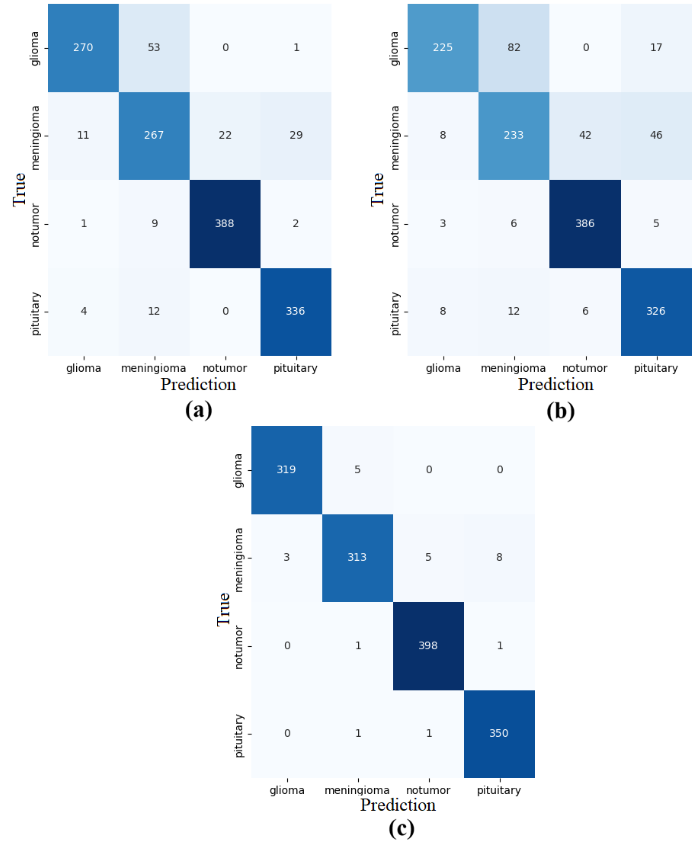

Figure 9 Confusion matrix obtained from training the DeiT3 model; a) Original dataset, b) LBP dataset, c) Grad-CAM dataset.

Table 3 Metric results obtained from the analyses of the DeiT3 model (%)

Dataset	Se	Pre	f-scr	Acc
Original	89.23	89.66	89.34	89.75
LBP	82.34	83.47	82.40	83.27
Grad-CAM	98.13	98.22	98.17	98.22

process was performed individually for all image sets (Original, Grad-CAM, LBP). As a result of feature fusion; [image count x 1536] from the (Grad-CAM & Original) set, [image count x 1536] from the (Original & LBP) set, (Grad-CAM & LBP) set [image count x 1536], (Grad-CAM & Original & LBP) set [image count x 2304]. The SVM method was used to evaluate the performance of the obtained feature sets. At this stage, the training and test ratios were selected in the same proportion as the model training (training data 80%, test data 20%). The confusion matrices obtained from the classification of the feature sets obtained as a result of feature fusion using the SVM method are given in Figure 10. The confusion matrix results are given in Table 4.

Table 5 shows that the (Grad-CAM & LBP) set achieved a general accuracy success rate of 99.42%, while the (Grad-CAM & Original) set achieved a general accuracy success rate of 99.36%. Therefore, it was determined that the (Grad-CAM & LBP) set provides better performance than the (Grad-CAM & Original) set. It was observed that the success of the new sets obtained with the feature

■ **Table 4** Metric results (%) of confusion matrices obtained after feature fusion

Set / Fusion	Class	Se	Pre	F-scr	Acc
GC & LBP	0	97.84	99.69	99.76	
GC & LBP	1	99.69	97.91	98.79	
GC & LBP	2	100	100	100	
GC & LBP	3	99.71	100	99.86	
					99.36
GC & ORJ	0	98.15	99.38	98.76	
GC & ORJ	1	99.39	98.19	98.78	
GC & ORJ	2	100	100	100	
GC & ORJ	3	99.71	99.71	99.71	
GC & ORJ					99.36
LBP & ORJ	0	95.68	99.04	96.91	
LBP & ORJ	1	98.78	94.20	96.32	
LBP & ORJ	2	99.75	99.75	99.75	
LBP & ORJ	3	98.58	99.43	98.97	
LBP & ORJ					98.15
GC & LBP & ORJ	0	97.84	99.37	98.37	
GC & LBP & ORJ	1	99.39	97.90	98.48	
GC & LBP & ORJ	2	99.75	100	99.87	
GC & LBP & ORJ	3	99.71	99.71	99.71	
GC & LBP & ORJ					99.22

■ **Table 5** Metric results (%) of confusion matrices obtained after cross-validation

Set / Fusion	Class	Se	Pre	F-scr	Acc
GC & LBP	0	99.07	99.69	99.38	
GC & LBP	1	99.21	98.85	99.03	
GC & LBP	2	99.80	99.65	99.72	
GC & LBP	3	99.21	99.32	99.26	
					99.42
GC & ORJ	0	99.01	99.63	99.32	
GC & ORJ	1	99.15	98.85	99.00	
GC & ORJ	2	99.85	99.75	99.80	
GC & ORJ	3	99.49	99.32	99.41	
					99.36

fusion approach was higher than that obtained from the DeiT3 model. The proposed approach had a positive impact on overall performance with all steps in the process.

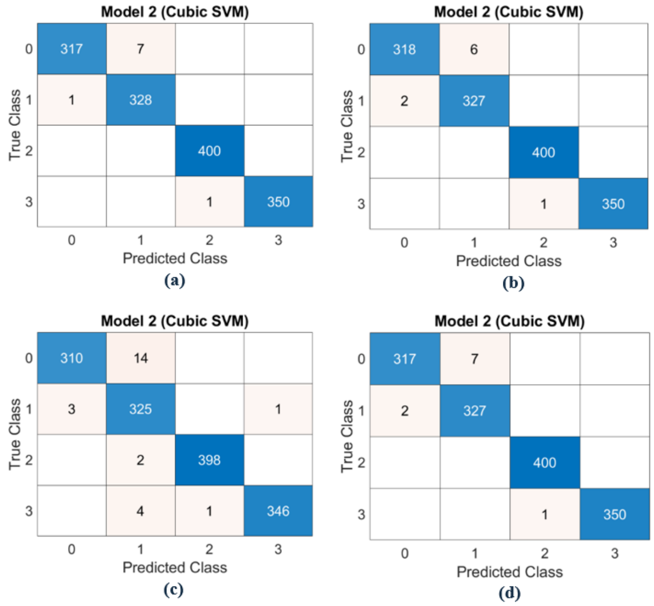

Figure 10 Confusion matrices obtained after feature fusion; a) GC & LBP, b) GC & ORJ, c) LBP & ORJ, d) GC & ORJ & LBP.

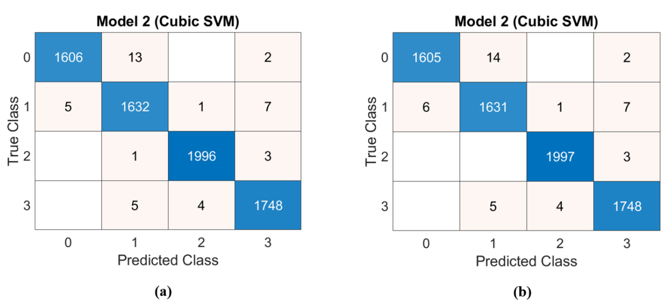

Figure 11 Confusion matrices obtained after cross-validation; a) GC & LBP, b) GC & ORJ

CONCLUSION

In this study, the model's effectiveness was enhanced by using the Grad-CAM and LBP techniques in the preprocessing steps. The regions of interest in the model's decision-making process were identified using the Grad-CAM-based visual interpretability approach. This step made the anatomical structures associated with the tumor more prominent while reducing the impact of the background and clinically insignificant regions. Thus, the deep learning model was directed to more meaningful regions in terms of classification, and the clinical consistency of the learned representations was increased. In addition, tissue features were extracted using the LBP method. LBP is an effective method, particularly for capturing microstructural differences between tumorous and healthy tissue. Thanks to this approach, local tissue variations specific to tumor regions were quantitatively represented and used as a complement to deep learning-based features. These different representations obtained during the preprocessing stage were converted into high-level features via the DeiT3 model. The combination of feature sets obtained from different preprocessing methods (feature fusion) created a richer and more discriminative feature structure compared to representations obtained from a single source. This directly contributed to strengthening the distinction between classes.

The preprocessing steps applied in the study played a critical role in reducing noise, highlighting tumor-related tissue and structural information, and making the model's learning process more targeted. The experimental results obtained demonstrate that these preprocessing strategies improve classification performance and that the proposed approach offers an effective and reliable solution for the brain tumor detection problem. The main reason for choosing the DeiT3 model in this study is that it can offer high generalization performance even without a very large dataset and provides stability during the training process. Although ViT approaches have the ability to effectively learn visual features, they often require a large amount of labeled data and complex training strategies. DeiT3, on the other hand, significantly alleviates these structural constraints through methods that balance the training process and limit the model's tendency to overfit.

The reason for using SVM in the classification stage is based on the method's ability to create effective decision boundaries in high-dimensional feature structures. Particularly when features obtained from deep learning-based models exhibit non-linear distributions, SVM's margin-based optimization approach offers more stable and discriminative classification performance. In addition, SVM has been preferred because it produces more consistent results by reducing the risk of overfitting under limited data conditions. The proposed approach offers several advantages in the analysis of brain tumor MRI images. Combining high-level features derived from the deep learning-based DeiT3 architecture with textural features extracted using the LBP method has created a complementary representation structure that incorporates both global and local information. This multi-feature representation has contributed to strengthening the distinction between classes and supported the improvement of the obtained classification performance. Furthermore, the use of Grad-CAM has enabled the visualization of the regions on which the model focuses during the decision-making process, allowing for a clearer analysis of tumor-related anatomical structures and increasing the clinical interpretability of the method.

However, the proposed method also has limitations. Its multi-stage structure, which includes preprocessing, feature extraction, feature fusion, and classification steps, increases computational cost and requires a more complex workflow compared to end-to-end learning approaches. Furthermore, the method's performance is sensitive to the quality of the preprocessing steps and the selected features, and parameter choices in these stages can influence the results. Finally, while the results are promising, validating the method's generalizability on larger, multi-center datasets is important for increasing its reliability for clinical applications. The proposed approach can be considered as a decision support tool that can assist the clinical diagnosis process through the automatic analysis of brain tumor MRI images. Especially under heavy clinical workload, it can contribute to accelerating the preliminary evaluation phase, allowing specialist physicians to focus on more complex cases. Furthermore, Grad-CAM-based visualizations enhance the clinical interpretability of the findings by making the model's decision-making process more understandable. In this regard, the proposed method can be used in clinical practice as a tool that supports the diagnostic process rather than replacing the physician's final decision.

The experimental results show that the proposed hybrid approach offers an effective method for the accurate and rapid diagnosis of brain tumor disease. The results of the study reveal that new-generation ViT models such as Deit3 are much more effective than traditional diagnostic methods. During the analyses in

Table 6 Comparison of studies using the same dataset

Research	Number of Images	Class	Model/Method	Değer (%)
(Gómez-Guzmán *et al.* 2023)	7.023	4	ResNet50, InceptionV3, InceptionResNetV2, Xception, MobileNetV2 and EfficientNetB0	%97.1
(Bayaral *et al.* 2025)	7.023	4	VGG16, VGG19, ResNet50, MobileNetV2 and SVM / XGBoost	%97.8
This research	7.023	4	DeiT3 model, Grad-CAM and LBP image processing techniques, Cubic SVM method	%99.5

this study, Grad-CAM and LBP image processing techniques were applied to the original dataset. The feature fusion resulted in the best combination set not being the original dataset, but rather the combination of Grad-CAM and LBP, demonstrating that the image processing techniques used provided a meaningful contribution. The use of the DeiT3 model in model training required powerful hardware during training. Therefore, it may be difficult to apply the proposed approach on low-performance systems. A comparison of studies conducted with the same dataset is provided in Table 6.

Ethical standard

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

Availability of data and material

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of interest

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

LITERATURE CITED

Almohamade, A., S. Kammoun, and F. Alsolami, 2025 Local binary pattern–cycle generative adversarial network transfer: Transforming image style from day to night. Journal of Imaging **11**: 108.

Aslan, E., 2024 Classification of brain tumors from mr images using the lstm-esa hybrid model. Adıyaman University Journal of Engineering Sciences **11**: 63–81.

Aslan, M., 2022 Deep learning-based automatic brain tumor detection. Fırat University Journal of Engineering Sciences **34**: 399–407.

Attallah, O., 2025 Multi-domain feature incorporation of lightweight convolutional neural networks and handcrafted features for lung and colon cancer diagnosis. Technologies (Basel) **13**: 173.

Aydemir, E., 2022 Speaker recognition and speaker verification by comparing mfcc and lbp methods. Turkish Informatics Foundation Journal of Computer Science and Engineering **15**: 104–109.

Bayaral, S., E. Gül, and D. Avcı, 2025 Classification of brain tumors using artificial intelligence. International Journal of Innovative Engineering Applications **9**: 8–22.

Bhuvaji, S., 2025 Brain tumor classification (mri). Kaggle dataset, Accessed: 10/15/2025.

Çelik, Ö. and B. C. Koç, 2021 Classification of turkish news texts using tf-idf, word2vec, and fasttext vector model methods. Deu Faculty of Engineering Science and Engineering **23**: 121–127.

Cheng, J., 2024 Brain tumor dataset. Figshare dataset, Accessed: 10/15/2025.

Das, D., C. Sarkar, and B. Das, 2025 Real-time detection of meningiomas by image segmentation: A very deep transfer learning convolutional neural network approach. Tomography **11**: 50.

Demirel, C. and E. Soylu, 2024 Comparison of transfer-based deep learning algorithms for tumor detection in mri data. Black Sea Journal of Science **14**: 1322–1339.

Devi, A. G. *et al.*, 2023 An improved chi2 feature selection based on a two-stage prediction of comorbid cancer patient survivability. Revue d'Intelligence Artificielle **37**: 83–92.

Ennab, M. and H. Mcheick, 2025 Advancing ai interpretability in medical imaging: A comparative analysis of pixel-level interpretability and grad-cam models. Machine Learning and Knowledge Extraction **7**: 12.

Erçelik, Ç. and K. Hanbay, 2023a Classification of brain tumors using gaussian filtering and the resnet50 model. Computer Science .

Erçelik, Ç. and K. Hanbay, 2023b Effects of histogram equalization method on some deep learning models in brain tumor classification. Computer Science .

Gómez-Guzmán, M. A. *et al.*, 2023 Classifying brain tumors on magnetic resonance imaging by using convolutional neural networks. Electronics (Basel) **12**: 955.

Gür, Y. E., M. Toğaçar, and B. Solak, 2025 Integration of cnn models and machine learning methods in credit score classification: 2d image transformation and feature extraction. Computational Economics **65**: 2991–3035.

Haldar, B., H. Joardar, A. K. Mondal, N. H. Alrasheedi, R. Khan, *et al.*, 2025 Machine-learning-driven analysis of wear loss and frictional behavior in magnesium hybrid composites. Crystals (Basel) **15**: 452.

Hamada, A., 2020 Br35h: Brain tumor detection. Kaggle dataset, Accessed: 10/15/2025.

Kuştaşı, E. and M. Yağanoğlu, 2024 Classification of variable stars using deep learning and transfer learning methods. Batman University Journal of Life Sciences **14**: 81–97.

Livieris, I. E., E. Pintelas, N. Kiriakidou, and P. Pintelas, 2023 Explainable image similarity: Integrating siamese networks and grad-cam. Journal of Imaging **9**: 224.

Nickparvar, M., 2021 Brain tumor mri dataset. Kaggle dataset, Accessed: 10/15/2025.

Rahman, S., S. N. F. Mursal, M. A. Latif, Z. Mushtaq, M. Irfan,

et al., 2023 Enhancing network intrusion detection using effective stacking of ensemble classifiers with multi-pronged feature selection technique. In *2023 2nd International Conference on Emerging Trends in Electrical, Control, and Telecommunication Engineering (ETECTE)*, pp. 1–6, IEEE.
Senjoba, L., H. Ikeda, H. Toriya, T. Adachi, and Y. Kawamura, 2024 Enhancing interpretability in drill bit wear analysis through explainable artificial intelligence: A grad-cam approach. Applied Sciences **14**: 3621.
Serttaş, S. and E. Deniz, 2023 Disease detection in bean leaves using deep learning. Communications Faculty of Sciences University of Ankara Series A2-A3 Physical Sciences and Engineering **65**: 115–129.
Sevinc, A., M. Ucan, and B. Kaya, 2025 A distillation approach to transformer-based medical image classification with limited data. Diagnostics **15**: 929.
Taşdemir, B. and N. Barışçı, 2024 Brain tumor segmentation with deep learning. Journal of Information Technologies **17**: 159–174.
Toğaçar, M., 2025 Time-frequency imaging for epilepsy seizure detection: Feature fusion with transformer model. Journal of Engineering Sciences and Research **7**: 93–102.
Tubog, M. V., K. Emsellem, and S. Bouissou, 2025 Detection of agricultural terraces platforms using machine learning from orthophotos and lidar-based digital terrain model: A case study in roya valley of southeast france. Land (Basel) **14**: 962.
Yang, P., R. Zhang, C. Hu, and B. Guo, 2025 Gmdnet: Grouped encoder-mixer-decoder architecture based on the role of modalities for brain tumor mri image segmentation. Electronics (Basel) **14**: 1658.
Yenikaya, M. A. and O. Oktaysoy, 2023 The use of artificial intelligence applications in the healthcare sector: Preliminary diagnosis using deep learning methods. Sakarya University Business Institute Journal **5**: 127–131.
Yılmaz, S., 2023 Design of a decision support system based on the yolov7 algorithm for brain tumor diagnosis. Kocaeli University Journal of Science **6**: 47–56.

How to cite this article: Kahraman, S., and Toğaçar, M. Classification of Brain MRI Images using Deep Learning: The DeiT3 Model and the Use of Feature Fusion Methods. *Computers and Electronics in Medicine*, 3(1), 77-85, 2026.