Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 22 No. 8 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v22i08.61385

PAPER

Brain Tumor Classification Using a Hybrid Learning Strategy Integrating GAN-Augmented Neural Networks

Iliass Zine-dine (✉), Jamal Riffi, Khalid El Fazazy, Ismail El Batteoui, Mohamed Adnane Mahraz, Hamid Tairi

Sidi Mohamed Ben Abdellah University (U.S.M.B.A), Fès, Morocco

iliass.zinedine@usmba.ac.ma

ABSTRACT

Brain tumor (BT) classification plays a vital role in computer-aided diagnosis, as early detection directly influences treatment planning and patient survival. Advances in artificial intelligence, particularly in machine learning (ML) and deep learning (DL), have greatly improved automated tumor identification from magnetic resonance imaging (MRI), reducing reliance on manual assessment. This study proposes a hybrid framework that combines generative adversarial networks (GANs), convolutional neural networks (CNNs), and Vision Transformers (ViTs) for accurate tumor classification. GANs are used to generate realistic synthetic MRI images, addressing data scarcity and increasing dataset diversity. CNNs extract discriminative deep features, accelerating training while limiting overfitting. These features are then processed by a ViT model, capable of capturing complex spatial relationships within medical images. Experimental results show that the proposed ViT–CNN–GAN approach achieves 97.36% accuracy, outperforming conventional methods. Overall, this framework demonstrates the potential of advanced DL models to strengthen MRI-based tumor diagnosis and support more reliable clinical decision-making.

KEYWORDS

brain tumor (BT), generative adversarial networks (GANs), VGG-16, ResNet50, Vision Transformers (ViTs)

1 INTRODUCTION

The brain is commonly recognized as the most intricate organ of the human body and a core element of the nervous system. It is composed of highly specialized tissues and neural cells responsible for controlling vital functions such as breathing, movement coordination, and sensory processing. Given its essential role, the timely identification of brain tumors remains a major priority in both clinical care and biomedical research [1].

Zine-dine, I., Riffi, J., El Fazazy, K., El Batteoui, I., Mahraz, M. A., Tairi, H. (2026). Brain Tumor Classification Using a Hybrid Learning Strategy Integrating GAN-Augmented Neural Networks. *International Journal of Online and Biomedical Engineering (iJOE)*, 22(8), pp. 4–24. https://doi.org/10.3991/ijoe.v22i08.61385

Article submitted 2026-03-05. Revision uploaded 2026-04-23. Final acceptance 2026-04-23.

Cancer is still one of the top diseases causing death around the world and a significant factor affecting life expectancy across nations. In 2020, brain tumors accounted for 1.6% of all tumors in the human body and 2.5% of tumor-related deaths. Given the continually rising frequency of brain cancers worldwide, the World Health Organization (WHO) has highlighted their screening, diagnosis, and treatment as a critical public health priority [2]. The abnormal development of brain tissue is the hallmark of a brain tumor. Unlike many other tumor types, brain tumors typically expand locally and rarely metastasize beyond the brain. This localized growth makes classification particularly challenging, as brain tumors can develop in different parts of the nervous system [3]. Our study's main goal is to enhance brain tumor (BT) classification using magnetic resonance imaging (MRI) images and neural network architectures.

Numerous techniques have been created to automate the prediction of tumors in medical images, many of which leverage machine learning (ML) techniques. These methods generally consist of two primary steps: feature extraction (FE) and the application of a classification model. Sharma et al. [4] highlighted the vital importance of human life in society, particularly within the medical field, emphasizing the necessity of timely intervention, identification, and categorization of brain tumors. Addressing the challenge of classification often involves FE and classification through ML algorithms. In this context, Saba et al. [5] presented a novel method for BT early detection and classification, such as ischemic strokes and gliomas, utilizing ML techniques combined with FE algorithms.

Although these approaches have shown positive results, they face limitations, particularly when working with large datasets. In contrast, deep learning (DL)-based techniques have proven highly effective in managing extensive data. Heba Mohsen et al. [6] proposed an innovative method employing a deep neural network (DNN) for categorizing BT images. They segmented images employing fuzzy C-means, using Discrete Wavelet Transformations (DWT) to extract features, and applying principal component analysis (PCA) to reduce dimensionality. Similarly, a method using convolutional neural networks (CNNs) was presented by Milica M. Badža et al. [7] to identify and categorize brain malignancies, namely meningiomas, gliomas, and pituitary tumors.

Transfer learning is a method that allows information gathered from a model that's already been trained to be used with a fresh dataset. Arshia et al. [8] utilized and analyzed various CNN models, such as AlexNet, GoogleNet, and VGG-16, to enable the identification and classification of brain cancers based on MRI images.

This study introduces a creative approach for BT classification by integrating GANs, CNNs, and Vision Transformers (ViTs). Our approach addresses critical difficulties in using GANs for medical image analysis for synthetic image generation to enhance data diversity, CNNs for effective FE and optimized model training, and ViTs for capturing intricate spatial dependencies. The proposed methodology ensures improved classification accuracy, increased robustness against variations, and a promising possibility of early BT identification, thereby contributing to advancements in computer-aided diagnosis.

The rest of this paper is structured as follows: Section 2 presents an overview of related work, highlighting key advancements and challenges. Section 3 outlines the methodologies and techniques utilized in our proposed approach. In Section 4, we introduce and analyze the experimental solutions attained. Lastly, Section 5 provides a conclusion summarizing our findings and suggests directions for subsequent research that are outlined.

2 RELATED WORK

This unit aims to provide a complete analysis of previous studies focused on the application of FE techniques, ML, DL architectures, auto-encoders, and transformer models in identifying and categorizing brain cancers. Numerous studies in this area emphasize early disease detection through non-invasive CAD techniques, effectively reducing reliance on surgical or invasive procedures. Sekhar et al. [9] developed a fully automated CAD system designed to improve the precision of early BT detection. Their methodology incorporates an automatic feature selection mechanism combined with a hybrid approach utilizing support vector machine (SVM) and K-nearest neighbor (KNN) classifiers. This framework significantly advances the field by enhancing both the accuracy and efficiency of BT classification. Several image-processing architectures have been explored for disease detection and classification. Our study focuses on the earlier prediction of brain tumors, nearly aligning with the referenced work. These investigations operate in the similar field of medical imaging, employing similar computer vision techniques. However, our proposed approach demonstrated superior performance, surpassing numerous modern models of brain tumor classification accuracy.

Several deep CNNs have been pre-trained and are actively utilized for extracting high-level features from magnetic resonance images. These networks are essential to capturing intricate patterns and structural details within medical imaging, improving the precision of disease identification and classification. Amin et al. [10] developed a methodology for segmenting and classifying brain tumors utilizing DNN and MRI. DL has shown itself to be a successful strategy for solving classification challenges. Sultan et al. [11] presented a DL technique that employs CNNs to detect and classify several kinds of brain tumors, which include Meningioma, Glioma, and Pituitary tumors. In a similar vein, Anaraki et al. [12] proposed a new technique that combines CNNs with a genetic algorithm to categorize brain tumors using MRI. In another study, Hüseyin and Engin [13] presented an innovative technique that makes use of CNNs' FE capabilities to categorize brain and liver cancers, the signal processing potential of DWT, and the long-term memory feature for signal classification. Additionally, Parnian Afshar et al. [14] developed a technique that divides brain tumors into three groups: Gliomas, Pituitaries, and Meningiomas, employing CNNs within the CapsuleNet architecture and applying MRI images. These approaches are widely used for quick recognition and prediction of brain tumor–related outcomes. This study conducts a comparative analysis to determine the approach that works well for the given assignment and to present the corresponding findings. While these approaches demonstrate high accuracy, their effectiveness often diminishes when applied to large-scale datasets, posing challenges in handling extensive medical image databases. Javeria et al. [15] employed a FE approach that integrated multiple techniques, including Gabor wavelet features (GWF), histograms of oriented gradients (HOG), local binary patterns (LBP), and segmentation-based fractal texture analysis (SFTA). By combining these methods, they generated a hybrid feature vector designed to enhance texture representation and improve classification performance. In the field of computer vision, a promising strategy for enhancing classification performance involves integrating multiple techniques, such as merging and fusing feature vectors obtained from different extraction methods. This approach leverages complementary information from diverse features, leading to more robust and accurate classification results. Mostafa et al. [16] utilized a fusion-based strategy that integrated texture and morphometric characteristics to investigate a possible biomarker for diagnosis for brain tumors. Their study aspired to enhance category accuracy by leveraging the complementary strengths of these features, ultimately improving the reliability of tumor identification. An alternative method that emphasizes capturing direct correlations between images could offer enhanced effectiveness for brain image analysis when compared to traditional convolutional neural networks. Hossain et al. [17] leveraged ViTs as a classification framework and introduced a transfer learning-based multiclass classification model to enhance classification accuracy. Within this scope, convolutional autoencoders serve as a widely adopted neural network architecture for image processing tasks, primarily focusing on capturing spatial dependencies within input data to support diverse image analysis applications. Nayak et al. [18] conducted an inventive deep autoencoder-based exploratory analysis of BT data. These autoencoders were used to create a condensed version of the input data, which was then rebuilt to resemble the initial data. This approach facilitated both the prediction and the disease's classification by leveraging the representations of the learned data. Anaya-Isaza et al. [19] proposed a different strategy utilizing transformers for BT classification, using self-attention processes to record complex spatial features and enhance diagnostic accuracy. Afshar et al. [20] integrated advanced Capsule Networks (CapsNets) into their framework, allowing the model to process contextual information from surrounding brain tissues while maintaining a strong focus on tumor localization. By incorporating tumor boundary details as supplementary inputs within the pipeline, their approach enhanced CapsNet's ability to capture spatial hierarchies and improve classification accuracy.

The previously discussed FE and classification techniques face several challenges, including reliance on extensive datasets, high computational costs, and difficulties in interpretability. To overcome these limitations, ongoing advancements in computer vision are needed to develop models that are more efficient, require less training data, and provide greater transparency in decision-making. Enhancing FE techniques with improved architectures can lead to better generalization capabilities, reducing dependency on large annotated datasets while maintaining high accuracy in tumor classification. Exploring innovative approaches that surpass these constraints is essential for improving predictive performance in medical imaging.

In medical diagnostics, real-time decision-making is critical, particularly in urgent scenarios where timely and accurate assessments can significantly impact patient outcomes. Evaluating the inference of a model's speed and computational efficiency is important to ensure its practical deployment in clinical environments. Achieving a balance between rapid processing and high diagnostic accuracy is necessary for developing reliable automated systems that assist healthcare professionals. Therefore, optimizing classification models to operate efficiently while maintaining precision remains a crucial objective in advancing methods for identifying and classifying brain tumors [21].

The essential contributions of this study may be summarized in the following order: Standard image processing methods were used throughout the preprocessing stage to improve data Integrity and optimize subsequent analysis. FE was conducted using two DL architectures, VGG-16 and ResNet50, to capture pertinent patterns, accelerate training, reduce overfitting, and improve classification reliability. The extracted features were then fused into a unified vector and fed into a ViT for final classification. To ensure the efficacy of the proposed approach, a comprehensive evaluation was performed at each stage to refine the methodology. The findings demonstrate the capacity of this framework for early BT diagnosis using cutting-edge computer vision and DL methods for medical applications.

3 COMPUTATIONAL FRAMEWORK AND METHODS

This section presents the overall structure of our suggested methodology. Figure 1 offers a detailed breakdown of the architectural parts intended for BT classification. The proposed strategy is structured into three main phases: data preparation, identifying characteristics using DL (GAN, VGG-16, and ResNet50), and classification using ViT, each playing a pivotal role in the workflow. Initially, we describe the dataset employed in this study (Section 3.1). Before moving on to the next stage (Section 3.2), the MRI pictures go through preprocessing techniques to improve quality. In the next phase, a GAN architecture is utilized to produce synthetic yet realistic images, while CNN-based architectures are leveraged to extract features, speed up model training, and reduce overfitting (Section 3.3). Lastly, a RandomForestRegressor is applied to assess the GAN-generated images and retain the most relevant ones for classification with ViTs (Section 3.4). The following subsections provide an in-depth discussion of these fundamental components:

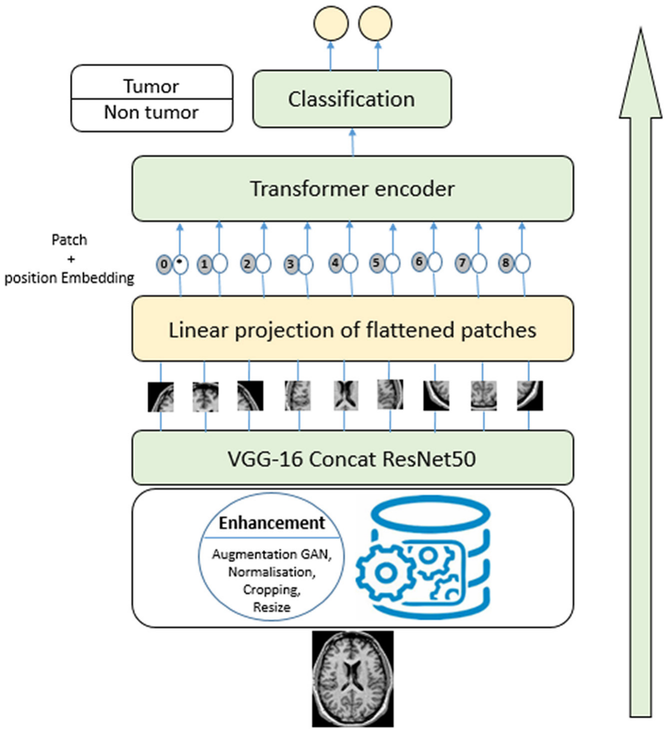

Fig. 1. GAN–CNN–ViT-based architecture for BT classification

3.1 Overview of the data

In this study, we utilized brain MRI images from the freely accessible Kaggle MRI dataset for BT [22]. This dataset comprises 253 MRI scans categorized into two classes: 155 images labeled as 'yes' for tumorous cases and 98 images labeled as 'no' for non-tumorous cases. These MRI scans serve as the foundational data for our study, enabling the exploration of distinct imaging characteristics associated with brain tumors. To enhance the dataset, improve classification performance, and prevent overfitting, we employed a Generative Adversarial Network to generate additional synthetic MRI images, increasing the dataset tenfold (from 253 to 2530 images) in practice. The extracted features from both real and synthetic images were then processed using CNN to capture meaningful patterns. Finally, a Vision Transformer model was utilized for classification, leveraging its ability to analyze spatial dependencies and improve diagnostic accuracy. This hybrid approach integrates the strengths of GANs and ViTs to achieve a robust and efficient classification framework.

To ensure a robust and reliable evaluation of the proposed model, a 5-fold cross-validation strategy was adopted during the dataset partitioning process. Specifically, the dataset was divided into five equally sized subsets, where in each iteration, 80% of the data were used for training and 20% for validation. The final performance was reported as the average of the results obtained across all folds, thereby reducing evaluation bias and improving the generalization capability of the model. The distribution of the dataset across different categories is presented in Table 1.

Table 1. Dataset partitioning using 5-fold cross-validation for the GAN-augmented dataset

Step	Number of Images			Percentage of Images [%]		
	Tumorous	Non-Tumorous	Total	Tumorous	Non-Tumorous	Total
Train	1518	506	2024	75	25	80
Validation	380	126	506	75	25	20

3.2 Data preparation

The implementation of preprocessing techniques to refine datasets and extract meaningful information for classifier input has gained significant traction in the domains of computer vision and image processing [23]. These methods enhance data quality, improve feature representation, and ultimately contribute to more accurate predictive outcomes. In this study, we employ widely recognized image-processing techniques, with the essential strategies utilized in this phase detailed in the following sections.

- **Data Crop**

  A significant portion of the brain MRI images in our dataset contains unnecessary background regions, which can negatively impact classification accuracy. To address this issue, it is essential to crop the images to retain only the most relevant information [24]. In this study, we use a cropping technique that identifies prominent feature points and extracts a specific region of interest based on these boundaries. There are five essential phases in the process: We load the actual MRI pictures first. Next, we use thresholding to generate binary masks, which help in segmenting the area of interest. In the third step, we perform dilation and erosion operations to minimize noise and enhance the quality of segmentation. Then, we identify the four extreme points—top, bottom, left, and right—based on the largest contour of the thresholder image. Finally, we crop the image according to these extreme points to isolate the brain region, followed by bicubic interpolation to standardize the image size. Figure 2 provides a visual representation of the cropping procedure applied to our dataset.

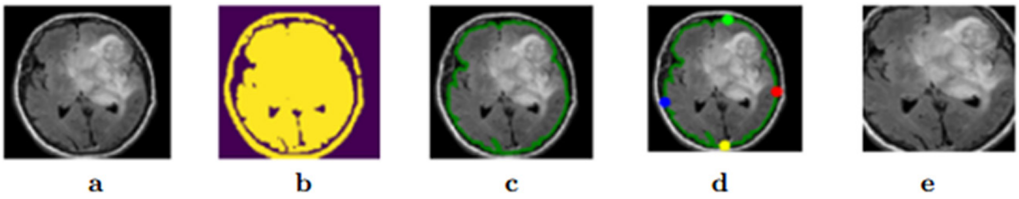

Fig. 2. Sequential steps in the image cropping process: (a) original MRI image; (b) image after binary thresholding; (c) detected external contour shown in green; (d) identified extreme boundary points (R, G, B, Y); (e) final region of interest after cropping

- **Generative Adversarial Networks (GANs)**

Deep learning models, notably convolutional neural networks, require extensive datasets to achieve robust training and prevent overfitting. Conventional data augmentation methods primarily utilize geometric transformations, including image rotation, scaling, image resizing, controlled noise addition, spatial translation, and horizontal/vertical flipping [25]. However, these techniques may not be entirely effective for medical imaging datasets due to their inherent complexity and domain-specific constraints. An emerging approach for augmenting medical image datasets is based on GANs, a DL framework that generates realistic images from random noise. This technique expands dataset size, thereby improving model generalization and performance [26].

A GAN is made up of the discriminator and generator neural networks, which operate in an opposing training framework. The generator is responsible for creating synthetic data samples that closely resemble real dataset instances, whereas the discriminator makes an attempt to distinguish between created and authentic data [27]. This competitive learning paradigm enhances learning efficiency, as the generator progressively improves its ability to produce highly realistic examples to trick the discriminator. Concurrently, the discriminator refines its capacity to distinguish real images from artificial ones [28]. Through this continuous feedback loop, GANs enhance data augmentation by generating high-quality synthetic images that closely mimic real medical scans.

Despite their potential, GANs exhibit several challenges that affect their practical implementation. A significant limitation is their convergence instability, as there is no universally established criterion to determine when the generator has reached optimal performance [29]. Standard loss functions fail to provide a reliable indication of GAN training completion, making it difficult to assess whether high-quality synthetic images have been achieved. Additionally, traditional GANs rely on the Jensen-Shannon divergence for measuring distribution similarity, which may not always be optimal for training stability. To address these issues, researchers have explored alternative strategies, such as modifying the loss function and incorporating the Earth Mover distance metric to enhance convergence. In this study, to address the limited dataset size, improve classification performance, and prevent overfitting, a standard GAN model was employed for data augmentation, increasing the dataset tenfold (from 253 to 2530 images) in practice, while being trained over 10 epochs to ensure convergence stability and assess generator and discriminator losses. Figure 3 presents an overview of the operation of the proposed GAN model.

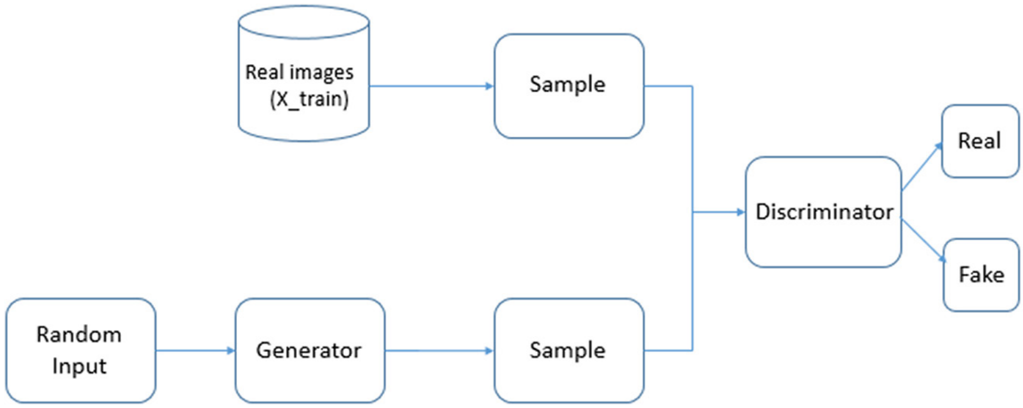

Fig. 3. The main processing stages of the GAN architecture

3.3 Deep FE

A specific class of DL models, known as convolutional neural networks, is designed to handle tasks related to image processing and pattern recognition. Their architectures efficiently extract hierarchical features and spatial relationships from images through the use of convolutional, pooling, and fully connected layers. This capability makes CNNs particularly well suited for applications such as object detection, image classification, and medical imaging–based analysis [30].

One of the main advantages of CNNs is their ability to automatically extract relevant features from raw images without requiring manual feature engineering. This FE process enhances model efficiency by reducing input dimensionality while preserving critical information. Additionally, CNNs incorporate regularization strategies like dropout and batch normalization in order to reduce overfitting and enhance generalization using new data [31].

In this study, we used VGG-16 and ResNet50 for deep FE, aiming to enhance model training efficiency and minimize overfitting. These models' convolutional, pooling, and normalizing layers were utilized while omitting the fully connected classification layers. The extracted features were then processed in the Vision Transformer model for classification. The subsequent sections provide a detailed overview of the VGG-16 and ResNet50 architectures and their contributions to our approach.

VGG-16: A large collection of images in several categories has been used to pretrain the deep convolutional neural network model VGG-16. It is specifically designed for image classification tasks and belongs to the VGGNet (Visual Geometry Group Network) family. This model is characterized by its deep architecture and consistent layer design. The name "VGG-16" originates from its structure, which consists of 16 trainable weight layers. The design is useful for FE in a variety of image-processing applications since it consists of thirteen convolutional layers, followed by fully linked and output layers. [32].

In this work, deep feature representations are obtained by retaining only the convolutional, pooling, and normalization components of the VGG-16 architecture, while the final fully connected classification layers are excluded. The convolutional blocks rely on compact 3×3 kernels applied with a unit stride, whereas the pooling operations use a stride of 2 to progressively reduce the spatial resolution of the feature maps. In the final stage, the retrieved features were aggregated to create a dimension-based feature vector (1,1,1000). This vector was then concatenated with the feature representation obtained from ResNet50, which maintains the same dimensional structure, ensuring a comprehensive feature set for subsequent classification. Figure 4 presents the structured pipeline illustrating the successive transformations and processing steps of the VGG-16 model applied to the dataset.

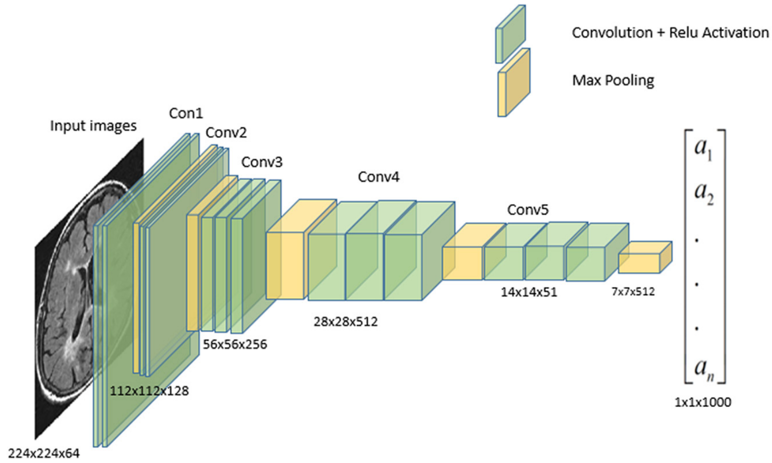

Fig. 4. VGG-16 processing pipeline

ResNet50 is a deep CNN designed for computer vision tasks, where residual functions with regard to the input layer are taught to its weighted layers. The ResNet architecture was introduced to address the issues with disappearing and expanding gradients that occur in extremely deep networks, allowing for improved training stability and performance [33]. The ResNet50 model consists of 50 layers structured into different stages, the architecture is structured around an input stage, followed by an initial convolutional module, four successive levels built from residual units, a global average pooling operation, and a final dense layer. For this study, the model leverages the entire network depth, apart from the terminal dense layer dedicated to classification, which was omitted to focus on extracting deep features. The extracted feature maps were then aggregated into a vector of dimension (1,1,1000), ensuring compatibility for concatenation with the corresponding feature vector obtained from VGG-16. Figure 5 depicts the systematic processing pipeline detailing the sequential transformations associated with the ResNet50 model when applied to the dataset.

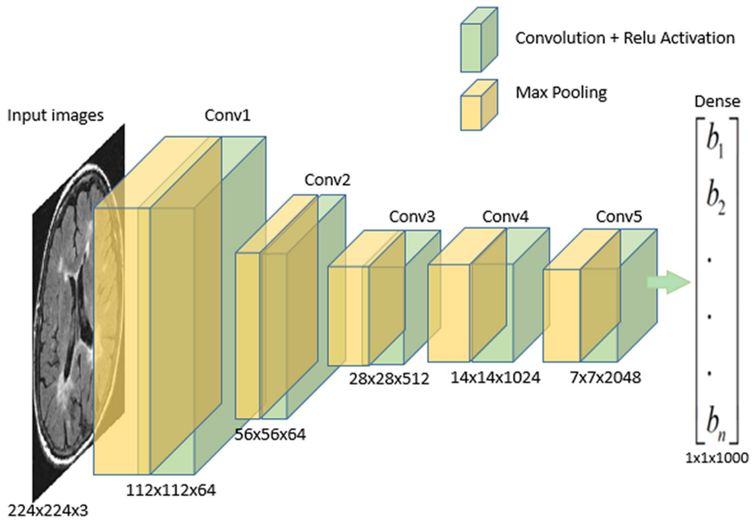

Fig. 5. ResNet50 processing pipeline

The concatenation of feature vectors in this study enhances model performance by integrating essential feature representations. This approach enhances feature representation, thereby enabling a more comprehensive understanding of the data. This approach improves the model's ability to generalize effectively, producing categorization results that are more trustworthy. In addition, by compacting the feature space, the concatenation strategy reduces the risk of overfitting and improves interpretability, which ultimately strengthens the overall BT classification pipeline.

Hyperparameters: All experiments were implemented using a TensorFlow 2.6.4 and Keras 2.6.0 implementation. The VGG-16 and ResNet-50 models rely on fixed architectural hyperparameters, including kernel size, padding, and stride, which remain unchanged during training. In contrast, the trainable parameters of the networks, namely weights and biases, are optimized iteratively. Stochastic Gradient Descent was employed as the optimization algorithm, with the learning rate set to 0.001 to ensure stable convergence and efficient learning. Model performance during training was evaluated using the cross-entropy loss function, which quantifies the divergence between predicted class probabilities and reference labels. The network was trained for 40 epochs with a batch size of 32 samples. These hyperparameter settings were selected to provide a suitable trade-off between computational efficiency and predictive accuracy while promoting stable convergence and improved generalization.

Discussion: The proposed framework leverages the complementary strengths of two convolutional architectures, namely VGG-16 and ResNet50, to enhance both predictive capability and computational practicality. Owing to its regular and sequential design, VGG-16 is particularly well suited for capturing progressively refined feature representations, which are essential for accurate image classification. In contrast, ResNet50 introduces residual learning mechanisms that enable efficient gradient propagation across deeper layers. This design choice alleviates optimization difficulties commonly encountered in deep networks, thereby supporting stable training while maintaining strong classification performance.

The joint utilization of VGG-16 and ResNet-50 within the framework leverages their complementary properties, combining expressive hierarchical feature learning with improved optimization dynamics through residual connections. This synergy not only enhances model training but also aids in reducing overfitting, ultimately improving classification accuracy and the overall robustness of our approach. Furthermore, integrating these feature representations within a ViT model enables more refined decision-making, leveraging both local and global image features for precise tumor classification.

3.4 Vision transformer–based classification

The Vision Transformer represents a neural network model specifically designed for image analysis [34]. Unlike CNNs, which rely on convolutional and pooling operations, ViT utilizes a self-attention mechanism, a concept inspired by human cognition that selectively focuses on the most relevant aspects of an image while disregarding less significant details [35]. This approach allows ViT to extract relevant information and long-range relationships than typical CNN-based architectures.

Within a ViT framework, an image is segmented into fixed-size patches, and the resulting representations are first reshaped into a one-dimensional format and organized as a sequence of tokens. This token sequence is then forwarded to a stack of Transformer blocks, where self-attention mechanisms and non-linear operations are applied to refine the feature representations [36]. As a result, ViT generates a feature-rich vector representation for each patch, which is later aggregated to facilitate jobs like classifying images. Figure 6 provides a visual representation of how our ViT model operates.

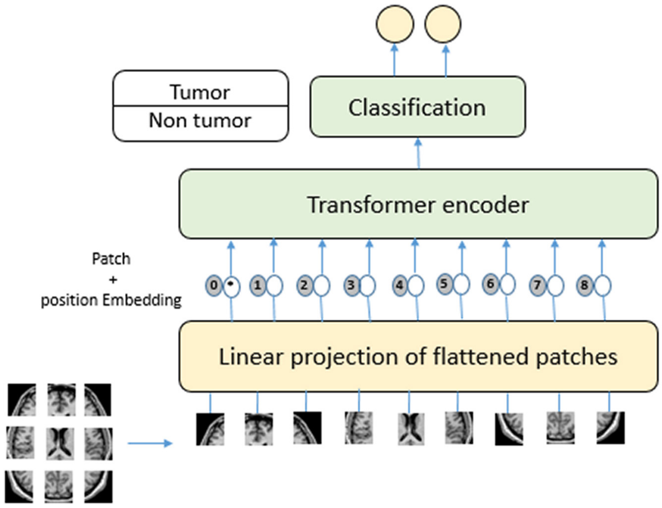

Fig. 6. Sequential processing stages in the ViT model

Discussion: In our strategy, the incorporation regarding the ViT technique provides notable benefits compared to conventional convolutional architectures. Unlike CNNs, which primarily focus on local FE through convolutional filters, ViT relies on self-attention operations to establish global dependencies between image patches, enabling comprehensive contextual modeling. This characteristic allows ViT to achieve an additional thorough comprehension of visual data, making it especially effective in complex classification tasks. Additionally, ViT's ability to process images as sequences of patches eliminates the need for convolutional inductive biases, enabling greater adaptability to diverse image variations. By integrating ViT into our structure, we improve feature representation capabilities and improve classification performance, demonstrating its effectiveness in the given use.

3.5 Performance analysis and discussion

Our method is structured around a carefully created three-stage approach with the goal of optimizing the efficiency and accuracy of our study framework. The first phase focuses on dataset preprocessing and augmentation, where a GAN is employed to enhance data diversity and realism. This stage not only cleans up the raw data but also strengthens feature representation, leading to improved generalization. In the second phase, we adopt a novel FE strategy by combining the outputs of two well-established neural network models, VGG16 and ResNet50. This feature obtaining leverages the complementary strengths of both architectures, facilitating accelerated training in the Vision Transformer model while mitigating overfitting. By integrating these deep feature representations, our approach ensures a more comprehensive understanding of spatial dependencies within the data, ultimately enhancing classification performance. The final phase involves utilizing a ViT model specifically designed for classification, capitalizing on its capacity to extract spatial relationships and identify complex patterns from images. This structured methodology enables precise and nuanced classification, reinforcing the robustness of our approach. Additionally, given the critical role of medical diagnostics performance in real time, particularly in emergency scenarios, we assessed the inference time of the model. Our approach successfully achieves classification results in under a minute, skillfully striking a balance between accuracy and speed to guarantee practical applicability in applied contexts and medical settings.

4 EXPERIMENTAL EVALUATION AND FINDINGS

This section presents the experimental framework and the corresponding results obtained from the proposed GAN-enhanced neural network approach for BT classification. It first summarizes the methodological setup and the implementation environment adopted for model development. The evaluation criteria employed to assess classification performance are then detailed. Finally, a comprehensive comparative analysis of the different experimental configurations investigated in this study is provided, we analyze our results and propose potential avenues for future research.

4.1 Description of the technical execution environment

To carry out this study under optimal technical conditions, a computing environment capable of handling intensive DL operations was carefully selected. The experiments were conducted using Google Colab, which offered access to GPU acceleration and approximately 12 GB of RAM, enabling efficient processing of computationally demanding stages such as training, validation, and performance testing of the proposed architecture.

The implementation of the models was developed using TensorFlow (version 2.6.4) in combination with Keras (version 2.6.0), providing a flexible and stable framework for designing and optimizing deep neural networks. This configuration facilitated the seamless integration of the generative adversarial network, multiple convolutional neural network architectures, and the Vision Transformer within a unified hybrid framework dedicated to BT classification. Overall, the selected setup ensured both computational efficiency and methodological consistency throughout the experimental phase.

4.2 Analytical findings

To rigorously evaluate the predictive capability of the proposed classification framework, a comprehensive set of performance metrics was systematically employed. Specifically, precision, recall, overall accuracy, F1-score, and loss values were computed to capture complementary aspects of model behavior, including class-wise discrimination, balance between sensitivity and specificity, and optimization stability. The joint consideration of these indicators ensures a robust and multidimensional assessment of classification effectiveness, particularly in the context of medical image analysis where class imbalance and misclassification costs are critical concerns.

The subsequent sections provide an in-depth quantitative analysis of these metrics, supported by detailed visualizations of training and validation loss and accuracy trajectories. The examination of these learning curves enables a thorough evaluation of convergence dynamics, generalization capacity, and potential overfitting phenomena, thereby offering deeper insight into the stability and reliability of the proposed approach.

- **Model Effectiveness Evaluation**

  Assessing the effectiveness of an ML model for a specific task necessitates the use of multiple evaluation criteria and analytical techniques. These measures provide a comprehensive understanding of the model's predictive accuracy, generalization capacity, and reliability on previously unseen data. By systematically applying such metrics, it becomes possible not only to quantify performance but also to detect potential limitations, including overfitting or underfitting, thereby guiding further model refinement and optimization [37]. To assess the success of our experiment, we utilized several evaluation indicators tailored to the classification task, including precision, recall, accuracy, and F1-score.

  Precision: the percentage of results that are relevant, and it is defined as:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

  Recall: quantifies the proportion of actual positive instances that are correctly identified by the proposed algorithm. Formally, it is defined as:

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

  Accuracy: measures the overall proportion of correctly classified instances among the total number of samples. It reflects the global predictive correctness of the model across all classes. Mathematically, it is expressed as:

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

  The F1-score: is a composite evaluation metric commonly used in classification tasks to provide a balanced assessment between precision and recall. It represents the harmonic mean of these two indicators, thereby emphasizing the trade-off between false positives and false negatives. The F1-score is mathematically defined as:

$$F_1 = 2 * \frac{P * R}{P + R} \tag{4}$$

  In this context, *TP*, *TN*, *FP*, and *FN* correspond to the counts of correctly identified positive samples, correctly identified negative samples, incorrectly predicted positive instances, and incorrectly predicted negative instances, respectively. While overall accuracy offers a broad measure of predictive correctness, it does not always provide a reliable reflection of model performance, particularly when the class distribution is uneven. Under such circumstances, additional evaluation measures are required to obtain a more balanced and informative assessment of the classifier's behavior.

  The experimental validation was conducted using a publicly available BT dataset obtained from a Kaggle challenge dedicated to tumor classification tasks. This dataset served as the empirical foundation for evaluating the proposed framework. The core objective of the study was to leverage deep feature representations extracted from two established convolutional neural network architectures, namely VGG-16 and ResNet50. These high-level representations were subsequently employed as input embeddings for a Vision Transformer-based classifier, enabling the integration of convolutional feature learning with transformer-based modeling.

To further enhance discriminative capacity, a spatial feature refinement process was applied prior to the final classification stage, aiming to improve representational quality and predictive robustness. The following section presents the primary evaluation criteria, specifically loss and accuracy, which are used to quantitatively analyze the learning behavior and overall effectiveness of the proposed approach.

- **Quantitative Analysis of Predictive Accuracy and Loss Convergence**

Accuracy: reflects the proportion of correctly classified samples relative to the total number of observations, providing a global indication of predictive reliability. It expresses the extent to which the model's outputs align with the true labels. The following figures illustrate the quantitative results obtained from these predictions.

Loss: quantifies the deviation between the predicted outputs and the corresponding ground-truth labels. It represents the aggregated error computed across all samples within the training or validation set, serving as the objective function minimized during optimization [38].

The experimental analysis was conducted using the publicly available BT (Two-Class) dataset from Kaggle. Following a preprocessing stage based on computer vision techniques, deep representations were extracted from VGG-16 and ResNet50 architectures and subsequently provided to a ViT classifier to enhance spatial dependency modeling beyond conventional CNN limitations. Figures 7 and 8 illustrate the corresponding accuracy and loss curves for each evaluated model.

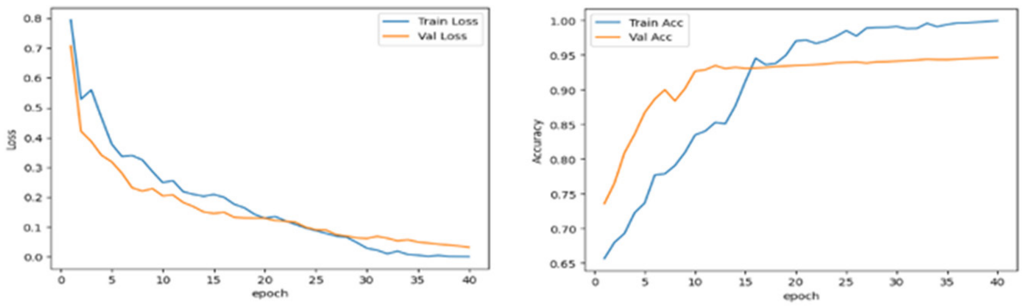

Fig. 7. ResNet50–ViT accuracy and loss curves

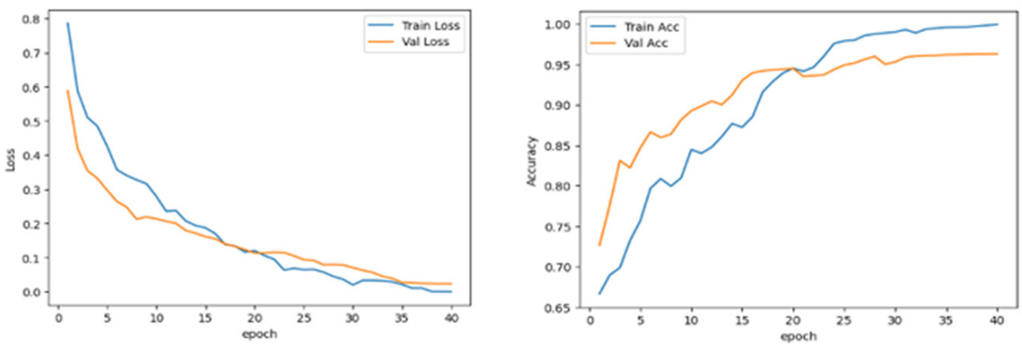

Fig. 8. VGG-16–ViT accuracy and loss curves

In summary, the experimental results indicate that a single CNN-based feature extractor already achieves competitive classification performance relative to comparable studies using the same dataset. To further improve predictive capability, a dual-FE strategy combining VGG-16 and ResNet50 was adopted, where the fused representations were provided to the classifier to enhance robustness and reduce generalization issues. Figure 9 presents the corresponding accuracy and loss trends of the proposed framework.

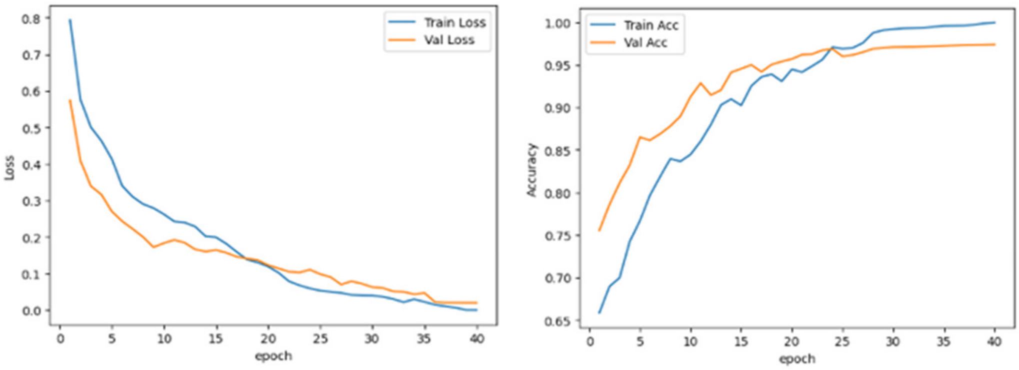

Fig. 9. Accuracy and loss trajectories of the proposed hybrid architecture

A detailed examination of the obtained results confirms the strong effectiveness of the proposed framework, which surpasses several recent state-of-the-art approaches. This improvement can be attributed to the synergistic combination of VGG-16 and ResNet50 feature extractors with a ViT module, enabling enhanced representation learning and improved modeling of long-range spatial dependencies.

4.3 Analytical interpretation of the results

The proposed framework combines advanced image preprocessing procedures with GAN-based data augmentation to enrich the diversity and realism of the training samples. By integrating VGG-16 and ResNet50 for local FE with a Vision Transformer for global dependency modeling, the method achieves a comprehensive and robust representation of image characteristics. Table 2 summarizes the classification performance obtained at each stage of the proposed pipeline.

Table 2. Summary of classification performance across different feature configurations

Feature Strategy	Classifier Applied	Accuracy (%)
Raw dataset (no augmentation)	VGG16	76.18
GAN-enhanced data	AlexNet	79.94
GAN-enhanced data	DenseNet	88.74
GAN-enhanced data	ViT	89.83
GAN+ResNet50 features	ViT	94.63
GAN+VGG16 features	ViT	96.29
GAN + (VGG-16 Concat ResNet50) features Proposed	**ViT**	**97.36**

The proposed hybrid framework exhibits improved computational efficiency while providing complementary advantages over standalone CNN and transformer-based architectures. By jointly modeling fine-grained local patterns and broader contextual dependencies, the approach achieves a more comprehensive feature representation. The table above illustrates the corresponding performance results of the developed model.

The superior performance of the proposed framework becomes particularly apparent when combining a Vision Transformer with concatenated deep representations extracted from VGG-16 and ResNet50 architectures. While convolutional networks are known for their stable optimization behavior and strong discriminative capacity, their receptive field mechanism inherently limits their ability to explicitly model long-range spatial interactions and relative positional relationships between distant features. Consequently, CNN-based representations may insufficiently encode global contextual dependencies across the image.

To address this limitation, the ViT component introduces a self-attention mechanism that enables direct modeling of pairwise feature interactions, allowing the network to dynamically weigh the relevance of different regions. By segmenting the input image into structured patches, the transformer architecture processes both fine-grained local cues and broader contextual information within a unified framework. Experimental observations indicate that this hybrid integration enhances representational richness, improves scalability, and maintains architectural efficiency. Nevertheless, transformer-based models remain sensitive to dataset size and require careful optimization to achieve stable convergence.

The proposed framework integrates advanced deep FE strategies with a ViT-based prediction module, resulting in improved classification performance compared to existing approaches in BT analysis. Furthermore, the computational efficiency of the model was assessed under real-time conditions, achieving an accuracy of 97.36% with an inference time below one minute. These findings highlight both the predictive strength and the practical feasibility of the proposed method for real-world clinical deployment.

Our results demonstrate strong and reliable performance, outperforming several recent approaches, as shown in Table 3. This effectiveness stems from a well-structured pipeline that integrates careful preprocessing, texture FE, and the combined use of deep representations from VGG-16 and ResNet50, further enhanced by a ViT to capture long-range dependencies through self-attention. In addition, the use of computer vision techniques and data enrichment strategies, including GAN-based augmentation, has contributed to improved training efficiency, reduced overfitting, and enhanced the overall robustness of the classification framework.

Table 3. Comparative performance of the proposed method against various CNN-based approaches using the same dataset

Reference	Classification Method	FE Method	Accuracy
[39]	FC Layer	CNN	90%
[40]	VGG-16	Lu-Net	90%
[41]	Stacking	HOG + Wavelet	92%
[42]	ResNet-50	CNN	96%
[43]	AlexNet	CNN	96%
Proposed	**ViT**	**GAN + (VGG-16 Concat ResNet50)**	**97%**

5 CONCLUSION AND PERSPECTIVE

This study introduces a hybrid DL framework for BT classification based on complementary neural network architectures. The proposed methodology is structured into three main stages. The first stage involves dataset preparation, including image normalization, resizing, cropping, and augmentation to ensure data consistency and variability. The second stage focuses on deep FE using VGG-16 and ResNet50 networks to capture discriminative visual patterns. Finally, the extracted representations are provided to a ViT model, which performs the classification task by modeling both local and global contextual dependencies.

The effectiveness of the proposed framework lies in the complementary integration of two deep FE backbones, VGG-16 and ResNet50. The high-level representations generated by these architectures are fused through feature vector concatenation, forming a unified embedding space that enriches the descriptive capacity of the model. This consolidated representation is subsequently processed by a ViT classifier, enabling enhanced predictive performance through joint modeling of local and global dependencies.

The experimental findings demonstrate the effectiveness of the adopted FE strategies, as both VGG-16 and ResNet50 successfully derive high-level representations from MRI images. The fusion of these complementary deep features further enhances the descriptive strength of the learned embedding space. Moreover, the ViT component effectively models spatial dependencies across image regions, contributing to improved classification performance. Overall, this study advances the development of reliable computer-aided diagnostic systems within the domain of digital pathology.

Numerous studies in disease classification have yet to meet the standards expected by medical professionals, often due to approaches that exhibit limited accuracy, high data demands, or computational inefficiency in DL models. Future work will extend this study to additional brain disease datasets, with the goal of developing methods that address these shortcomings while maintaining robust performance and efficiency.

6 REFERENCES

[1] S. Khan, A. Khan, N. Hameed, M. A. Taufiq, and S. Riaz, “Localizing epileptogenic zone from high density EEG data using machine learning,” *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 17, no. 6, pp. 73–86, 2021. https://doi.org/10.3991/ijoe.v17i06.18653

[2] D. N. Louis, “The 2016 World Health Organization classification of tumors of the central nervous system: A summary,” *Acta neuropathologica*, vol. 131, pp. 803–820, 2016. https://doi.org/10.1007/s00401-016-1545-1

[3] L. M. DeAngelis, “Brain tumors,” *New England Journal of Medicine*, vol. 344, no. 2, pp. 114–123, 2001. https://doi.org/10.1056/nejm200101113440207

[4] K. Sharma, A. Kaur, and S. Gujral, “Brain tumor detection based on machine learning algorithms,” *International Journal of Computer Applications*, vol. 103, no. 1, pp. 7–11, 2014. https://doi.org/10.5120/18036-6883

[5] T. Saba, A. Sameh Mohamed, M. El-Affendi, J. Amin, and M. Sharif, “Brain tumor detection using fusion of hand crafted and deep learning features,” *Cognitive Systems Research*, vol. 59, pp. 221–230, 2020. https://doi.org/10.1016/j.cogsys.2019.09.007

[6] H. Mohsen, E.-S. A. El-Dahshan, E.-S. M. El-Horbaty, and A.-B. M. Salem, "Classification using deep learning neural networks for brain tumors," *Future Computing and Informatics Journal*, vol. 3, no. 1, pp. 68–71, 2018. https://doi.org/10.1016/j.fcij.2017.12.001

[7] M. M. Badža and M. Č. Barjaktarović, "Classification of brain tumors from MRI images using a convolutional neural network," *Applied Sciences*, vol. 10, no. 6, 2020. https://doi.org/10.3390/app10061999

[8] A. Rehman, S. Naz, M. I. Razzak, F. Akram, and M. Imran, "A deep learning-based framework for automatic brain tumors classification using transfer learning," *Circuits, Systems, and Signal Processing*, vol. 39, no. 2, pp. 757–775, 2020. https://doi.org/10.1007/s00034-019-01246-3

[9] A. Sekhar, S. Biswas, R. Hazra, A. K. Sunaniya, A. Mukherjee, and L. Yang, "Brain tumor classification using fine-tuned GoogLeNet features and machine learning algorithms: IoMT enabled CAD system," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 3, pp. 983–991, 2022. https://doi.org/10.1109/jbhi.2021.3100758

[10] J. Amin, M. Sharif, M. Yasmin, and S. L. Fernandes, "A distinctive approach in brain tumor detection and classification using MRI," *Pattern Recognition Letters*, vol. 139, pp. 118–127, 2020. https://doi.org/10.1016/j.patrec.2017.10.036

[11] H. H. Sultan, N. M. Salem, and W. Al-Atabany, "Multi-classification of brain tumor images using deep neural network," *IEEE Access*, vol. 7, pp. 69215–69225, 2019. https://doi.org/10.1109/ICASSP.2019

[12] A. Kabir Anaraki, M. Ayati, and F. Kazemi, "Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 1, pp. 63–74, 2019. https://doi.org/10.1016/j.bbe.2018.10.004

[13] H. Kutlu and E. Avcı, "A novel method for classifying liver and brain tumors using convolutional neural networks, discrete wavelet transform and long short-term memory networks," *Sensors*, vol. 19, no. 9, 1992. https://doi.org/10.3390/s19091992

[14] P. Afshar, K. N. Plataniotis, and A. Mohammadi, "Capsule networks for brain tumor classification based on MRI images and coarse tumor boundaries," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* IEEE, 2019, pp. 1368–1372. https://doi.org/10.1109/ICASSP.2019.8683759

[15] J. Amin, M. Sharif, M. Raza, and M. Yasmin, "Detection of brain tumor based on features fusion and machine learning," *Journal of Ambient Intelligence and Humanized Computing*, vol. 15, no. 1, pp. 983–999, 2024. https://doi.org/10.1007/s12652-018-1092-9

[16] A. M. Mostafa, M. A. El-Meligy, M. A. Alkhayyal, A. Alnuaim, and M. Sharaf, "A framework for brain tumor detection based on segmentation and features fusion using MRI images," *Brain Research*, vol. 1806, p. 148300, 2023. https://doi.org/10.1016/j.brainres.2023.148300

[17] S. Hossain, A. Chakrabarty, T. R. Gadekallu, M. Alazab, and M. J. Piran, "Vision transformers, ensemble model, and transfer learning leveraging explainable AI for brain tumor detection and classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 3, pp. 1261–1272, 2024. https://doi.org/10.1109/JBHI.2023.3266614

[18] D. R. Nayak, N. Padhy, P. K. Mallick, and A. Singh, "A deep autoencoder approach for detection of brain tumor images," *Computers and Electrical Engineering*, vol. 102, p. 108238, 2022. https://doi.org/10.1016/j.compeleceng.2022.108238

[19] A. Anaya-Isaza, L. Mera-Jiménez, L. Verdugo-Alejo, and L. Sarasti, "Optimizing MRI-based brain tumor classification and detection using AI: A comparative analysis of neural networks, transfer learning, data augmentation, and the cross-transformer network," *European Journal of Radiology Open*, vol. 10, p. 100484, 2023. https://doi.org/10.1016/j.ejro.2023.100484

[20] P. Afshar, A. Mohammadi, and K. N. Plataniotis, "BayesCap: A Bayesian approach to brain tumor classification using capsule networks," *IEEE Signal Process. Lett.*, vol. 27, pp. 2024–2028, 2020. https://doi.org/10.1109/LSP.2020.3034858

[21] G. Li, J. Sun, Y. Song, J. Qu, Z. Zhu, and M. R. Khosravi, "Real-time classifcation of brain tumors in MRI images with a convolutional operator-based hidden Markov model," *Journal of Real-Time Image Processing*, vol. 18, no. 4, pp. 1207–1219, 2021. https://doi.org/10.1007/s11554-021-01072-4

[22] M. Nickparvar, "Brain Tumor MRI Dataset," *Kaggle*, 2021. https://doi.org/10.34740/KAGGLE/DSV/2645886

[23] K. Dang, T. Vo, L. Ngo, and H. Ha, "A deep learning framework integrating MRI image preprocessing methods for brain tumor segmentation and classification," *IBRO Neuroscience Reports*, vol. 13, pp. 523–532, 2022. https://doi.org/10.1016/j.ibneur.2022.10.014

[24] J. Yan, S. Lin, S. B. Kang, and X. Tang, "Learning the change for automatic image cropping," in *Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 971–978. https://doi.org/10.1109/CVPR.2013.130

[25] M. S. Hasan, M. S. Uddin, K. J. Hasan, M. Rahman, and M. Ali, "Deep learning for cultural heritage: A mobile app for monument recognition using convolutional neural networks," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 19, no. 3, pp. 22–40, 2025. https://doi.org/10.3991/ijim.v19i03.50935

[26] A. Elazab *et al.*, "GP-GAN: Brain tumor growth prediction using stacked 3D generative adversarial networks from longitudinal MR Images," *Neural Networks*, vol. 132, pp. 321–332, 2020. https://doi.org/10.1016/j.neunet.2020.09.004

[27] S. Anjana, P. M. Siva Raja, K. Rejini, M. Garuba, and A. Ananth, "Brain tumor detection with bi-directional cascade Gaussian kernel feature-generative adversarial networks," *Biomedical Signal Processing and Control*, vol. 107, p. 107838, 2025. https://doi.org/10.1016/j.bspc.2025.107838

[28] S. Karpakam and N. Kumareshan, "A multi-modality framework for precise brain tumor detection and multi-class classification using hybrid GAN approach," *Biomedical Signal Processing and Control*, vol. 104, p. 107559, 2025. https://doi.org/10.1016/j.bspc.2025.107559

[29] M. Zhou, M. W. Wagner, U. Tabori, C. Hawkins, B. B. Ertl-Wagner, and F. Khalvati, "Generating 3D brain tumor regions in MRI using vector-quantization Generative Adversarial Networks," *Computers in Biology and Medicine*, vol. 185, p. 109502, 2025. https://doi.org/10.1016/j.compbiomed.2024.109502

[30] I. Zine-dine, J. Riffi, K. El Fazazi, M. A. Mahraz, and H. Tairi, "Brain tumor classification using machine and transfer learning," in *Proceedings of the 2nd International Conference on Big Data, Modelling and Machine Learning*, 2021, pp. 566–571. https://doi.org/10.5220/0010762800003101

[31] J. Sachdeva, V. Kumar, I. Gupta, N. Khandelwal, and C. K. Ahuja, "A package-SFERCB-'Segmentation, feature extraction, reduction and classification analysis by both SVM and ANN for brain tumors'," *Applied Soft Computing*, vol. 47, pp. 151–167, 2016. https://doi.org/10.1016/j.asoc.2016.05.020

[32] Z.-d. Iliass, R. Jamal, E. F. Khalid, M. Mohamed Adnane, and T. Hamid, "Alzheimer's disease classification using Histogram of Oriented Gradient, transfer learning, and capsules network," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 4, pp. 5335–5350, 2024. [Online]. Available: https://ijisae.org/index.php/IJISAE/article/view/7325

[33] A. Deshpande, V. Vania Estrela, and P. Patavardhan, "The DCT-CNN-ResNet50 architecture to classify brain tumors with super-resolution, convolutional neural network, and the ResNet50," *Neuroscience Informatics*, vol. 1, no. 4, p. 100013, 2021. https://doi.org/10.1016/j.neuri.2021.100013

[34] D. Pan, J. Shen, Z. Al-Huda, and A. A. Mohammed Al-qaness, "VcaNet: Vision Transformer with fusion channel and spatial attention module for 3D brain tumor segmentation," *Computers in Biology and Medicine*, vol. 186, 2025. https://doi.org/10.1016/j.compbiomed.2025.109662

[35] I. Zine-dine, J. Riffi, K. El Fazazy, I. El Batteoui, M. A. Mahraz, and H. Tairi, "A hybrid model for Alzheimer's disease classification based on neural network architectures enhanced by GAN model," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 21, no. 8, pp. 23–40, 2025. https://doi.org/10.3991/ijoe.v21i08.54363

[36] I. S. Abed, A. T. Khalil, H. M. Amer, S. M. Mohamed, and M. M. Ata, "Advancing osteoporosis diagnosis through state-of-the-art CNNs and vision transformers with ensemble strategies," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 21, no. 11, pp. 116–131, 2025. https://doi.org/10.3991/ijoe.v21i11.56865

[37] G. L. Baldwin, V. Booth-Womack, S. E. LaRose, C. S. Stwalley, and R. M. Stwalley III, "Using support networks as a predictor of success for STEM degrees: Preliminary results detailing a selection process for test subjects engaged in a longitudinal study of low Socioeconomic status American undergraduate students," *International Journal of Engineering Pedagogy (iJEP)*, vol. 12, no. 6, pp. 16–49, 2022. https://doi.org/10.3991/ijep.v12i6.25547

[38] A. Saeed, J. Baber, M. Z. Abbas, A. Sajid, H. Razzaq, and A. A. Khan, "Improving the imbalanced data accuracy using CNN and ReLU," *IETI Transactions on Data Analysis and Forecasting (iTDAF)*, vol. 2, no. 3, pp. 50–58, 2024. https://doi.org/10.3991/itdaf.v2i3.51013

[39] J. Kang, Z. Ullah, and J. Gwak, "MRI-based brain tumor classification using ensemble of deep features and machine learning classifiers," *Sensors*, vol. 21, no. 6, p. 2222, 2021. https://doi.org/10.3390/s21062222

[40] H. M. Rai and K. Chatterjee, "Detection of brain abnormality by a novel Lu-Net deep neural CNN model from MR images," *Machine Learning with Applications*, vol. 2, p. 100004, 2020. https://doi.org/10.1016/j.mlwa.2020.100004

[41] I. Zine-dine *et al.*, "Brain tumor classification using feature extraction and ensemble learning," *Machine Graphics & Vision*, vol. 33, nos. 3/4, pp. 3–28, 2024. https://doi.org/10.22630/mgv.2024.33.3.1

[42] M. Jha, R. Gupta, and R. Saxena, "A framework for in-vivo human brain tumor detection using image augmentation and hybrid features," *Health Information Science and Systems*, vol. 10, no. 1, p. 23, 2022.

[43] A. Çinar and M. Yildirim, "Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture," *Medical Hypotheses*, vol. 139, p. 109684, 2020. https://doi.org/10.1016/j.mehy.2020.109684

7 AUTHORS

Iliass Zine-dine, PhD in Computer Science at the Laboratory of Informatics, Signals, Automation, and Cognitivism (LISAC), Faculty of Sciences Dhar El Mehraz (FSDM), Sidi Mohamed Ben Abdellah University (USMBA), Fes, Morocco (E-mail: iliass.zinedine@usmba.ac.ma).

Jamal Riffi, Professor in Computer Science at the Laboratory of Informatics, Signals, Automation, and Cognitivism (LISAC), Faculty of Science, Dhar El Mahraz (FSDM), Sidi Mohamed Ben Abdellah University (USMBA), Fes, Morocco.

Khalid El Fazazy, Professor in Computer Science at the Laboratory of Informatics, Signals, Automation, and Cognitivism (LISAC), Faculty of Science, Dhar El Mahraz (FSDM), Sidi Mohamed Ben Abdellah University (USMBA), Fes, Morocco.

Ismail El Batteoui, Professor in Computer Science at the Laboratory of Informatics, Signals, Automation, and Cognitivism (LISAC), Faculty of Science, Dhar El Mahraz (FSDM), Sidi Mohamed Ben Abdellah University (USMBA), Fes, Morocco.

Mohamed Adnane Mahraz, Professor in Computer Science at the Laboratory of Informatics, Signals, Automation, and Cognitivism (LISAC), Faculty of Science, Dhar El Mahraz (FSDM), Sidi Mohamed Ben Abdellah University (USMBA), Fes, Morocco.

Hamid Tairi, Professor in Computer Science at the Laboratory of Informatics, Signals, Automation, and Cognitivism (LISAC), Faculty of Science, Dhar El Mahraz (FSDM), Sidi Mohamed Ben Abdellah University (USMBA), Fes, Morocco.