Brain Tumor Classification Using Feature Extraction and Ensemble Learning

Iliass Zine-dine*, Anass Fahfouh, Jamal Riffi, Khalid El Fazazy, Ismail El Batteoui, Mohamed Adnane Mahraz and Hamid Tairi
Laboratory of Informatics, Signals, Automatics and Cognitivism (LISAC),
Faculty of Sciences Dhar El Mehraz (FSDM), Sidi Mohamed Ben Abdellah University (USMBA), Fès, Morocco
**Corresponding author: Iliass Zine-dine (zinedine.iliass@gmail.com)*

Abstract Brain tumors (BT) are considered the second-principal cause of human death on our planet. They pose significant challenges in the field of medical diagnosis. Early detection is crucial for effective treatment and improved patient outcomes. As a result, researchers' studies that deal with tumor detection play a vital role in early disease prediction in the field of medicine. Despite advancements in medical imaging technologies, accurate and efficient classification of BT remains a complex task. This study aims to address this challenge by proposing a novel method for brain tumor classification utilizing ensemble learning techniques combined with feature extraction from neuroimaging data. In the present paper, we present a novel approach for brain tumor classification that contains ensemble learning methods following the extraction of important features from brain tumor images. Our methodology involves the preprocessing of neuroimaging data, followed by feature extraction using descriptor techniques. These extracted features are then utilized as inputs to ensemble learning classifiers. Experimental results demonstrate the efficacy of the proposed approach in accurately classifying brain tumors with high precision and recall rates. The ensemble learning framework, combined with feature extraction, outperforms several benchmark models and methods commonly used in brain tumor classification, including AlexNet, VGG-16, and MobileNet, in terms of classification accuracy and computational efficiency. The proposed method that integrates ensemble learning techniques with feature extraction from neuroimaging data offers a promising solution for improving the accuracy and efficiency of brain tumor diagnosis, thereby facilitating timely intervention and treatment planning. The findings of this study contribute to the advancement of medical imaging-based classification systems for brain tumors, with implications for enhancing patient care and clinical decision-making in neuro-oncology.

Keywords: brain tumor, Histogram of Oriented Gradients, Discrete Wavelet Transform, ensemble learning.

1. Introduction

The brain is the principal organ of the nervous system, and it is the most complex organ in the human body. It consists of nerve cells and tissues to control the most basic functions of the body, such as muscular movement, breathing, and senses. For the mentioned reasons, early detection of brain tumors represents a crucial task in the medical field [45]; a brain tumor is a form of cancer that affects the central nervous system, according to the definition provided by the World Health Organization (WHO). It was categorized as a deadly disease in 2016 [2]. In general, a brain tumor is described as a collection of abnormally growing brain cells. The purpose of this study is to classify cancer in MRI images.

The classification of brain tumors is one of the most challenging tasks in the medical field because various criteria are dependent on the structure of the nervous system's tissue and cells [50]. Currently, the development of technology and the digitalization of medical devices help doctors properly detect and classify brain tumors in the early stages; the field of machine learning is focusing on this task. Rajan and Sundar [36] present an architecture for classifying brain cancers using support vector machines (SVM) classifiers and feature extraction like K-means clustering. This classifier is integrated with Fuzzy C- Means (KMFCM) and active contour by level set for tumor segmentation. In another study, Murugan et al. [5] proposed a system including enhancement, transformation, feature extraction, and classification with machine learning models.

The architectures of deep learning are a subset of machine learning that allows computers to make predictions and describe conclusions based on data thanks to their capacity to learn data representations. These methods are widely utilized in medical imaging categorization and are considered one of the most important computational intelligence techniques. In this context, Das et al. [11] presented an approach that consists of two principal steps. First, preprocess the images using different image processing techniques, then classify the preprocessed images using convolutional neural networks (CNN). In this regard, Paul et al. [34] applied two types of neural networks, fully connected and convolutional neural networks, which were used to classify brain images with different tumor types. Thus, Shree et al. [22] proposed a probabilistic neural network (PNN) approach that relies on feature extraction techniques (such as noise suppression, gray level co-occurrence matrix (GLCM), and the growth of brain tumor region segmentation based on DWT) to reduce complexity and improve performance. However, the field of deep learning has several limitations. Firstly, there is a strong use of neural network methods in this field of brain tumor classification, with the lack of feature extraction functions; secondly, deep learning architectures do not work perfectly with small datasets; they need greedy datasets; and finally, deep network training necessitates the meticulous tuning of numerous parameters, and suboptimal tuning can lead to overfitting or underfitting. These restrictions do not allow researchers to find better performance metrics. In consequence, we thought about a method that gives us better prediction scores.

Recently, new research has been based on the extraction of features from the image, which will be fed to the classifiers. Mircea et al. [15] have proposed an approach using different wavelet transforms and support vector machines to detect and classify the brain tumor. In this regard, Nabizadeh and Kubat [30] have studied a method based on the extraction of features with the Gabor wavelet that is able to detect slices that include tumors and delineate the tumor area. Another type of image feature extraction based on the extraction of information from image textures is widely used in medical image analysis. Singh et al. [43] proposed a hybrid technique for automatic classification of MRI images by first extracting the features using Principal Component Analysis (PCA) and Gray-Level Co-occurrence Matrix (GLCM). The extracted features are fed as input to a Support Vector Machine (SVM) classifier, which classifies the brain image as normal or abnormal. Thus, LBP (Local Binary Pattern) is a commonly used technique for texture description and pattern recognition in images [46]. Another well-known feature extraction technique is the histogram of oriented gradients, which is a computer vision and image processing technique used to detect objects [25].

In this study, we have proposed an approach based on ensemble learning using the stacking model, we utilized two feature descriptors, HOG and DWT, to capture discriminative information from images for use as inputs to classification algorithms. This not only accelerates the training of these models but also helps prevent overfitting. The extracted characteristics were combined into a vector, evaluated, and selected using a RandomForestRegressor to evaluate and select the most essential features. This process reduces the dimensionality of our dataset, improves the interpretability and reliability of our model's decisions, and provides refined inputs for subsequent models. This process ultimately enhances the overall effectiveness and robustness of the classification framework. The proposed system recorded a more satisfactory classification performance reaches 94%. It appears that the functionalities used in this classification task have an important and effective role. Therefore, our approach demonstrates significant value in the early prediction of AD through the advancement of computer vision and machine learning methods, applied within the medical domain.

The following is how the rest of the paper is organized. The related studies are presented in the second Section. Also, our main contribution introduced in this study is described in the final part of that Section. It is followed by a description of the methods and techniques used in our approach in Section 3. Within this Section, in Subsections 3.1 and 3.1.1 the dataset used is presented and descried, and in Subsection 3.1.2 the data preprocessing, including data augmentation, is presented. In the following Subsection 3.2 the features derived from images are described, including the Histogram of Oriented Gradients in Subsection 3.2.1, the Discrete Wavelet Transform in Subsection 3.2.2, and their concatenation and feature selection with the RandomForestRegressor in Subsection 3.2.3. The classification methodology is presented in Subsection 3.3, divided into machine learning methods in Subsection 3.3.1 and ensemble learning methods in Subsection 3.3.2. The whole Section 3 on materials and methods is concluded with its discussion in Subsection 3.4.

The experiments and their results are presented in Section 4, with four Subsections: 4.1 on the experimental setting, 4.2 on performance evaluation measures, 4.3 on the results and finally Subsection 4.4 on their discussion. The whole paper is concluded and the perspective for future work is outlined in Section 5.

2. Related work

The primary goal of this section is to review the existing research on employing extraction features, machine learning, and deep learning models to identify and classify brain tumors. There are several works in this area that deal with the early prediction of brain disease, which is based on computer-aided diagnostic methods (CAD) without surgery or invasive methods. Sobhaninia et al. [44] proposed an architecture that is based on an encoder layer and uses post-max-pooling features for residual learning for brain tumor classification.

In this regard, there are several image-processing architectures interested in detecting and classifying tumors. Our current research focuses on the early prediction of brain tumors, which is similar to the work that will be cited. These studies belong to the same medical field and use the same techniques and methods of computer vision and infographics for this task. However our proposed method provided effective results, surpassing various state-of-the-art experiments on the topic of brain tumors in terms of accuracy.

Various deep convolution neural networks have already been trained are used to extract deep features from magnetic resonance (MR) brain images. Kang et al. [19] presented an architecture for classifying brain cancers using a collection of deep features and machine learning classifiers. In this area, Deepak et al. [13] proposed a method for classifying MRI images; this method is based on transfer learning by applying several models of machine learning to the MRI image dataset of the brain tumor, which is already pre-trained on a VGG-16 model of convolutional neural network. Despite the accurate results of these methods, they remain poor, mainly when dealing with large databases. Ari et al. [4] presented a method based on a pre-processing brain tumor dataset (resize, crop lesion, segment lesion, etc.). Kaplan et al. [20] used a feature extraction approach called Local Binary Patterns (LBP), which is a statistical image processing technique that allows us to extract useful and important characteristics from images. In the domain of computer vision, another approach that gives better results at the classification level is based on the combination of methods like concatenation and confusion of vectors of extracted features, there are several approaches use multiple techniques combined to obtain a model more efficient and effective than a model built with a single algorithm. Abbasi et al. [1] used techniques for segmentation and detection to distinguish between different brain regions based on feature extraction from MRI images or learning features like the local binary pattern (LBP) and the histogram of oriented gradients (HOG). Another type of method based on deep learning proves to be very effective while managing vast amounts of data. Mohsen et al. [29] presented a new method for classifying brain tumor images using a deep neural network (DNN) learning method that used fuzzy C-means to segment the images, discrete wavelet transforms (DWT) to extract the features, principle component analysis (PCA) to reduce the dimensions, and DNN for classification.

The majority of existing medical MR imaging research focuses on the automatic classification and segmentation of tumor regions. Several researchers have recently presented various strategies for detecting and segmenting the tumor region in MR images. Convolutional neural networks are powerful architectures based on deep convolutional layers that automatically extract robust functionality from the input space related to traditional neural network layers. Rehman et al. [40] proposed CNN models such as AlexNet, GoogLeNet, and VGGNet to classify MRI images of brain tumors.

Timely, deep learning is generally applied in the medical industry. The fundamental CNNs that are applied for classification tasks have similar architectures. A CNN architecture is made up of a series of feed-forward layers that employ convolutional filters and pooling layers, following the last pooling layer, CNN uses many fully connected layers to turn the previous layers' 2D feature maps into 1D vectors for classification. In summary, CNNs rely on three characteristics. Firstly, each layer's units get input from the previous layer's units, which are all in the same tiny neighborhood. This technique allows for the extraction of basic features such as edges and corners. Secondly, in the subsequent layers, these features will be merged to detect higher-order features. The concept of shared weights, which is the second crucial attribute, means employing similar feature detectors throughout the image. Thirdly, CNNs frequently have many sub-sampling layers, which are either advantageous or harmful because this information varies for different situations according to the specific position of characteristics [41].

Although CNNs are beneficial in a variety of applications, they have several flaws, particularly in the sub-sampling layers, which provide only a tiny amount of translational invariance and lose the precise location of the most active feature detectors. A capsule neural network (CapsNet) is a sort of artificial neural network (ANN) that can be used to improve model hierarchical relationships in a machine learning system. To classify brain tumors, Afshar et al. [3] proposed a model based on the architecture of CapsNet that allows access to the tumor tissue without distracting it from the central target.

The majority of extant medical MR imaging research focuses on the automatic classification and segmentation of tumor regions. Several researchers have recently presented various strategies for detecting and segmenting the tumor region in MR images; Table 1 represents previous work carried out on different datasets.

Recent research has shown that deep learning techniques are widely used in expert and intelligent systems as well as in medical image analysis. The methodologies described previously presented limits at the level of data processing, more precisely in the feature extraction phase. These approaches took into account only the binary categorization (normal and abnormal) of the MRI image dataset, and they ignored extracting the crucial features and from the images. In addition, throughout the course of our investigation, it became evident that the referenced models exhibit a scarcity of data,

Tab. 1. Related work approaches to classification methods, feature extraction, and accuracy of brain tumor classification.

Authors	Feature Extraction and Classification Methods	Dataset	Accuracy
Díaz-Pernas et al. [14] 2021	Multi-pathway convolutional neural network (CNN)	3064 MRI	97.3%
Das et al. [12] 2019	Advanced Deep Learning-based Solutions (CNN)	3064 MRI	94.39%
Paul et al. [34] 2017	Fully connected and CNN	3064 MRI	91.43%
Kumar and Shree. [22] 2018	Probabilistic neural network (PNN)	650 MRI	95%
Khawaldeh et al. [21] 2017	CNN	587 MRI	91.16%
Hemanth et al. [16] 2019	CNN	220 MRI	94.5%

intricate computational processes, and suboptimal performance. For the mentioned reasons, it is recommended to search for a new approach that exceeds these constraints and gives us better prediction scores.

Real-time performance is a critical factor in medical diagnosis, particularly in emergency situations where timely and accurate decisions are essential for patient treatment and prognosis. Evaluating a model's inference time and computational resource requirements ensures its suitability for real-world applications. Models designed for such scenarios must balance speed and accuracy to provide reliable diagnostics without compromising computational efficiency [23].

The main contribution of this study can be summarized as follows: during the preprocessing phase, we employed common computer vision and infographic techniques to facilitate subsequent tasks. Then, we utilized two descriptors, HOG and DWT, to extract relevant and significant features, accelerate model training, avoid overfitting, and thereby enhance the overall effectiveness and robustness of the classification framework. These extracted characteristics were combined into a vector, evaluated, and selected using a RandomForestRegressor, and then considered as inputs for classification machine learning algorithms. To validate the effectiveness of our approach, experiments were conducted on a well-known brain tumor dataset, and the results were compared with existing methodologies. Our approach has shown considerable value in the early prediction of brain tumors through advancements in computer vision and machine learning methods and their applications in the medical domain.

3. Material and methods

The general design of our suggested method is described in this section. We have used an approach that consists of two complementary main phases: feature extraction and classification. The features recovered from the first phase will be considered, such as the inputs from the machine learning classifiers of the second phase. The main objective of

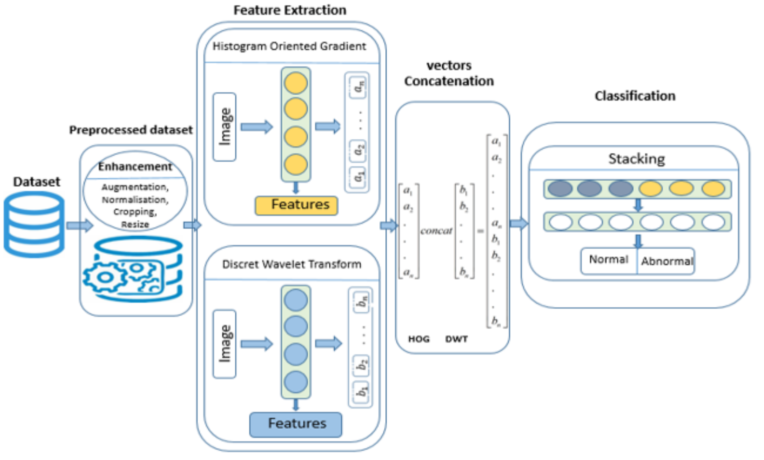


Fig. 1. Proposed architecture of histogram of oriented gradients, discrete wavelet transform, and ensemble learning (Stacking) for brain tumor classification.

this study is to find a higher classification score. Figure 1 illustrates the architecture of our suggested approach to classify brain tumors, which will be described in detail in Sec. 3.3. We have detailed each component of the proposed approach with greater clarity in the following Sections.

3.1. Dataset

The brain tumor dataset utilized in our research is crucial for classification, offering real-world data reflecting clinical complexities. It enables algorithm development and evaluation, facilitating supervised learning and serving as a benchmark for advancing medical image analysis. In this research, we used the free Kaggle brain tumor dataset [32].

3.1.1. Data description

The dataset we have used contains 253 brain MRI images split into two groups: 'yes' contains 155 tumorous brain MRI images, and 'no' contains 98 non-tumorous brain MRI images. We started to preprocess the dataset by applying imaging methods like normalization, resizing, cropping, and augmentation, to facilitate the employment of the following functions. These techniques applied in the data processing phase allowed us to have 2065 images. Distribution of the dataset across different categories is presented in Table 2.

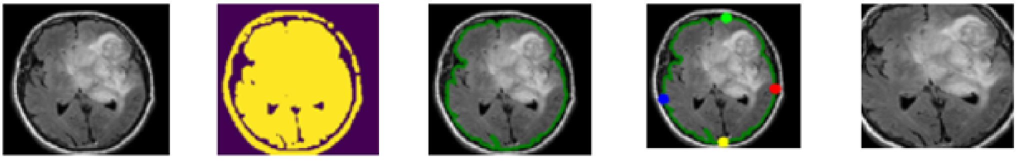


Fig. 2. Image representation of different stages of cropping images from the dataset: (**a**) original image; (**b**) thresholded; (**c**) outer contour (green); (**d**) edge points (R, G, B, Y); (**e**) cropped image.

3.1.2. Data pre-processing

The trend of processing datasets containing images for predictive purposes has gained prominence in the domains of computer graphics and computer vision [28]. In this manuscript, we will use image-processing functions. The primary techniques used in this part of the treatment will be discussed below.

Data crop: Nearly all of the images in our brain MRI datasets have undesirable spaces. Hence, it results in subpar classification performance. Therefore, it is vital to crop the pictures in order to eliminate unnecessary portions and use only the pertinent information on [49]. In this study, we employed the cropping approach that computes extreme points and returns a geographic subset of an object based on specifications provided by an extent object. Figure 2 illustrates how the MR images were cropped using an extreme point computation. This cropping method consists of five phases: 1^o we load the original MR images. 2^o We apply thresholding to the MR images in order to create binary images. 3^o We also undertake dilation and erosion processes to reduce image noise. 1^o We determine the four extreme points of the images (extreme top, bottom, right, and extreme left) using the largest contour of the threshold images. 5^o We crop the image based on the edge and extreme point data. Bicubic interpolation is used to enlarge cropped tumor images.

Tab. 2. Division of the dataset of images into training, validation, and testing, and the tumorous and non-tumorous classes.

Step	No. of images			Percentage of images [%]		
	Tumorous, 'yes'	Non-tumorous, 'no'	Total	Tumorous, 'yes'	Non-tumorous, 'no'	Total
Train	885	560	1445	61.3	38.7	70
Validation	190	120	310	61.3	38.7	15
Test	190	120	310	61.3	38.7	15
Total	1265	800	2065	61.3	38.7	100

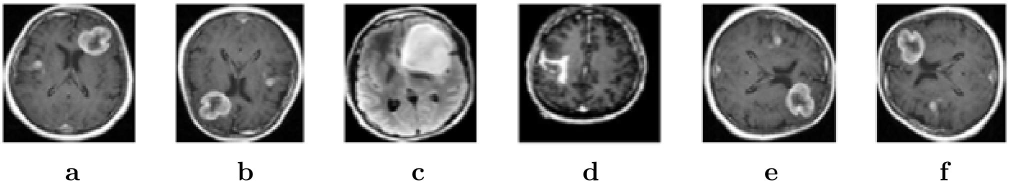


Fig. 3. Example of data augmentation: (**a**) vertical flip; (**b**) horizontal flip; (**c**) brightness increased; (**d**) vertical shift; (**e**) rotation +90°; (**f**) rotation −90°.

The cropping function plays a crucial role in feature extraction for classification tasks. Cropping involves removing the outer parts of an image to focus on the most relevant region, which can enhance the performance of classification algorithms. By isolating the area of interest, cropping reduces noise and irrelevant information, leading to a more accurate representation of the essential features. This process not only helps in concentrating on the significant aspects of the image but also reduces the computational complexity by decreasing the amount of data that needs to be processed. Consequently, cropping contributes to improving the efficiency and accuracy of the classification model.
Data augmentation: Due to the relatively modest size of our MRI dataset, we performed image augmentation to increase the size of the dataset. Data augmentation is a technique that involves transforming the original dataset to produce a synthetic dataset, it is a procedure that generates additional training data by applying transformations to existing data to obtain new data [26]. This method involves creating numerous duplicates of the original image with various scales, orientations, locations, brightness, and other characteristics. Results from previous related work showed that augmenting existing data can increase accuracy model classification, rather than collecting new data. Figure 3 illustrates the augmentation techniques applied to the original dataset to generate the new dataset.

The function of image data augmentation is highly beneficial for feature extraction in classification tasks in machine learning. Data augmentation involves creating new training samples by applying random transformations such as rotation, scaling, translation, and flipping to the original images. This technique helps to increase the size and diversity of the training dataset, which is particularly valuable when dealing with limited data. By providing more varied examples, data augmentation allows the model to generalize better to new, unseen data, thereby enhancing its robustness and accuracy. Moreover, it helps to prevent overfitting by ensuring that the model does not memorize the training data but learns to identify the underlying features that are relevant for classification. Consequently, image data augmentation is a crucial step in improving the performance and reliability of machine learning models in image classification tasks.

3.2. Feature extraction

Feature extraction refers to the process of transforming raw data into digital features. These features will be processed while preserving the information from the original dataset. This method performs better than directly applying machine learning to raw data [47]. After the dataset had been pre-processed, we used a descriptor-based approach as a feature extractor to extract pertinent characteristics, speed up the models training; avoid overfitting, and thereby augmenting the overall effectiveness and robustness of the classification framework. Then, we concatenated these characteristics gleaned by HOG and DWT to create a unified input suitable for feeding into the final classifier. Lastly, we trained these extracted features using a RandomForestRegressor model to select the crucial features. In the following two sections, we describe the two descriptors used in our approach HOG and DWT.

3.2.1. Histogram of Oriented Gradients

Histogram of Oriented Gradients (HOG) is a feature descriptor used in computer vision and image processing for object detection; in other words, HOG is a technique for characterizing textures and shapes in an image, similar to Canny Edge Detector and Scale-Invariant Feature Transformation (SIFT) [33].
The HOG has the purpose of quantifying the distribution of local gradient orientations in an image. This approach counts occurrences of gradient orientation in localized parts of an image. This method is comparable to edge orientation histograms, scale-invariant feature transformation descriptors, and shape contexts, but it is more accurate because it is computed on a dense grid of equally spaced cells and employs overlapping local contrast normalization. Four types of normalization are explored. The unnormalized vector containing all the histograms of a single block is denoted by v, its k-norm by $\|v\|_k$, and e is a low-value constant. The normalization factor is then defined by:

- L2-norm:

$$f = \frac{v}{\sqrt{\|v\|_2^2 + e^2}} \tag{1}$$

- L2-hys: L2-norm followed by clipping (limiting the maximum values of v to 0.2) and renormalizing.
- L1-norm:

$$f = \frac{v}{\|v\|_1 + e^2} \tag{2}$$

- L1-sqrt:

$$f = \sqrt{\frac{v}{\|v\|_1 + e^2}} \tag{3}$$

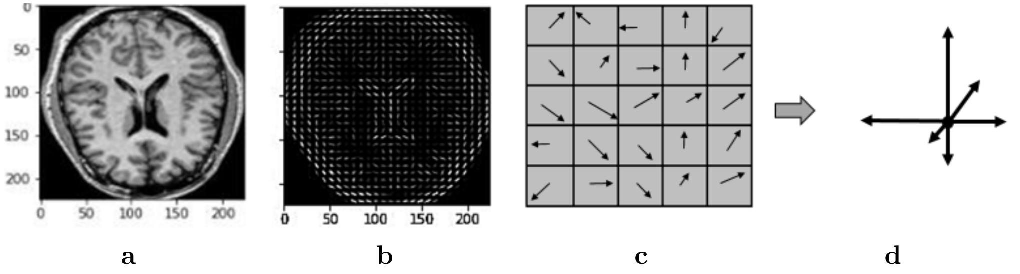


Fig. 4. Representation of features extracted from brain images by the HOG algorithm. (**a**) Original image; (**b**) representation of the image with the application of a HOG; (**c**) gradients in a cell; (**d**) Histogram of Oriented Gradients in six directions in 3D.

The L2-Hys, L2-norm, and L1-sqrt norms achieve similar performance, while L1-norm performs worse, but still significantly outperforms no normalization. In our approach we applied the first L2-norm normalization. Figure 4 illustrates the application of the HOG algorithm to our dataset.

The integration of the HOG function into our approach holds significant importance for several reasons. Firstly, HOG is widely acknowledged for its ability to capture texture and shape information within an image, making it a powerful tool for feature extraction. By leveraging HOG as an image descriptor in our methodology, we can extract relevant and discriminative features, which are crucial for the task of medical image classification. Furthermore, HOG provides a compact representation of the extracted features, aiding in reducing the dimensionality of the data and enhancing the efficiency of subsequent classification algorithms. By incorporating HOG into our approach, we can improve the quality of the extracted features and, consequently, the accuracy and robustness of our classification model. For our task, we aggregated the features extracted by HOG into a vector of dimensions $(1, 1, 1000)$ to later concatenate it with another vector extracted by DWT, which has the same dimensions.

3.2.2. Discrete Wavelet Transform

Discrete wavelet transform is a data transformation technique that allows the signal to be represented in the form of wavelet coefficients, which can be useful for data compression, feature detection, noise reduction, frequency analysis, and other tasks [27].

Nowadays, DWT is widely used to extract the most relevant features at different orientations and scales from temporal signals or time series data, which can facilitate the modeling and analysis of these data, i.e., Gabor-wavelets capture the local structure of the image corresponding to spatial frequency (scales), space localization, and orientation selectivity [24]. The wavelet coefficients obtained through DWT can be used as features to train machine learning models for classification, regression, or other data analysis tasks, as it provides localized time-frequency information of a signal using cascaded filter banks of high-pass and low-pass filters to extract features in a hierarchical manner [30].

The principle of the algorithm consists of dividing the image into four blocks at each iteration: three blocks concerning the details of the image, and the fourth corresponding to the most important information for the eye (low frequencies), which serves as a basis for the next iteration. In mathematics, a wavelet ψ is a summable square function of Hilbert space $Ł^2(R)$, the more often oscillating and with zero mean, chosen as a multi-scale analysis and reconstruction tool. Wavelets are generally found in families, made up of a mother wavelet and the set of its images by the elements of a subgroup of the group of affine transformations of R^n. We thus define a family $\psi_{s,\tau}$ where $(s,\tau) \in R^{+*} \times R$, of wavelets from the mother wavelet ψ:

$$\forall t \in R, \psi_{s,\tau}(t) = \frac{1}{\sqrt{s}}\,\psi\,\frac{t-\tau}{s} \tag{4}$$

By extension, families of functions on submanifolds of $\mathbb{R}$ invariant by a transformation group locally isomorphic to the affine group can also be qualified wavelet families.

We use wavelet coefficients for generating the initial features. The wavelet transform is traditionally used for feature extraction. The provision of localized frequency information about a function of a signal is the main advantage of wavelets and is particularly beneficial for classification. Earlier, wavelets were used as a feature extraction method for discrimination; they have advantages in fields like image processing, image watermarking, medical imaging, image compression, and many more. They are also used to denoise medical images. Orthogonal wavelets have always played a main role in biomedical image processing [48].

By applying DWT, we are able to decompose an image into the corresponding subbands with their relative DWT coefficients. The DWT is implemented using cascaded filter banks, in which the low pass and high pass filters satisfy certain specific constraints [8]. At each scale of feature extraction by this technique, there are four subband images (LL, LH, HH, and HL). The LH, HL, and HH sub-bands may be thought of as the detailed components of the image, while the LL sub-band can be thought of as the approximation component. For DWT decomposition at the next scale, only the sub-band LL is utilized for feature extraction. Additionally, the output feature vector uses the LL sub-band at the final level. Figure 5 illustrates DWT decomposition and its application to our dataset.

The use of DWT in feature extraction for classification tasks offers several key advantages. DWT is highly effective at capturing both spatial and frequency information from images, making it an excellent tool for identifying relevant features. This transform decomposes the image into different frequency components, allowing for the isolation of important details and patterns at various scales. By leveraging DWT, we can extract multi-resolution features that are essential for distinguishing between different classes in

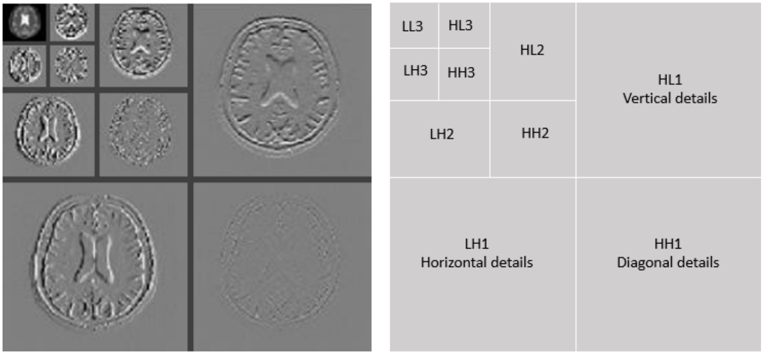


Fig. 5. Representation of the three steps of the features extracted from the brain image by the DWT algorithm.

medical image classification. Additionally, DWT helps to reduce the dimensionality of the data, which not only enhances computational efficiency but also mitigates the risk of overfitting. Consequently, incorporating DWT into our feature extraction process can significantly improve the accuracy and robustness of the classification model.

3.2.3. Concatenation

Following feature extraction, the subsequent procedural step is the concatenation of the extracted features. In the field of machine learning, the concatenation of input vectors in a model is a function that combines the inputs into a single input vector, puts them end to end, and then processes them according to the chosen method [37]. This technique allows for the combination of various sources of information, capturing diverse and complementary aspects of the data, thereby enhancing the overall data representation. Additionally, this approach can aid in reducing the dimensionality of the data by combining multiple features into a single vector, thus facilitating further processing and analysis. Furthermore, concatenating features can lead to more robust and effective classification models by integrating information from different modalities or sources, potentially resulting in improved prediction performance [39]. In our study, we concatenated a vector of (1.1.1000) of feature extracted by HOG with a vector of (1.1.1000) of feature extracted by DWT into a vector of (1.1.2000), then evaluated them by a RandomForestRegressor to select the most relevant ones, reduce the dimensionality of our dataset, and bolster the interpretability and dependability of our model's decisions. Finally, the resulting features were employed as inputs for the classifier. In this research, we employed a 10-fold cross-validation approach to ensure the reliability and statistical significance of the results, as well as to rigorously evaluate the model's performance. The dataset was divided into 10 equal subsets (folds). In each iteration, one fold was designated as the test set, while the remaining nine folds were used for training the model. This process was repeated 10 times, with each fold serving as the test set exactly once. By averaging the performance metrics across all iterations, we achieved a comprehensive and unbiased assessment of the model's effectiveness.

3.3. Classification

Automatic classification, or supervised classification, is the algorithmic categorization of objects based on statistical data [6]. In our study, our goal is to classify brain tumors. We started by processing our dataset using usual image processing functions such as normalization, resizing, augmentation, and cropping.

Then, we applied two descriptor functions, HOG and DWT, to extract more information and to facilitate the task at the classification stage, to accelerate model training, to prevent overfitting, and thus to enhance the overall efficacy and resilience of the classification framework. The classification step consists of applying several machine learning classification models to our extracted features in order to find a model with high classification quality measures, that would give good results of identification of brain tumors.

Now we shall describe the elements of the classification methodology in detail. Its general structure has been already shown in Fig. 1, p. 9.

3.3.1. Machine learning models

Machine learning models can be conceptualized as algorithms trained to discern patterns in novel data and formulate predictions. These models are mathematically represented as functions designed to process input data, predict outcomes, and yield corresponding outputs [17]. Generally, these models undergo training on a designated dataset and are parameterized to extrapolate predictions for previously unseen data. Below, we list the different classification models used in our approach.

Support Vector Machine: It is one of the most efficient classification algorithms, having been proposed by Vapnik [9]. SVM converts the initial data space into a new space with a higher dimension using the kernel function $K(x_n, x_i)$. The following definition fits the hyperplane function used to separate the data:

$$f(x_i) = \sum_{n=1}^{N} \alpha_n y_n K(x_n, x_i) + b\,, \tag{5}$$

where x_n is support vector data (features extracted from brain MR image), α_n is Lagrange multiplier, and y_n represents a target class.

Gaussian Naïve Bayes: The machine learning classifier known as the Naive Bayes classifier operates under the assumption of conditional independence between the attributes and the class [31]. In this study, one of our ML classifiers for classifying brain tumors is the Gaussian NB classifier. The conditional probability $P(Y|X)$ in the Gaussian NB classifier is determined as the sum of the individual conditional probabilities under the naive independence assumption as follows:

$$P(Y|X) = \frac{P(Y) \times P(X|Y)}{P(X)} = \frac{P(Y)\prod_{i=1}^{n} P(x_i|Y)}{P(X)} \tag{6}$$

where X is the presented data instance (an extracted deep feature from brain MR image) which is represented by its feature vector $(x_1, ..., x_n)$, y is a class target (type of brain tumor) with two classes (normal and tumor) for two MRI datasets ensemble learning models.

K-Nearest Neighbors (k-NN): One of the simplest classification methods is k-NN. It makes predictions right away using the training set that is memorized. For example, to categorize a new data instance (a deep feature from a brain MR image), k-NN selects the set of k objects from the training instances that are closest to the new data instance by calculating the distance and assigns the label with two classes (normal or tumor). The selection is based on the majority vote of its k neighbors for the new data instance. The most popular methods for evaluating how close new data instances are to training data examples are Manhattan distance and Euclidean distance [38]. In this study, we applied the k-NN method using the Euclidean distance metric. Data points x and y's Euclidean distance d is determined as follows:

$$d(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{7}$$

Random Forest: Breiman's ensemble learning technique RF [7] classifies new data instances (a deep feature of a brain MR image) into a class target (a type of brain tumor) with two classes (normal and tumor) for two MRI datasets. It does this by building multiple decision trees using the bagging approach. When building the decision trees, RF randomly chooses n features or attributes to determine the best split point using the Gini index as a cost function. This random selection of attributes results in less correlation between the trees and lower ensemble error rates. To predict the class target of a new data instance, the new observation is fed into all classification trees of the RF. When RF collects the predictions for each class, it chooses the class with the most votes as the new data instance's class label.

3.3.2. Ensemble learning models

Ensemble methods are machine learning techniques that combine several base models to produce one optimal predictive model. In other words, ensemble methods are techniques that combine the predictions of multiple machine learning models (called base learners or learners) to obtain a more robust and accurate prediction. Ensemble methods are commonly used to improve the performance of machine learning models by reducing overfitting and increasing generalization [42]. One of the most well-liked ensemble machine learning strategies is stacking, which is used to forecast several nodes to create a new model and enhance model performance (in this research, it is the model that gave us effective results among the models used). By stacking, we can train many models to tackle related issues and then create a new model with higher performance based on the output of all the trained models. The three main classes of ensemble learning methods are bagging, stacking, and boosting.

Bagging (Bootstrap Aggregating): This method creates multiple data samples from the training set using bootstrapping (sampling with replacement) and trains a base model on each sample. The predictions from these models are then aggregated (usually using a majority vote) to form a final prediction.

Boosting methods assign weights to training examples based on the performance of previous models. The base models are trained iteratively, with an emphasis on misclassified examples. The predictions from each model are weighted to obtain a final prediction.

Stacking combines several base models using a meta-learning model that learns from the predictions of the base models. The meta-learning model takes the predictions from the base models as input and generates the final prediction.

3.4. Discussion on data and methods

Smaller datasets, despite their limitations, can enhance model generalization when combined with proper preprocessing, enhancement methods, and regularization. They encourage the model to focus on core features, reducing the risk of overfitting. Data augmentation techniques artificially increase variability, further improving generalization. Enhancement methods such as HOG and DWT can also compensate for the small dataset size by effectively extracting critical features. With careful management, small datasets can support a balanced and effective learning framework.

Several attempts have been made to classify brain tumors based on MRI using various machine and ensemble learning classifiers. In this research, we employed seven well-known and diverse classifiers, including Naive Bayes, k-NN, RF, SVM, Gradient Boosting, XGBoost, and Stacking, to determine which classifier works best for MRI-based brain tumor classification. A crucial factor in effectively building the model for MRI-based brain tumor classification is designing a method to generate a discriminative and informative feature from brain MR images. This is because the performance of machine and ensemble learning classifiers heavily depend on the input feature type, i.e., the features extracted from the image and the parameters used in the classifier.

Real-time performance is crucial in medical diagnosis, especially in emergencies requiring fast and accurate decisions. In our study, we evaluated the model's inference time and achieved classification results within 4 seconds, ensuring a balance between speed and accuracy for practical applications.

4. Experiments and results

In this section, we present the experimental setup and results of our study on brain tumor classification using feature extraction and ensemble learning techniques. Initially, we provide an overview of our approach and the frameworks employed for implementing the code. Following this, we detail the evaluation metrics used to assess the performance of our classification models. Subsequently, we conduct a comprehensive comparison between our approach and related works in the field. Finally, we delve into a discussion of our findings and offer insights into future research directions.

4.1. Experimental setting

In this experiment, following the image pre-processing phase that included normalization, resizing, augmentation, and cropping, we employed two descriptor functions HOG and DWT as feature extractors to extract pertinent features speed up the models training, avoid overfitting, and thereby augmenting the overall effectiveness and robustness of the classification framework. Then, we concatenated these characteristics, evaluated, selected and used them as inputs for the machine learning classifiers. This novel strategy improved our classification prediction score. All trials were carried out on a computer with an NVIDIA GeForce GTX 1070 Ti GPU.

4.2. Performance evaluation

Evaluating the performance of a machine learning model involves employing a range of metrics and techniques to gauge its effectiveness, accuracy, and ability to generalize to new data. These metrics help assess how well the model might perform on unseen data and identify issues like overfitting or underfitting. Our experiment's effectiveness was determined using specific performance metrics tailored to the classification task, including precision, recall, accuracy, and F1-score.

Precision: It represents the percentage of relevant results and is defined as:

$$\mathrm{Precision} = \frac{\mathrm{TruePositive}}{\mathrm{TruePositive} + \mathrm{FalsePositive}} \tag{8}$$

Recall: It denotes the percentage of correctly classified total relevant results by the proposed algorithm and is defined as:

$$\mathrm{Recall} = \frac{\mathrm{TruePositive}}{\mathrm{TruePositive} + \mathrm{FalseNegative}} \tag{9}$$

Tab. 3. Comparison between the performance of machine learning and ensemble learning algorithms with the HOG descriptor as the only feature extraction algorithm.

Group	Model	Performances measures [%]			
		Accuracy	Precision	Recall	F1-Score
Machine Learning	Random Forest	82.3	82.3	82.3	82.3
	Support Vector Machine	88.1	86.4	93.0	89.6
	K-Nearest Neighbors	88.1	88.1	88.1	88.1
	Gaussian Naïve Bayes	61.5	63.4	61.5	60.7
Ensemble Learning	Gradient Boosting	84.2	84.2	84.2	84.2
	XGBoost	83.4	83.4	83.4	83.4
	Stacking	90.6	90.6	90.6	90.6

Accuracy: Formally, accuracy has the following definition:

$$\text{Accuracy} = \frac{\text{TruePositive} + \text{TrueNegative}}{\text{Total}} \tag{10}$$

F1-score: It is a machine learning measure commonly used in classification models and is defined as:

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{11}$$

4.3. Results

The empirical results were derived from the 'Brain MRI Images for Brain Tumor Detection' dataset, obtained from a Kaggle competition dedicated to brain tumor classification tasks. The primary aim of this experiment was to extract features using two distinct descriptors, namely HOG and Wavelet, with the intent to expedite machine learning model training, prevent overfitting, and enhance the overall effectiveness and robustness of the classification framework. These extracted features were concatenated into a vector, evaluated, and selected using the RandomForestRegressor to identify the most significant features, thereby reducing the dimensionality of our dataset, and enhancing the interpretability and reliability of our model's decisions, serving finally as inputs for machine learning classifiers. Subsequently, the outputs of these classifiers were aggregated to identify the most accurate predictions and improve the new model's performance through stacking, resulting in a high-performance classification. Tables 3 and 4 provide the prediction scores achieved by each descriptor across various machine and ensemble learning algorithms.

In summary, our empirical findings indicate that using a single descriptor for feature extraction, combined with the machine learning algorithms in our methodology, results in

Tab. 4. Comparison between the performance of machine learning and ensemble learning algorithms with DWT as the only feature extraction algorithm.

Group	Model	Performances measures [%]			
		Accuracy	Precision	Recall	F1-Score
Machine Learning	Random Forest	80.4	80.4	80.4	80.4
	Support Vector Machine	86.3	84.6	93	87.5
	K-Nearest Neighbors	85.8	85.8	85.8	85.8
	Gaussian Naïve Bayes	60.2	61.7	59.8	59.1
Ensemble Learning	Gradient Boosting	82.3	82.3	82.3	82.3
	XGBoost	81.7	81.7	81.7	81.7
	Stacking	88.5	88.5	88.5	88.5

Tab. 5. Comparison between the performance of machine learning and ensemble learning algorithms with the combination of features extracted from HOG and DWT.

Group	Model	Performances measures [%]			
		Accuracy	Precision	Recall	F1-Score
Machine Learning	Random Forest	82.6	82.6	82.6	82.6
	Support Vector Machine	89.7	88.7	93.2	89.7
	K-Nearest Neighbors	89.8	89.8	89.8	89.8
	Gaussian Naïve Bayes	62.3	63.8	61.5	61.4
Ensemble Learning	Gradient Boosting	85.3	85.3	85.3	85.3
	XGBoost	85.6	85.6	85.6	85.6
	Stacking	91.7	91.7	91.7	91.7

a notably larger classification score compared to similar studies using the same dataset and methodology. To further enhance this classification score, we propose combining both descriptors—HOG and wavelet—to extract additional features. This approach allows us to provide our classifier with enriched inputs, reducing the risk of overfitting.

We then use these extracted features to train a RandomForestRegressor model, which helps identify and select the most essential features, thereby reducing the dimensionality of our dataset and enhancing the interpretability and reliability of our model's decisions. These selected features are subsequently used as inputs for our models. Table 5 shows the prediction scores for each algorithm using this vector concatenation approach.

Based on a comprehensive analysis of our results, we confidently assert that our approach has yielded robust outcomes, outperforming numerous state-of-the-art experiments, as shown in Tab 6. This success can be attributed to the meticulous phases of data preprocessing, extraction of image texture characteristics, and the synergistic

Tab. 6. Performance comparison between our proposed method and different CNN approaches on the same dataset.

Publication	Classification method	Feature extraction methods	Accuracy
[18]	ResNet-50	CNN	95%
[35]	VGG-16	Lu-Net	90%
[10]	AlexNet	CNN	96%
[19]	FC Layer	CNN	90%
Proposed	Stacking	HOG+Wavelet	**92%**

use of two descriptors—HOG and DWT—enhanced by a stacking algorithm that integrates multiple base models. Throughout these phases, we emphasize the crucial role of computer vision and infographic functions in enriching our dataset. Additionally, the effectiveness of HOG and DWT techniques in extracting essential features, expediting model training, and preventing overfitting has significantly boosted the overall efficacy and resilience of our classification framework.

The concatenation, assessment, and selection of feature vectors to identify critical features, reduce dataset dimensionality, and enhance the interpretability and reliability of our model's decisions have further contributed to the success of our approach. Moreover, the incorporation of machine learning algorithms such as RF, SVM, and K-NN, along with the stacking technique, has streamlined the integration of their predictions, thereby improving classification accuracy. In summary, the fusion of these methods has resulted in a novel approach distinguished by superior accuracy.

4.4. Discussion of results

Prior research on early brain tumor prediction has shown varying levels of accuracy, all addressing the same fundamental challenge. The techniques used in each approach—from dataset preprocessing to feature extraction and classification—are crucial factors that highlight the value of our proposed method. In our study, we leverage common computer vision and graphics functions, including normalization, resizing, augmentation, and cropping, along with HOG and DWT feature extraction techniques, to ensure robust and informative input data for the classifier, expedite model training, and prevent overfitting.

Additionally, we combine multiple machine learning algorithms to construct a resilient model, resulting in improved classification scores. The following table presents the classification prediction precision of various methods applied to the same dataset, which originally contained 253 images before preprocessing. The significance of our study lies in utilizing advanced imaging functions and descriptors to extract relevant information, reduce model training time, and prevent overfitting. These extracted features serve as inputs for our ML classifiers, along with the integration capability of machine learning algorithms such as SVM, K-NN, and RF, which consolidate their predictions into a single model for efficient classification through stacking.

Our research introduces a novel approach to brain tumor classification, combining the advantages of computer vision functions, feature extraction through descriptors, and classification using machine learning models. Unlike prior studies that often overlook preprocessing, we prioritize this stage by incorporating common computer vision and graphics techniques. Additionally, we use HOG and DWT as image descriptors to extract relevant and significant features, ensuring high-quality input data for our model, expediting training, and reducing overfitting—thereby creating a strong foundation for further processing. This integration of preprocessing and feature extraction allows us to capture crucial patterns and nuances essential for accurate disease classification.

Finally, we have employed a stacking model that combines the outputs of three machine learning classifiers to improve classification accuracy. Our approach combines advanced feature extraction techniques and accurate prediction models to surpass previous methods in brain tumor classification. Additionally, we evaluated the model's real-time performance, achieving classification results within a 4-second time frame and a classification accuracy of 92%, demonstrating its efficiency and suitability for practical applications.

5. Conclusion and perspective

In this paper, we have introduced an ensemble learning approach for brain tumor prediction and classification. Our study focuses on two primary phases. Firstly, after preprocessing the dataset using common computer vision and graphics functions including normalization, resizing, augmentation, and cropping, we proceed to the feature extraction phase. Here, we utilize HOG and DWT descriptors to extract salient and relevant features, thereby accelerating ML model training, preventing overfitting, and bolstering the overall effectiveness and robustness of the classification framework. Subsequently, we concatenate these extracted features into a single vector, evaluate, and select them using a RandomForestRegressor, then utilize them as input in the Stacking model to classify the results.

Our approach's effectiveness lies in combining features extracted by the HOG and DWT descriptors and integrating outputs from machine learning classifiers, resulting in a more efficient classification model through stacking. This technique maximizes the predictive power of individual classifiers, enhancing the overall accuracy and efficiency of our classification framework.

The experimental outcomes highlight the remarkable capability of the HOG and

DWT descriptors in extracting crucial features from MR images. Furthermore, the effectiveness of amalgamating ML models is evident in achieving efficient classification metrics. This study holds promise for advancing computer-assisted diagnosis in digital pathology, indicating potential breakthroughs in medical imaging analysis.

Several studies in disease classification often fall short of meeting medical experts' expectations due to issues such as poor performance, data dependency, or reliance on computationally complex deep learning models. In our future endeavors, we aim to investigate alternative large-scale datasets and devise methodologies to overcome these limitations, striving to advance the field of medical image analysis and provide more reliable tools for disease diagnosis and prognosis.

References

[1] S. Abbasi and F. Tajeripour. Detection of brain tumor in 3D MRI images using local binary patterns and histogram orientation gradient. *Neurocomputing*, 219:526–535, 2016. doi:10.1016/j.neucom.2016.09.051.

[2] P. Afshar, A. Mohammadi, and K. N. Plataniotis. Brain tumor type classification via capsule networks. In: *Proc. 2018 25th IEEE International Conference on Image Processing (ICIP)*, pp. 3129–3133, 2018. doi:10.1109/ICIP.2018.8451379.

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

[4] A. Ari and D. Hanbay. Deep learning based brain tumor classifcation and detection system. *Turkish Journal of Electrical Engineering and Computer Sciences*, 26(5):9, 2018. doi:10.3906/elk-1801-8.

[5] M. Arunachalam and S. Royappan Savarimuthu. An efficient and automatic glioblastoma brain tumor detection using shift-invariant shearlet transform and neural networks. *International Journal of Imaging Systems and Technology*, 27(3):216–226, 2017. doi:10.1002/ima.22227.

[6] A. Barragán-Montero, U. Javaid, G. Valdés, D. Nguyen, P. Desbordes, et al. Artificial intelligence and machine learning for medical imaging: A technology review. *Physica Medica: European Journal of Medical Physics*, 83:242–256, 2021. doi:10.1016/j.ejmp.2021.04.016.

[7] L. Breiman. Random forests. *Machine Learning*, 45:5–32, 2001. doi:10.1023/A:1010933404324.

[8] S. Charfi, R. Lahmyed, and L. Rangarajan. A novel approach for brain tumor detection using neural network. *International Journal of Research in Engineering & Technology*, 2:93–104, 2014. `http://www.impactjournals.us/download/archives/2-77-1406198782-13.Eng-ANovelApproachforbraintumordetectionusing-SaidCharfi.pdf`.

[9] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20:273–297, 1995. doi:10.1007/BF00994018.

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

[11] S. Das, O. F. M. R. R. Aranya, and N. N. Labiba. Brain tumor classification using convolutional neural network. In: *Proc. 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, pp. 1–5, 2019. doi:10.1109/ICASERT.2019.8934603.

[12] S. Das, S. Bose, G. K. Nayak, and S. Saxena. Deep learning-based ensemble model for brain tumor segmentation using multi-parametric MR scans. *Open Computer Science*, 12(1):211–226, 2022. doi:10.1515/comp-2022-0242.

[13] S. Deepak and P. M. Ameer. Brain tumor classification using deep CNN features via transfer learning. *Computers in Biology and Medicine*, 111:103345, 2019. doi:10.1016/j.compbiomed.2019.103345.

[14] F. J. Díaz-Pernas, M. Martínez-Zarzuela, M. Antón-Rodríguez, and D. González-Ortega. A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network. *Healthcare*, 9(2):153, 2021. doi:10.3390/healthcare9020153.

[15] M. Gurbină, M. Lascu, and D. Lascu. Tumor detection and classification of MRI brain image using different wavelet transforms and Support Vector Machines. In: *Proc. 2019 42nd International Conference on Telecommunications and Signal Processing (TSP)*, pp. 505–508, 2019. doi:10.1109/TSP.2019.8769040.

[16] G. Hemanth, M. Janardhan, and L. Sujihelen. Design and implementing brain tumor detection using machine learning approach. In: *Proc. 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 1289–1294, 2019. doi:10.1109/ICOEI.2019.8862553.

[17] L. Hussain, S. Saeed, I. A. Awan, A. Idris, M. S. A. Nadeem, et al. Detecting brain tumor using machines learning techniques based on different features extracting strategies. *Current Medical Imaging*, 15(6):595–606, 2019. doi:10.2174/1573405614666180718123533.

[18] 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*, 10(1):23, 2022. doi:10.1007/s13755-022-00193-9.

[19] J. Kang, Z. Ullah, and J. Gwak. Mri-based brain tumor classification using ensemble of deep features and machine learning classifiers. *Sensors*, 21(6):2222, 2021. doi:10.3390/s21062222.

[20] K. Kaplan, Y. Kaya, M. Kuncan, and H. M. Ertunç. Brain tumor classification using modified local binary patterns (LBP) feature extraction methods. *Medical Hypotheses*, 139:109696, 2020. doi:10.1016/j.mehy.2020.109696.

[21] S. Khawaldeh, U. Pervaiz, A. Rafiq, and R. S. Alkhawaldeh. Noninvasive grading of glioma tumor using magnetic resonance imaging with convolutional neural networks. *Applied Sciences*, 8(1), 2018. doi:10.3390/app8010027.

[22] T. N. R. Kumar and N. V. Shree. Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network. *Brain Informatics*, 5(1):23–30, 2018. doi:10.1007/s40708-017-0075-5.

[23] G. Li, J. Sun, Y. Song, J. Qu, Z. Zhu, et al. Real-time classification of brain tumors in MRI images with a convolutional operator-based hidden Markov model. *Journal of Real-Time Image Processing*, 18(4):1207–1219, 2021. doi:10.1007/s11554-021-01072-4.

[24] Y. Liu, M. Muftah, T. Das, K. Robson, and D. Auer. Classification of MR tumor images based on Gabor-wavelet analysis. *Journal of Medical and Biological Engineering*, 32(1):22–28, 2011.

[25] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004. doi:10.1023/B:VISI.0000029664.99615.94.

[26] K. Maharana, S. Mondal, and B. Nemade. A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, 3(1):91–99, 2022. Proc. International Conference on Intelligent Engineering Approach (ICIEA-2022). doi:10.1016/j.gltp.2022.04.020.

[27] S. Mallat and W. Hwang. Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory*, 38(2):617–643, 1992. doi:10.1109/18.119727.

[28] R. F. Mansour, J. Escorcia-Gutierrez, M. Gamarra, V. G. Díaz, D. Gupta, et al. Artificial intelligence with big data analytics-based brain intracranial hemorrhage e-diagnosis using CT images. *Neural Computing and Applications*, 35(22):16037–16049, 2023. doi:10.1007/s00521-021-06240-y.

[29] 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*, 3(1):68–71, 2017. doi:10.1016/j.fcij.2017.12.001.

[30] N. Nabizadeh and M. Kubat. Brain tumors detection and segmentation in MR images: Gabor wavelet vs. statistical features. *Computers & Electrical Engineering*, 45:286–301, 2015. doi:10.1016/j.compeleceng.2015.02.007.

[31] N. Naik, M. Matonkar, P. Shatandel, M. A. O. Verekar, and V. Salkar. Detection and classification of brain tumor using naïve Bayes and J48. *International Journal of Science Engineering and Information Technology Research*, 9(2):19–28, 2019.

[32] M. Nickparvar. Brain tumor MRI dataset. Kaggle, 2021. doi:10.34740/KAGGLE/DSV/2645886.

[33] T. Ojala, M. Pietikäinen, and D. Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1):51–59, 1996. doi:10.1016/0031-3203(95)00067-4.

[34] J. S. Paul, A. J. Plassard, B. A. Landman, and D. Fabbri. Deep learning for brain tumor classification. In: *Proc. Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging*, vol. 10137, p. 1013710. SPIE, 2017. doi:10.1117/12.2254195.

[35] 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*, 2:100004, 2020. doi:10.1016/j.mlwa.2020.100004.

[36] P. Rajan and C. Sundar. Brain tumor detection and segmentation by intensity adjustment. *Journal of Medical Systems*, 43(8):282, 2019. doi:10.1007/s10916-019-1368-4.

[37] V. Rajinikanth, A. N. Joseph Raj, K. P. Thanaraj, and G. R. Naik. A customized VGG19 network with concatenation of deep and handcrafted features for brain tumor detection. *Applied Sciences*, 10(10):3429, 2020. doi:10.3390/app10103429.

[38] R. H. Ramdlon, E. Martiana Kusumaningtyas, and T. Karlita. Brain tumor classification using MRI images with K-Nearest Neighbor method. In: *Proc. 2019 International Electronics Symposium (IES)*, pp. 660–667, 2019. doi:10.1109/ELECSYM.2019.8901560.

[39] R. Ranjbarzadeh, A. Caputo, E. B. Tirkolaee, S. Jafarzadeh Ghoushchi, and M. Bendechache. Brain tumor segmentation of MRI images: A comprehensive review on the application of artificial intelligence tools. *Computers in Biology and Medicine*, 152:106405, 2023. doi:10.1016/j.compbiomed.2022.106405.

[40] 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*, 39(2):757–775, 2020. doi:10.1007/s00034-019-01246-3.

[41] J. Seetha and S. S. Raja. Brain tumor classification using convolutional neural networks. *Biomedical & Pharmacology Journal*, 11(3):1457–1461, 2018. doi:10.13005/bpj/1511.

[42] A. Shafi, M. B. Rahman, T. Anwar, R. S. Halder, and H. E. Kays. Classification of brain tumors and auto-immune disease using ensemble learning. *Informatics in Medicine Unlocked*, 24:100608, 2021. doi:10.1016/j.imu.2021.100608.

[43] D. Singh and K. Kaur. Classification of abnormalities in brain MRI images using GLCM, PCA and SVM. *International Journal of Engineering and Advanced Technology*, 1(6):243–248, 2012. `https://www.ijeat.org/portfolio-item/F0676081612/`.

[44] Z. Sobhaninia, N. Karimi, P. Khadivi, R. Roshandel, and S. Samavi. Brain tumor classification using medial residual encoder layers. *arXiv*, 2019. ArXiv.2011.00628. doi:10.48550/arXiv.2011.00628.

[45] Z. N. K. Swati, Q. Zhao, M. Kabir, F. Ali, Z. Ali, et al. Brain tumor classification for MR images using transfer learning and fine-tuning. *Computerized Medical Imaging and Graphics*, 75:34–46, 2019. doi:10.1016/j.compmedimag.2019.05.001.

[46] F. Tajeripour, E. Kabir, and A. Sheikhi. Fabric defect detection using modified local binary patterns. *EURASIP Journal on Advances in Signal Processing*, 2008(1):783898, 2008. doi:10.1155/2008/783898.

[47] D. P. Tian. A review on image feature extraction and representation techniques. *International Journal of Multimedia and Ubiquitous Engineering*, 8(4):385–395, 2013.

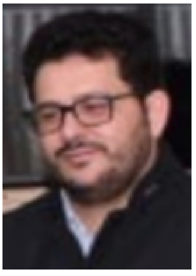
[48] S. Wang, M. Chen, Y. Li, Y. Zhang, L. Han, et al. Detection of dendritic spines using wavelet-based conditional symmetric analysis and regularized morphological shared-weight neural networks. *Computational and Mathematical Methods in Medicine*, 2015(1):454076, 2015. doi:10.1155/2015/454076.

[49] J. Yan, S. Lin, S. B. Kang, and X. Tang. Learning the change for automatic image cropping. In: *Proc. 2013 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 971–978, 2013. doi:10.1109/CVPR.2013.130.

[50] I. Zine-dine, J. Riffi, K. El Fazazi, M. A. Mahraz, and H. Tairi. Brain tumor classification using machine and transfer learning. In: *Proc. 2nd International Conference on Big Data, Modelling and Machine Learning (BML)*, pp. 566–571. INSTICC, 2022. doi:10.5220/0010762800003101.

Iliass Zine-dine: Ph.D. in Computer Science at the Laboratory of Computer Science, Signals, Automation, and Cognitivism (LISAC), Faculty of Science, Dhar El Mahraz.

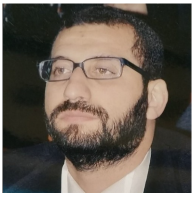
Anass Fahfouh: Doctor of Computer Science at the Laboratory of Computer Science, Signals, Automation, and Cognitivism (LISAC), Faculty of Science, Dhar El Mahraz.



Jamal Riffi: Doctor of Computer Science at the Laboratory of Computer Science, Signals, Automation, and Cognitivism (LISAC), Faculty of Science, Dhar El Mahraz.

Khalid El Fazazy: Doctor of Computer Science at the Laboratory of Computer Science, Signals, Automation, and Cognitivism (LISAC), Faculty of Science, Dhar El Mahraz.

Ismail El Batteoui: Doctor of Computer Science at the Laboratory of Computer Science, Signals, Automation, and Cognitivism (LISAC), Faculty of Science, Dhar El Mahraz.



Mohamed Adnane Mahraz: Doctor of Computer Science at the Laboratory of Computer Science, Signals, Automation, and Cognitivism (LISAC), Faculty of Science, Dhar El Mahraz.

Hamid Tairi: Doctor of Computer Science at the Laboratory of Computer Science, Signals, Automation, and Cognitivism (LISAC), Faculty of Science, Dhar El Mahraz.