

Original Article

# Explainable and Interpretable Model for Brain Tumor Classification with Optimized Transfer Learning

Sandip Desai<sup>1</sup>, Milind Mushrif<sup>2</sup>

<sup>1</sup>Department of Electronics Engineering, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India.

<sup>1,2</sup>Department of Electronics and Telecommunication Engineering, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India.

<sup>1</sup>Corresponding Author : [sad.ycce@gmail.com](mailto:sad.ycce@gmail.com)

Received: 18 February 2025

Revised: 20 March 2025

Accepted: 19 April 2025

Published: 29 April 2025

**Abstract** - Brain neoplasms represent the tenth most prevalent cause of morbidity among all oncological conditions. The manual identification of cerebral neoplasms via Magnetic Resonance Imaging (MRI) is laden with inaccuracies, as disparate radiologists may interpret identical imaging studies in divergent manners. This proposed research showcases an automated system for the identification of brain neoplasm that employs a pretrained Convolutional Neural Network (CNN) architecture and transfer learning techniques to classify brain MRI scans into four categories: glioma, meningioma, pituitary adenoma, and absence of tumor. Pretrained architectures such as ResNet50, EfficientNetB1, Xception, MobileNet, VGG19, InceptionResNetV2, and ConvNeXtLarge were utilized to extract complex features from MRI scans. The models were trained employing three distinct optimization algorithms: Stochastic Gradient Descent (SGD), ADAM, and NADAM. In this study, we implement explainable AI using Grad-CAM to enhance trust in tumor detection mechanisms by highlighting the specific regions in MRI scans that inform decision-making. Empirical findings reveal that the EfficientNetB1 architecture, when paired with the ADAM optimizer, demonstrated improved performance compared to the other models, attaining training and validation accuracies of 95.17% and 89.26%, respectively. The proposed model exhibited remarkable performance metrics, achieving an F1 score, accuracy, recall, and precision value of 100%.

**Keywords** - Brain tumor detection, EfficientNetB1, Explainable AI, Grad-CAM, Transfer Learning.

## 1. Introduction

In a fundamental context, neoplasms are categorized into two principal types: benign and malignant. Benign neoplasms are characterized by their non-cancerous nature, whereas malignant neoplasms are identified as cancerous entities. In instances of malignancy, the cancerous cells infiltrate the surrounding healthy cerebral tissues with considerable aggression, adversely affecting them and possessing the potential to metastasize to other regions of the body. Conversely, benign tumors remain localized and do not disseminate to other anatomical sites, thereby being classified as non-cancerous.

The protective membranes encompassing the brain and spinal cord are referred to as the meninges, which consist of three distinct layers of tissue. A neoplasm arising within this anatomical region is designated as a meningioma [1]. Frequently, the dimensions of this meningioma are comparable to that of a pea. Among the various types of brain tumors, this specific neoplasm poses a greater risk to pediatric and adult populations alike. In the United States, an estimated

170,000 individuals receive a diagnosis of meningioma annually.

Glioma constitutes approximately 33% of all existing brain tumors. The glioma tumor arises from the cerebral region's excessive proliferation of glial cells. These tumors are predominantly malignant in nature. They are primarily observed in the adult population. Frequently, they disseminate within the intracranial and spinal regions rather than metastasizing to other anatomical sites. Based on the various classifications of glial cells, gliomas are categorized into three principal types: glioblastoma, oligodendrogliomas, and ependymomas. Among the diverse brain tumors, glioblastoma represents a significant proportion, whereas oligodendrogliomas account for only 1% to 2%, and ependymomas contribute a mere 2%. Surgical intervention is the primary therapeutic modality recommended by medical professionals. During surgical procedures, the physician can typically excise only the visibly accessible tumor tissue. However, the delicate nature of certain brain areas renders complete resection challenging. To address the residual malignant cells, radiation therapy and chemotherapy are



advocated by clinicians as they assist in the comprehensive eradication of the affected cells. The five-year life expectancy for glioblastoma remains alarmingly low, ranging from 6% to 20% [2]. The pituitary gland is a diminutive organ, comparable in size to a pea, situated posterior to the nasal cavity and at the inferior region of the brain. Tumors that develop inside this gland are classified as pituitary tumors. Tumors in the pituitary gland may induce the overproduction or underproduction of hormones. These hormones are vital in controlling numerous physiological functions within the human body. In most instances, these types of tumors are benign in nature [3].

The presence of a brain tumor induces alterations in cerebral tissues and applies pressure on various cerebral regions. Such pressure disrupts the functional capabilities of humans. The manifestations in individuals may include cephalalgia, emesis, visual disturbances, compromised limb mobility, loss of equilibrium, speech difficulties, and cognitive impairment, among others. Should the brain tumor be classified within stages I or II, it may be amenable to therapeutic intervention; conversely, a tumor in stages III or IV is likely to culminate in mortality.

Brain tumors exhibit rapid proliferation within the cerebral region. Statistical data from the United States indicates that approximately 700,000 individuals have been diagnosed with primary brain tumors. In the year 2023, a total of 94,390 individuals received a diagnosis of brain tumors. Furthermore, 18,990 individuals succumbed to brain tumors in the same year. Brain cancer ranks as the tenth leading cause of morbidity among all cancer types [4]. This underscores the necessity for timely detection and precise classification of brain neoplasms.

Magnetic Resonance Imaging, commonly referred to as MRI, employs powerful magnets in conjunction with computer technology to generate images of the internal anatomical structures. A magnetic contrast agent is administered to the patient to enhance the visualization of anomalies within cerebral tissues. MRI represents the most valuable tool for diagnosis among imaging methods since it shows an exceptional ability to pinpoint tiny brain tumors.

MRI holds a greater preference than other imaging techniques in brain neoplasm diagnosis because its superior resolution surpasses CT scan images [5]. MRI scans reveal vital information about tumor anatomy in addition to tumor dimension measures and important characteristics. The process of identifying brain tumors manually in MRI images produces multiple possible inaccuracies. Healthcare professionals define how well brain tumors are identified through their experience and diagnostic skills. Several radiologists could probe different findings when working with identical MRI images. Utilizing the hands to specify tumor regions on MRI images demands prolonged dedication and

complex work procedures. When medical practitioners mistype the size of tiny tumors, it can lead to delayed medical treatment. The manual assessment of tumors faces high sensitivity to human mistakes due to differences in tumor sizes, shapes, and positions. The late diagnosis of brain tumors results in problematic treatment arrangements, which leads to deteriorated medical outcomes for patients. Automated brain tumor detection systems prove essential in such scenarios based on the requirement for such methodologies. Computer-Aided Detection (CAD) systems enable faster and more accurate tumor detection through their provided benefits. Through its diagnostic powers, CAD successfully detects tiny malignant masses that doctors would miss when conducting manual assessments. The detection system enables healthcare personnel to develop prompt treatment approaches for their patients [6, 7].

This manuscript advocates for forming an innovative computer-assisted automated deep learning framework designed to analyse brain MRI scans, facilitating timely diagnostic insights to enhance the efficacy of brain tumor detection.

The principal contributions of the presented investigation are delineated as follows:

1. A groundbreaking deep learning-oriented framework is introduced that leverages cutting-edge deep learning architectures, including ResNet50, EfficientNetB1, Xception, MobileNet, VGG19, InceptionResNetV2, and ConvNeXtLarge, specifically targeting MRI scans of brain neoplasms while implementing transfer learning methodologies on the dataset.
2. An initial four-phase image quality improvement protocol is adopted to ameliorate the suboptimal perceptual quality of the MRI scans.
3. A data augmentation methodology is utilized to produce superior results on limited datasets, thereby mitigating the challenges associated with overfitting.
4. Three distinct optimization algorithms, namely SGD, ADAM, and NADAM, are employed to augment classification efficacy.
5. The proposed framework has been evaluated based on an array of key performance evaluators, including F1-score, accuracy, recall and precision.
6. The suggested methodology aims to focus on particular regions to facilitate determinations concerning the identification of brain tumors by implementing the Grad CAM technique.

## 2. Literature Review

The proposed method employs a tailored GoogleNet deep learning network that evaluates brain tumor MRI images to extract vital attributes with deep Convolutional Neural Network (CNN) features. This set of attributes serves as a critical tool to classify brain tumors into their respective three diagnostic groups correctly. Researchers used softmax

classifiers as well as Support Vector Machines (SVM) and K-Nearest Neighbors (KNN) algorithms to assess these features because they work with MRI-derived data to increase classification precision [8]. The accuracy rate reached 92.3% through deep transfer learning in the experiments, yet integration of SVM and KNN with deep CNN features achieved 97.8% and 98.0% accuracy successively. Small training data sizes introduce overfitting problems that lead to the incorrect classification of meningioma entities within the model.

Two brain tumor classification tasks operate using the proposed neural network system, which handles glioma grading and brain tumor broad classification together. The designed CNN procedure operates through 16 layers, working with augmented images to obtain feature information. The network architecture implements convolutional layers for pattern recognition through filters and activation functions with Rectified Linear Unit (ReLU) to accelerate training together with pooling layers that reduce data dimensions [9]. During training, the application drops specific neurons while implementing dropout layers to reduce the chance of overfitting. The fully connected layer as the terminal element combines softmax function computation to calculate probabilities for different classes. The researchers achieved favorable classification results of 96.13% and 98.7% with data augmentation techniques implemented to expand the training dataset despite using a confined dataset. The suggested CNN required training for a different compelling reason because of its intended application in various medical domains. The advanced variant of Speeded Up Robust Features (SURF) is known as DSURF, which enables image processing to detect vital points within images. Through HOG methodology, image segmentation forms cells that help detect meaningful patterns by computing multi-angle edge orientations [10]. Researchers are developing a method to categorise cerebral neoplasms using support vector machines with multiple image processing techniques.

DWT enables the fusion of MRI T1, T1C, T2, and Flair sequences into one unified image, strengthening details by collecting various textures and structures from multiple sequences [11]. The process of separating images into different frequency bands displays an enhanced representation of tumor areas better than individual MRI images can show. PDDF achieves lesion enhancement via its noise reduction features that enhance the diagnostic quality of lesion images. The global thresholding method sorts pixels into two separate categories: foreground identification representing the lesion and background designation that depends on their intensity measurement values. The research endeavors to fuse MRI images exclusively since it does not explore integrating CT or PET scans with either the MRI images or other modalities.

Researchers developed Caps-VGGNet by integrating CapsNet and VGGNet to achieve exceptional performance in

classifying cerebral tumors into Normal, Pituitary, Meningioma, and Glioma groups [12]. A notable limitation inherent in this approach pertains to the complexity of the model.

According to [13], VSBEAM stands as a specialized deep learning model that detects and locates brain tumors in MRI images through its Voting Based Semi-Bayesian Ensemble Attention Mechanism architecture. The framework contains three core modules, which start with a “squeeze and excitation” step to control the value of each data feature. The second phase selects Bayesian learning because its prediction capabilities excel under the condition of sparse datasets and outlier data points. Ensemble learning combines multiple model predictions into a unified accurate classification using a voting system that uses collective prediction results. The contourlet transform performs feature extraction by combining wavelet transform advantages with directional filtering power to detect significant image characteristics. This detection method proves excellent at finding small elements and spatial orientations to advance the processing of images and recognition of objects. The extraction of statistical features from different frequency components through the contourlet transform produces more precise brain tumor detections in MRI imaging [14].

SSBTCNet utilizes a framework that enables the solution of two separate problems, starting from the data domain and proceeding to tumor classification, by integrating supervised and unsupervised learning concepts. The classification system achieves optimization thanks to a training methodology that mixes labeled and unlabeled data while using fuzzy logic and image enhancement techniques [15].

The Discrete Wavelet Transform (DWT) enables the conversion of images to preserve vital details while making computational operations faster. The four separate data bands (LL, LH, HL, HH) produced by DWT allow for feature extraction, such as contrast and energy. Applying Gabor filters produces texture elements, resulting in a new feature set featuring 220 elements that enhance detection precision [16].

Multiple brain tumor features are computed by the Gray Level Co-occurrence Matrix (GLCM) through its pixel intensity relationship analysis to enable classification by analyzing contrast, correlation and energy metrics. The team developed 48 specific features because they applied four angular displacements to create these analytical features, which advanced their detection performance. An image brightness assessment through histograms generates quantitative frequency data about pixel brightness spread through intensity-based feature extraction. Statistical indicators from the histogram analysis include the mean brightness measurement, standard deviation, and three dimensions: skewness, kurtosis, and entropy. Medical experts classify tumor documents based on vital form parameters,

including circularity, bending energy, and rectangularity, which assist MRI image tumor morphology analysis. The text explains calculating particular statistical moments by measuring distances between object boundary points and their central point. Such moments generate features to analyze tumor shape characteristics with roughness attributes. A selection of vital features obtained from the CVM method serves to train three machine learning models, including KNN and, mSVM and NN.

The author employs IFF-FLICM-based fuzzy segmentation combined with an optimized extreme learning machine model controlled by the MHS-SCA hybrid algorithm for precise tumor identification in MRI images. Proof shows that the proposed method achieves exceptional performance in detecting brain tumors, which makes it practical for clinical medicine used in early medical diagnosis.

Tumor segmentation in MRI images depends on thresholding methods and histogram-based techniques according to the approach described in [17]. The framework implements the Gray Level Co-occurrence Matrix (GLCM) to produce multiple statistical and textural image features through contrast evaluation, correlation counting, and energy measurements. The adopted features enable precise image representation using limited CPU power during the evaluation process.

The MRI image feature extraction process implemented GoogLeNet while SVM and KNN operated as pre-classifiers before implementing a fully connected layer that included softmax as the final decision-making component. The scientists redesigned the end portion of GoogLeNet to perform brain tumor image identification across three essential categories to support MR imaging data handling. Combining GoogLeNet with SVM or KNN exhibits superior performance [18].

The image classification field contains Vision Transformer (ViT) models originating from research methods originally designed for Natural Language Processing. The IVX16 approach unifies predictions obtained from VGG16, InceptionV3, and Xception because these networks demonstrated exceptional performance during previous studies. The Swin Transformer, together with the Compact Convolutional Transformer (CCT), strengthens performance by using shifted windows along with convolutional operations to enhance local image feature extraction. The authors established that IVX16 yielded superior performance when detecting tumor areas through their analysis using the LIME explainable artificial intelligence tool. The feature extraction process for MRI images utilized GoogLeNet while SVM, together with KNN, served as classifiers before implementing a fully connected layer with softmax as the last decision-making step. The researchers modified GoogLeNet's final segment to handle brain tumor image classification into three

specific categories, thus making it suitable for MRI data analysis. Combining GoogLeNet with SVM or KNN exhibits superior performance [18].

The Vision Transformer (ViT) models derive from original Natural Language Processing techniques to classify images. The IVX16 model functions by combining prediction results from VGG16, InceptionV3 and Xception architectures, which proved effective during previous evaluations, according to [19]. The Swin Transformer, together with the Compact Convolutional Transformer (CCT), optimizes image feature detection through shifted window approaches combined with convolutional operations, which leads to improved performance in identifying objects within the image frame. LIME explained that IVX16 achieved better tumor region recognition than other established models, which the authors confirmed through their research.

Active contour models serve as a methodology operating under the "snake" name to identify objects in processed images that include tumors in MRI scans. The snake method works by placing flexible curves around target areas before optimizing their energy value, enabling these curves to align with object edges. Internal energy, together with external energy are calculated independently to achieve snake smoothness and object edge attraction in the methodology. When these regions are located through identification, the specialized decision-making framework uses texture analysis to determine if tumors exist in these areas [20].

This investigation delves into an innovative model referred to as a triplet-based variational autoencoder, which is designed to discern patterns inherent in healthy brain MRI data while simultaneously mitigating noise present in the images. The researchers tackle a prevalent error in antecedent methodologies that presumed that healthy segments of the images would exhibit uniformity post-reconstruction. This notion is not invariably accurate in the presence of lesions.

The anomaly detection protocol delineated involves a training stage where exclusively healthy brain images are scrutinized to delineate the characteristics of a normal brain. During this training phase, noise is intentionally introduced into these images to simulate anomalies, thereby aiding the network in recognizing and reconstructing healthy brain slices even amidst the noise. In the subsequent testing phase, the network conducts a comparative analysis between the original images and their reconstructed counterparts to identify discrepancies, which may signify potential anomalies. It subsequently employs filtering techniques to enhance the results and eliminate any minor inaccuracies [21].

The authors introduce an innovative methodology termed the Two-Stage Generative Model (TSGM) to augment the detection and segmentation of cerebral neoplasms within magnetic resonance imaging (MRI) scans. This approach

integrates two sophisticated techniques: the Cycle Generative Adversarial Network (CycleGAN), which facilitates the generation of pathological images from normative counterparts, and a Variance Exploding stochastic differential equation (VE-JP) that assists in the reconstruction of healthy images while accentuating tumor regions[22]. Stage 1 of the framework entails the employment of a CycleGAN, a variant of artificial intelligence architecture, to produce paired images representing both normal (healthy) and abnormal (unhealthy) specimens.

This is accomplished through an educational process that enables the model to transmute images between categories via two distinct mapping functions: one that alters normal images into abnormal ones and another that performs the inverse transformation. Through the optimization of specific loss functions, the CycleGAN guarantees that the generated images exhibit visual similarities to authentic images within their respective classifications while preserving salient features. Stage 2 of the framework is concentrated on the generation of healthy cerebral images employing a conditional diffusion model.

This phase commences with data derived from Stage 1, wherein the model discerns the interrelationship between healthy images and their corresponding abnormal counterparts. By introducing stochastic noise to the healthy images and progressively transforming them into a Gaussian distribution, the model is subsequently capable of reversing this process to retrieve healthy images from the noisy data while being guided by the abnormal images.

The Multi-Modality MRI Ensemble process amalgamates images obtained from various MRI modalities to enhance the detection of brain tumors. It employs a computational formula to derive a new image by calculating the weighted differences between the original and reconstructed images for each modality, ensuring that the cumulative weights total one. This methodology facilitates the creation of an anomaly heatmap, which delineates areas of significance and is subsequently utilized to produce a segmentation mask that identifies tumor regions within the cerebral structure.

The author of [28] developed a specialized CNN model for brain tumor identification. The authors applied three explainable AI approaches to reveal the basis for CNN's decision-making. The system performs brain tumor classification between tumors and non-tumorous conditions. The proposed system uses pre-trained models combined with four-class glioma, pituitary, and meningioma and no tumor classification, which eliminates the need for extensive computation resources to train an entire model from scratch.

The study in [35] applied both DenseNet121 and InceptionNetV3 networks for performing feature extraction operations. Through voting procedures, the tumor gets

categorized into three categories known as glioma, meningioma or pituitary. The proposed model demonstrates improved results in terms of accuracy, recall, precision, and F1 score.

Our analysis included the selection of the most suitable optimizer and pre-trained model combination for brain tumor classification.

### 3. Methodology

#### 3.1. Dataset

Table 1. Dataset details

Total MRI scan	Brain Tumor Type	Number of MRI Scan
3285	Glioma	931
	Meningioma	942
	Pituitary	906
	No tumor	No tumor
<b>Total MRI scan</b>		<b>3285</b>

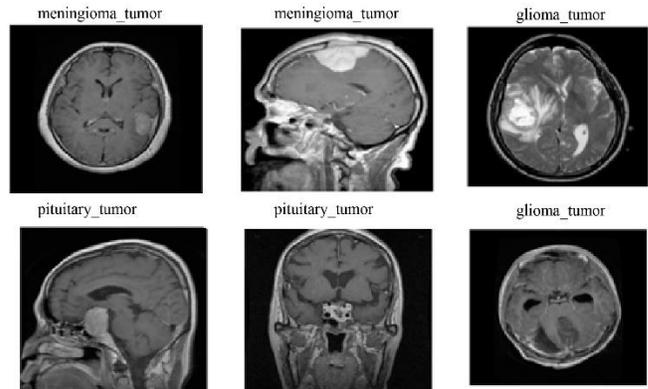


Fig. 1 Sample MR images of Brain tumor

The dataset utilized in this study has been procured from the Github repository. This dataset incorporates a total of 3285 MRI images. Each image within the dataset is categorized into one of four tumor classifications, specifically meningioma, glioma, pituitary, and the absence of a tumor. The proportional distribution of images across these four categories is presented in Table 1.

An examination of Table 1 clearly indicates that the first three tumor classifications possess approximately an equal number of images. Conversely, the no tumor category exhibits a comparatively lower number of samples relative to the other classifications, resulting in an imbalance within the dataset. Such class imbalance leads to the potential neglect of the no tumor category. Figure 1 illustrates magnetic resonance imaging samples across all four classifications present within the dataset. Figure 2 illustrates the schematic representation of the brain tumor detection system which is proposed in this research. Algorithm 1 presents a detailed explanation of the proposed model. The subsequent block delineates each process with meticulous detail.

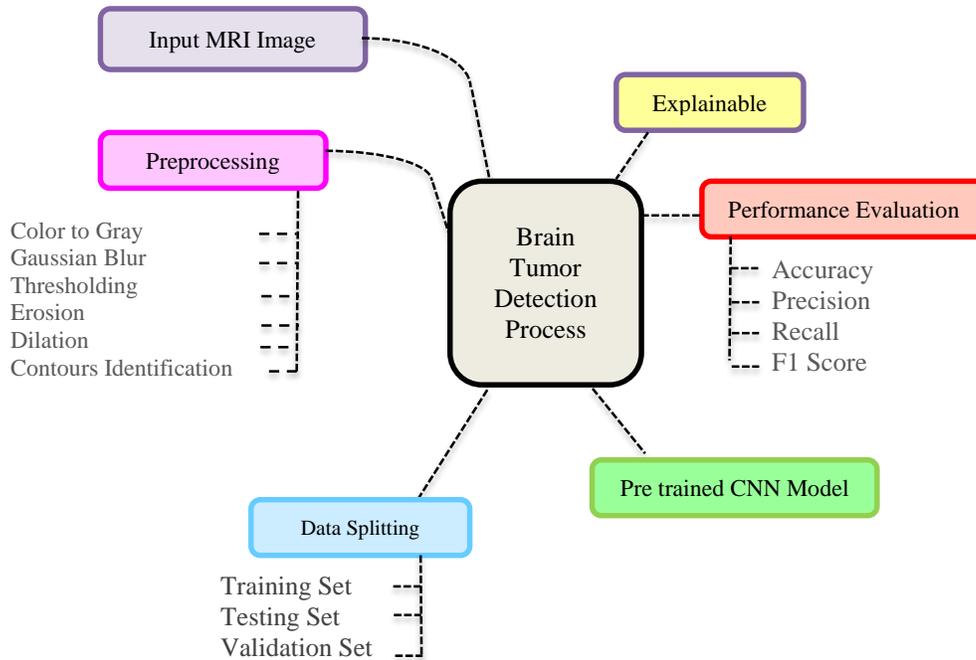


Fig. 2 Block diagram of brain tumor detection

Algorithm 1. Brain Tumor Classification with Optimized Transfer Learning

Input: Brain Tumor Dataset, Epoch - 50, Batch size - 32, Learning rate - 0.001

Step 1: Apply pre-processing to reduce noise and improve image quality.

Step 2: Apply data augmentation like rotation, flipping, Gaussian noise addition, etc.

For the model in pre-trained models(ResNet50, EfficientNetB1, MobileNet small, VGG19, ConvNeXtLarge, Xception and InceptionResNetV2)

Step 3: Load the pre-trained model with imagenet weights.

Step 4: Remove the classification layer.

Step 5: Add a new classification layer to classify Brain tumor in four classes.

For the optimizer (Adam, Nadam and SGD)

Step 6: For epoch 1 to 50, do

Step 7: Train the customised classification layer on the Brain Tumor dataset.

Step 8: Evaluate the model accuracy

If there have been no improvements in the past 5 iterations, do

Step 9: Reduce the learning rate by 0.3

End if

End for

End for

Step 10: Compare all trained models based on performance metrics and choose the best prediction model.

### 3.2. Image Pre-Processing

A classification model loses its performance capability substantially when MRI images contain noise. The presence of noise within the image generates erroneous information, thereby hindering the processes of feature extraction and generalization. A substantial decline in classification accuracy is observed when such noisy MRI images are introduced as input to the classification system. The adverse effects of noise can be alleviated by implementing noise-reduction techniques

specifically tailored for MRI images. To achieve noise reduction, the Gaussian Blur method is utilized on the RI images, which facilitates the smoothing of the image and the reduction of noise. This technique employs a Gaussian bell-shaped curve function on the image, which allocates greater weight to pixels that are in proximity to the center while assigning lesser weight to those that are more distant. Consequently, this leads to a phenomenon known as “blurring,” wherein the sharp intensity transitions of the image

are smoothed without compromising the underlying structural integrity.

According to the Gaussian principle, pixels receive their calculated values through a weighted average computation of adjacent pixels. The determination of weights via the Gaussian function is contingent upon the distance of each pixel from its center. The Gaussian function yields weights in a matrix format, which is referred to as the Gaussian kernel. The Gaussian function is mathematically expressed as

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Where,

$G(x, y)$ : Assigned weight to a pixel at a distance  $(x, y)$ .

$x, y$ : Distance of pixel from the centre of the kernel.

$\sigma$ : Standard deviation of the Gaussian distribution.

The detection algorithm must prioritize the object over the background for optimal identification of brain tumours. Thresholding techniques are employed to isolate the background from the object within an image effectively. The gray scale image undergoes conversion into a binary format by applying thresholding.

Various thresholding methodologies are at one's disposal, including binary thresholding, global thresholding, Otsu thresholding, adaptive thresholding, and inverse thresholding. Due to the inherent simplicity of binary thresholding, the present study advocates for its utilization.

Binary thresholding categorizes pixels into two distinct classifications based on pixel intensity. If the pixel intensity exceeds the designated threshold  $T$ , then a value of 255 is assigned; conversely, a value of 0 is allocated to pixels that do not meet this criterion. The determination of the threshold value is left to the programmer's discretion. This relationship can be mathematically articulated as

$$Output(x, y) = \begin{cases} 255; & I(x, y) > T \\ 0; & I(x, y) \leq T \end{cases}$$

### 3.2.1. Erosion and Dilation

The extraneous pixel located at the periphery of the image diminishes the dimensions of the object; consequently, erosion eliminates these pixels. It constitutes a morphological operation wherein a kernel is convolved with the image to yield a refined image representation.

Another prevalent morphological operation that complements erosion is dilation. Dilation introduces additional pixels to the foreground's boundary to augment an object's dimensions within an image. The outcome of dilation manifests as an expansion in the foreground pixels.

### 3.2.2. Contours

In the context of object shape analysis, the identification of object contours proves to be highly beneficial. The curve that connects all successive points along the boundary exhibiting uniform intensity is referred to as a contour. This process entails the recognition and organization of boundary pixels in accordance with the object's shape. Contour computation yields optimal results when applied to a binary image, which is typically generated during the thresholding phase.

### 3.2.3. Image Augmentation

Images can exhibit variations in numerous aspects, such as pose, occlusion, scale, and lighting. Because of these differences, computer vision models face extraction challenges when determining generalized image features. When the amount of data available for training a deep learning model is limited, the model often focuses on learning particular characteristics instead of broader patterns. This issue is known as overfitting in machine learning. In the case of overfitting, the model demonstrates satisfactory performance during the training phase, yet its performance deteriorates during the testing phase.

Augmentation introduces variability into the dataset, compelling the deep learning model to acquire a more robust representation. It produces modified versions of the existing images. This technique is particularly advantageous when the dataset is limited in size.

#### Rotation

The application of rotational transformation to training images generates new test images. A random rotation of a specific angle applies to images during the process, which generates variant images. The executed Figure 3 shows that a 30-degree rotation took place after implementing rotation augmentation on the sample MRI image.

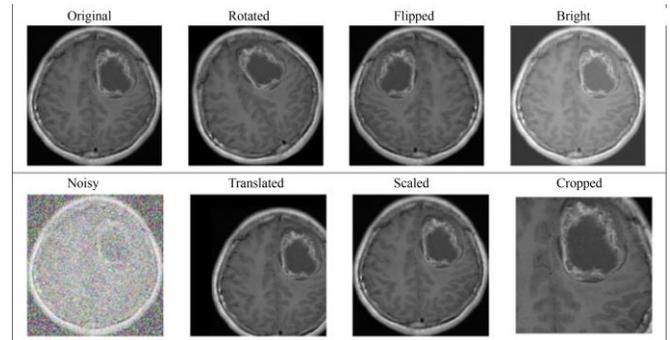


Fig. 3 Effect of augmentation on brain MRI image

#### Flipping

Flipping is an important augmentation practice that focuses on image flipping. Mirroring operations on images occur after selecting vertical or horizontal axis directions. A horizontal flipping operation affects this MRI image to move

the tumor from its original right position to the opposing left hemisphere of the brain.

#### *Brightness*

The training of computer vision machine learning models uses brightness augmentation as a widely recognized technique. Adjusting image brightness through this method improves the operational strength of the model. The procedure performs multiplication on pixel values using specified range limiting factors. The displayed sample MRI image's data matches the brightness modifications caused by brightness augmentation versus the original image.

#### *Noisy*

In real-world scenarios, it is feasible for the pre-processing techniques to introduce noise into MRI scans. To replicate this scenario, noise augmentation is utilized. This augmentation method involves adding a specific type of noise to an image, including Gaussian noise, salt and pepper noise, or speckle noise, among others. Figure 3 depicts the resultant MRI image following applying Gaussian noise augmentation.

#### *Translation*

Translation constitutes one of the various augmentation methods computer vision uses for model generalization enhancement. The content of an image under this augmentation technique is relocated vertically and horizontally while preserving all other elements of the image. The desire to pad or crop image pixels comes into effect when they surpass original boundary limits. During pixel displacement according to chosen locations. The ramifications of translation augmentation are observable in Figure 3, where the MRI image has been shifted to the right, with the pixels moving outside the boundary being cropped.

#### *Scaling*

Scaling augmentation is a preferred strategy among practitioners to foster diversity within the dataset. This augmentation pertains to the process of resizing an image by applying scaling transformations, which may either reduce or expand its dimensions. Figure 3 demonstrates the results of implementing scaling augmentation on a sample MRI image.

#### *Cropping*

Cropping augmentation is employed to artificially enlarge the dataset size. This technique involves extracting a smaller image segment and designating it as a new constituent of the expanded dataset. The various cropping options available encompass random cropping, scale-aware cropping, center cropping, and region-specific cropping, among others. Figure 3 illustrates the effects of random cropping applied to a sample MRI image.

Figure 3 illustrates the impact of diverse augmentation methodologies on the brain MRI dataset.

### **3.3. EfficientNet**

The conventional deep learning paradigm is impeded by the significant burden of resource consumption necessary for optimal model efficacy. The innovative aspect of EfficientNet lies in its implementation of compound scaling, which facilitates an enhanced equilibrium between model efficacy and resource allocation. In this methodology, the dimensions of height, width, and depth of the model are scaled uniformly.

The number of channels existing in each network layer defines the width within neural networks. An increased number of channels across every layer enables the capture of intricate patterns and features within an image, thereby enhancing classification accuracy. Conversely, depth within a neural network pertains to the network's total layers. A deeper architecture can capture more advanced representations, albeit at the cost of increased computational demands. Resolution pertains to the scaling of image dimensions. For optimal model performance, higher-resolution images are critical as they encapsulate detailed information, thus necessitating substantial memory and computational resources. The principal contribution of the EfficientNet architecture resides in its capacity to determine these three dimensions through a principled approach. The optimal depth, width, and resolution synthesis is ascertained using a grid search algorithm. The uniform scaling of these parameters is executed through a grid search guided by a user-defined compound coefficient " $\phi$ ". This compound coefficient uniformly scales all three dimensions. To optimize computational efficiency, the baseline model is employed as an initial reference. This baseline model constitutes a moderately sized neural network that exhibits satisfactory performance for the specified task yet requires considerable computational power. This situation engenders the necessity for a computationally efficient model architecture.

The user-defined parameter, the compound coefficient, is introduced to determine the scaling across each network dimension, aimed at enhancing computational efficiency. The scaling decisions pertaining to the model's height, width, and depth are made based on a suitable value of a singular scalar, identified as the compound coefficient. The overall complexity of the model and its corresponding computational resource requirements are modulated through adjustments to the compound coefficient.

In EfficientNet, depth signifies the scaling factor for the network's depth, governed by a constant  $\alpha$ . Width denotes the proportional scaling factor for the network's width, governed by a constant  $\beta$ . Resolution is scaled by multiplying the original image size by a constant  $\gamma$ . The optimal scaling configuration is contingent upon the optimal values of the constants  $\alpha$ ,  $\beta$ , and  $\gamma$ . These optimal values are derived through the grid search optimization process, which ensures the most favorable trade-off between model efficiency and computational resource requirements [24].

### 3.4. Transfer Learning

Transfer learning represents a methodology within Convolutional Neural Networks (CNNs) that leverages the knowledge acquired from a pre-trained network on an extensive dataset to address a new but related task. This approach retains the pre-trained network in its original form while the fully connected layer is replaced by the new fully connected layer specifically designed for the target task [25]. Essentially, CNNs are employed to extract features from an image, with the earlier layers capturing lower-level features and the subsequent layers capturing more complex features. Transfer learning affords the opportunity to unfreeze several of the terminal layers for retraining, thereby enabling the network to learn task-specific features [25].

### 3.5. Optimizers

The training process of Convolutional Neural Networks (CNNs) requires optimizers to accomplish their essential task of delivering complete training capabilities to these specialized artificial neural networks dedicated to image analysis. Extreme parametric adjustments through optimizers serve to minimize loss functions while boosting complete model performance metrics. Furthermore, optimizers are instrumental in regulating the manner in which the weights of the convolutional neural network are systematically updated throughout the training phases. Optimizers achieve effective learning capability by making proper weight and bias modifications according to data patterns. The model becomes unable to enhance its predictive skills when there is no optimizer; thus, its performance remains stagnant. The model learns images efficiently through optimal weight updates performed by optimizers to achieve a methodologically sound adjustment of weights and biases.

The selection of optimizers is a critical factor that significantly impacts various aspects of the performance, convergence speed, and generalization capability of convolutional neural networks, thereby underscoring the importance of careful consideration in their selection process.

#### 3.5.1. ADAM Optimizer

The Adaptive Moment Estimation approach, known as ADAM, represents one of the most frequently selected optimizers within deep learning research because it shows strong robustness and implements effective optimizations across multiple applications together with its high processing performance. Each ADAM parameter learning rate modification occurs independently based on the gradient magnitude and historical value calculations that enhance gradient sparsity and improve convergence speed. ADAM combines fundamental features of momentum acceleration with RMSProp adaptive learning rate adjustment that uses recent gradient averages. ADAM enhances its performance by implementing a bias correction system that maintains unbiased first and second-moment estimates for the first stages of training. ADAM has achieved widespread popularity

in deep learning because it reaches high convergence speeds compared to alternative optimization methods.

#### 3.5.2. SGD Optimizer

SGD is one of the basic yet frequently used optimization algorithms for training Convolutional Neural Networks (CNNs). The optimizer achieves decreasing loss through stepwise parameter adjustments that stem from computed gradients in an automated fashion. At each iteration, SGD calculates the model parameters' gradient from individual data points or small batches before performing computations to achieve a better training response.

#### 3.5.3. NADAM Optimizer

Nesterov-accelerated Adaptive Moment Estimation, commonly abbreviated as Nadam, represents an advanced optimization approach that builds upon the foundational principles established by the Adam optimizer while incorporating innovative ideas derived from Nesterov Accelerated Gradient (NAG). Nadam effectively merges the adaptive learning rate mechanism characteristic of Adam with the forward-looking gradient update strategy associated with NAG, which, in many instances, leads to a more rapid and stable convergence during the training process. By predicting the forthcoming position of the parameters and subsequently computing the gradient at that anticipated location, Nadam facilitates more informed and strategic updates to the parameters. Additional capabilities of Nadam involve an automatic parameter learning rate adjustment based on gradient magnitude, which consequently improves the optimization method's effectiveness [26].

### 3.6. Explainable AI

Prior to the advancement of explainable AI, deep learning models were commonly perceived as "black boxes." Convolutional neural network models lack clarity in elucidating the rationale behind the generated outputs. In the domain of healthcare, the predominant concern is "trust," which encompasses the necessity for decisions to be accurate, equitable, and dependable. When the decision-making process is rendered transparent, artificial intelligence can cultivate trust across various sectors.

Gradient-weighted Class Activation Mapping, or Grad-CAM, is an analytical technique employed to discern the regions a convolutional neural network emphasizes when predicting a specific class. Utilizing a class discriminative localization strategy, Grad-CAM produces visual explanations without altering the model architecture, thereby emphasizing critical segments of the input image that are pivotal for the class prediction.

A limitation of the class activation map is its requirement to incorporate a Global Average Pooling (GAP) layer before the final fully connected classification layer. This imposes additional structural demands on a convolutional network

aimed at enhancing explainability. Such modifications to the original architecture necessitate the retraining of the model to accommodate the newly integrated layers. The Grad-CAM method combines CAM methodology with gradients extracted from the final convolutional layer to enhance the target class interpretation process. During the training process, an image traverses through a CNN incrementally extracts semantic information pertinent to a designated object. Within a CNN framework, the final convolutional layer invariably contains the most significant semantic information while preserving spatial details, which is effectively utilized in Grad-CAM.

The gradient serves as a critical element in interpreting the output generated by the CNN model. Analysing the gradient yields insights into the most influential features present within the input image, which are used to predict a specific class. The calculation of the predicted class gradient aims to determine the relationship between the last convolutional layer feature map and the predicted output. The relationship facilitating this gradient computation is expressed as follows.

$$\frac{\partial y^T}{\partial A^{<k>}}$$

Where  $y^c$  represents the output class and  $A^{<k>}$  is the  $k$ th feature map. The last convolutional layer in any CNN architecture is responsible for extracting high-level features and encompasses information regarding the key components of the input image. This allows for understanding the weightage assigned to each feature map in relation to the predicted outcome.

The gradient consistently conveys information pertaining to the variation in the output class in relation to each feature map. It has been observed that certain feature maps exert a more substantial influence on the output class compared to others. To ascertain the influence of a feature map, a score must be calculated utilizing the global average pooling of the feature map [27].

$$\alpha_k^T = \frac{1}{Z} \sum_i \sum_j \frac{\partial y_T}{\partial A_{i,j}^{<k>}}$$

$$Z = W \times H$$

Where  $\alpha_k^T$  is the score calculated from the  $k$ th feature map.

A higher positive score implies a greater effect on the output class. These scores are referred to as weights. Through the utilization of these scores, a Grad-CAM heatmap is generated.

The weighted sum of the  $k$  feature map in  $A$  is computed using.

$$I = \sum_{k=1}^K \alpha_k^T A^{<k>}$$

The final heatmap is generated by applying element-wise ReLU.

$$L_{\text{GradCAM}}^T = \text{RELU}(I)$$

### 3.7. Experimental Setup

This scholarly investigation meticulously employed a diverse array of sophisticated and highly regarded libraries, including but not limited to Pandas, TensorFlow, Numpy, and Keras, in order to carry out a comprehensive analysis that is both rigorous and methodologically sound. The proposed innovative methodology utilized sophisticated features of Google Colab to develop an effective system. The evaluation of system effectiveness and overall performance was carried out through simulations that utilized devices with high processing power that contained a Core i5 Central Processing Unit to verify result accuracy.

**Table 2. Optimal environmental conditions incorporated in the model**

Configuration	Value	Hardware	Software
Training Set	80%	Memory Capacity – 51GB	Python - 3.10.12
Testing Set	20%	GPU – T4	Keras – 3.5.0
Batch Size	32		Tensor Flow – 2.17.1
Activation	Adam		Matplotlib - 3.8
Epochs	50		Numpy – 1.26.4
Learning Rate	0.0001		
Steps per epoch	82		

Table 2 serves as a detailed presentation of the specific platform that was utilized in the intricate process of detecting brain tumors, highlighting the technological foundation upon which the research is built. The aforementioned libraries and computational platforms collectively provided a resilient and high-performing structure, which was essential for deploying the proposed model and carrying out various experiments to thoroughly assess its performance in the complex tasks of detecting and classifying brain tumors with high accuracy.

The Adam optimizer, which is well-regarded in the field of machine learning for its adaptability and efficiency, was employed to significantly improve the performance of the proposed approach, initiating the optimization process with a carefully selected learning rate set at  $1e-4$  to ensure optimal convergence. As the performance metrics of the model reached a state plateau, a learning rate reduction function was subsequently applied to facilitate better convergence

outcomes, thereby improving the model's effectiveness in gaining knowledge from data. The training of the models was conducted over a total of 50 epochs, utilizing a batch size of 32, while incorporating L2-regularization techniques and an early stopping mechanism that would activate in the event that the model's performance showed no signs of improvement.

The model's hyperparameter optimization process that determines its performance quality is displayed in Table 3 through detailed insights about parameter adjustments throughout training. The experimental procedure incorporates an early stopping criterion, which halts the training of the Convolutional Neural Network (CNN) model in the event that the validation accuracy does not exhibit a substantial improvement.

The technique applies to transfer learning by replacing the terminal fully connected layer of a pre-trained model with our custom fully connected layer, succeeded by a Global Average Pooling (GAP) layer and concluding with a fully connected layer utilizing a softmax activation function comprising four neurons. To facilitate the extraction of task-specific features, we have unfrozen the final twenty layers of the pre-trained architecture.

### 3.8. Performance Evaluation Metric

#### 3.8.1. Accuracy

This particular measurement can be defined as the proportion or ratio of the number of accurate predictions made by the predictive model in question to the total number of predictions that the model has generated overall, thereby providing an in-depth evaluation of the model's overall predictive capabilities.

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

#### 3.8.2. Precision

This specific evaluative metric places a significant emphasis on the correct identification of positive predictions and is mathematically represented as the ratio of the number of rightfully identified positive instances to the total number of positive predictions made by the model, serving as an essential measure of the model's capability to minimize false positives.

$$\text{Precision} = \frac{\text{Number of Correct positive predictions}}{\text{Total number of positive predictions}}$$

It is widely regarded as a crucial metric, particularly in scenarios involving the potential for incorrect positive predictions, which holds paramount importance and relevance in the context of the medical field, where the implications of such errors can have significant consequences for patient outcomes and treatment efficacy.

#### 3.8.3. Recall

The ratio between accurate positive predictions made by the model and actual positive samples shows the model's ability to recognize all important examples of the positive class.

$$\text{Recall} = \frac{\text{Correct positive predictions}}{\text{Actual positive}}$$

#### 3.8.4. F1-Score

The evaluation method combines both precision and recall through harmonic means to achieve balanced results that perform well under skewed class distribution situations.

$$\text{F1 - Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

## 4. Results

The findings of this investigation elucidate that the utilization of transfer learning-based deep learning architectures for the diagnosis of brain tumor patients via MRI has demonstrated efficacy in accurately categorizing individuals into four discrete classifications: glioma tumor, meningioma tumor, pituitary tumor, and absence of tumor. The employment of a variety of pre-trained neural networks, encompassing ResNet50, EfficientNetB1, Xception, MobileNet, VGG19, InceptionResNetV2, and ConvNeXtLarge, has been pivotal in attaining optimal accuracy metrics for the classification endeavour.

Through extensive training cycles with these pre-trained networks, in conjunction with a diverse array of esteemed optimization algorithms such as ADAM, SGD, and NADAM, the research adeptly discerned the most performing amalgamation of model architecture and optimizer for the classification of brain tumor. This rigorous methodology not only augments the precision of classification outcomes but also guarantees the dependability of the predictions proffered by the system. The dataset utilized for the present research is acquired from Kaggle. For the training procedure, the dataset is partitioned into three distinct segments: training (70%), testing (20%), and validation (10%). Data pre-processing improves the quality of images by mitigating noise and selectively eliminating certain background elements. The table delineates the hyperparameters employed to train the proposed architectural framework. Moreover, integrating an explainable artificial intelligence framework has introduced an additional dimension of credibility to the model's predictions by elucidating the specific areas of the MRI scans that the model prioritizes in its decision-making process. This level of transparency is imperative for clinical applications, as it empowers healthcare practitioners to comprehend and substantiate the rationale behind the model's conclusions, ultimately enhancing the decision-making framework in identifying brain neoplasm.

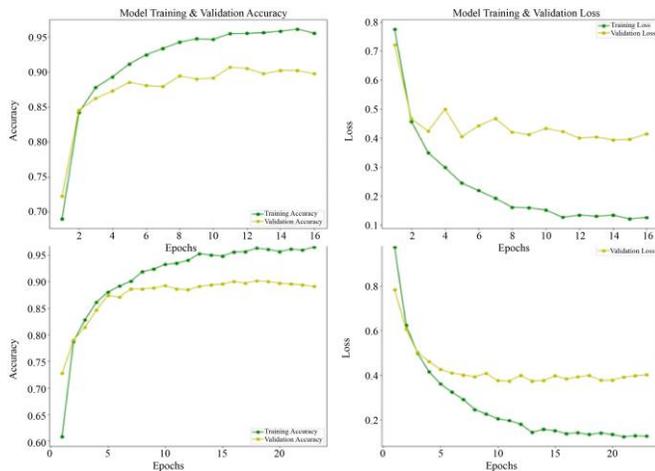
The research endeavor presented in this document is primarily centered on the objective of identifying the most effective combination of Transfer Learning (TL) based pre-trained models alongside the most suitable optimization algorithms to enhance performance. The accompanying Table 3 meticulously delineates the performance metrics associated with various pre-trained models when subjected to three distinct optimization techniques, specifically Stochastic Gradient Descent (SGD), Adaptive Moment Estimation (ADAM), and Nesterov-accelerated Adaptive Moment Estimation (NADAM). In the context of the ResNet50 architectural framework, it becomes evident that the SGD optimization method consistently surpasses the other techniques across all performance metrics, achieving a remarkable score of 100% as delineated in section 3.8 of this study.

Conversely, the EfficientNetB1 architecture emerges as the superior model, demonstrating exceptional performance by exceeding the benchmarks established by all three optimization algorithms, thereby attaining a flawless 100% accuracy, precision, F1 score, and recall. Furthermore, the InceptionResNetV2 model exhibits commendable performance when optimized using the NADAM algorithm, achieving an impressive 100% across all relevant metrics, while the ConvNeXtLarge model is optimally enhanced by either the ADAM or NADAM optimizers, also achieving a perfect accuracy rate of 100%. From the comprehensive analysis outlined in the preceding discussion, it is conclusively determined that the EfficientNetB1 architecture, when paired with any one of the optimization algorithms explored in this research, stands out as the most effective model available.

appropriate model and optimizer combination, particularly in fields where accuracy and precision hold paramount importance. The results of this investigation provide essential knowledge to enhance pre-trained model optimization in different applications.

**Table 3. Performance evaluation of trained model based on performance metric**

Pre-Trained CNN model with TL	Optimizer	Accuracy	Precision	Recall	F1-Score
ResNet50	SGD	1	1	1	1
	ADAM	0.9	0.93	0.9	0.9
	NADAM	0.95	0.96	0.95	0.94
Efficient NetB1	SGD	1	1	1	1
	ADAM	1	1	1	1
	NADAM	1	1	1	1
MobileNet small	SGD	0.25	0.13	0.25	0.16
	ADAM	0.25	0.06	0.25	0.1
	NADAM	0.3	0.19	0.3	0.18
VGG19	SGD	0.25	0.06	0.25	0.1
	ADAM	0.85	0.87	0.85	0.83
	NADAM	0.95	0.96	0.95	0.95
ConvNeXt Large	SGD	0.95	0.96	0.95	0.95
	ADAM	1	1	1	1
	NADAM	1	1	1	1
Xception	SGD				
	ADAM	0.95	0.96	0.95	0.95
	NADAM	0.95	0.96	0.95	0.95
Inception ResNetV2	SGD	0.95	0.96	0.95	0.95
	ADAM	0.95	0.96	0.95	0.95
	NADAM	1	1	1	1



**Fig. 4 Accuracy and loss performance of EfficientNet + ADAM and EfficientNet + SGD**

In the specific context of medical applications, it is imperative to note that precision serves as a critical parameter, which is notably maximized to an exceptional 100% in the case of the EfficientNetB1 architecture. Consequently, this research underscores the significance of selecting the

Upon careful examination of the data presented in Table 4, along with a thorough analysis of the accuracy and loss plots associated with both training and validation processes, it becomes remarkably evident that the ResNet50 architecture, when paired with the NADAM optimization algorithm, as well as the EfficientNetB1 model utilizing various optimizers such as SGD, ADAM, or NADAM, alongside the VGG19 framework also employing the NADAM optimizer, collectively demonstrate a considerable reduction in loss metrics and a significant enhancement in accuracy levels, all while effectively mitigating the risk of overfitting.

Conversely, it has been observed that certain models, particularly the MobileNetSmall variant, exhibit a tendency to overfit regardless of the optimization techniques applied, thereby compromising the generalizability of the model's ability across previously unobservable data. This phenomenon underscores the critical importance of selecting appropriate model architectures and optimization to overfit regardless of the optimization techniques applied, thereby compromising the generalizability of the model's ability across previously unobservable data. The correct model selection combined

with optimal optimization strategies enables achieving an ideal balance between loss reduction and accuracy improvement without causing unreliable predictive capabilities from model overfitting.

The EfficientNet model trained with ADAM optimized the accuracy of predicting different tumor classifications through visualization in Figure 5, which presents the confusion matrix results. During diagnosis, it becomes necessary to identify all patients who have active cancer because they are cancerous. The inability of deep learning model detection to identify brain tumors results in deadly health problems for patients. The proposed model achieves excellent brain tumor detection capability through its strong curve bending towards the top left corner that minimizes both false classifications and diagnoses. A plot representing the ROC curve of the optimal pre-trained CNN model combination with optimizers is presented in Figure 6.

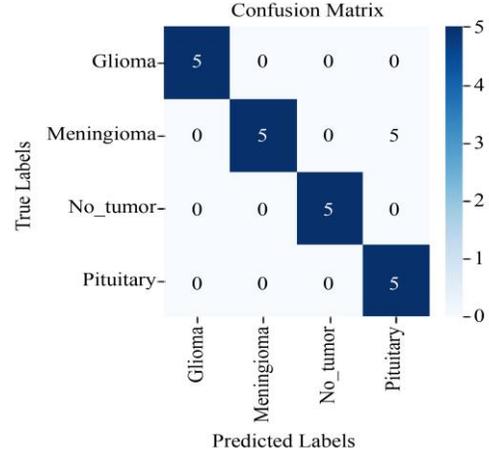


Fig. 5 Confusion matrix for EfficientNet + ADAM

Table 4. Performance of the model based on accuracy and loss

Pre-trained CNN Model with TL	Optimizer	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
ResNet 50	SGD	98.43	90.64	0.062	0.897
	ADAM	93.90	85.27	0.24	0.46
	NADAM	98.34	91.56	0.0093	0.41
Efficient Net B1	SGD	96.42	90.64	0.1178	0.4222
	ADAM	<b>95.17</b>	<b>89.26</b>	<b>0.1404</b>	<b>0.3751</b>
	NADAM	98.30	91.26	0.0645	0.4221
Mobile Net small	SGD	64.78	27.91	0.8565	1.5051
	ADAM	52.19	16.14	1.0597	1.7234
	NADAM	73.33	38.65	1.9540	0.6598
VGG19	SGD	26.82	NaN	27.81	NaN
	ADAM	98.22	88.50	0.0650	0.5318
	NADAM	99.77	93.56	0.0072	0.5611
ConvNe Xt Large	SGD	97.26	89.42	0.0676	0.4018
	ADAM	98.37	88.96	0.0654	0.3610
	NADAM	99.47	90.34	0.0176	0.4371
Xception	SGD	97.05	89.26	0.0925	0.3803
	ADAM	98.82	89.26	0.0558	0.3901
	NADAM	99.29	90.80	0.0237	0.4163
Inception ResNet V2	SGD	97.60	90.95	0.0755	0.3948
	ADAM	99.26	90.49	0.0329	0.4131
	NADAM	99.80	90.95	0.111	0.4867

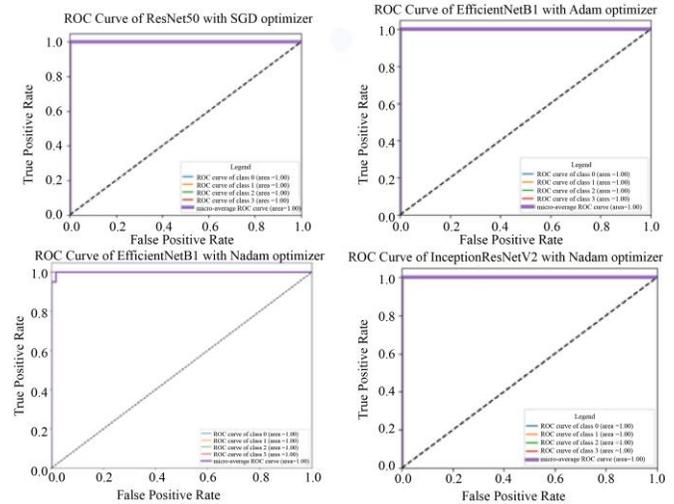


Fig. 6 ROC curve of trained models using the proposed approach

The Grad CAM model significantly enhances the interpretability of the predictions made by the models, as it effectively emphasizes the specific regions within the MRI scans that are instrumental in informing the Convolutional Neural Network (CNN) decision-making process. The Figure 6 presented illustrates the Heat Map that has been generated through the application of the EfficientNetB1 architecture, utilizing the ADAM optimization algorithm, while also depicting the superimposition of this heatmap onto the MRI scan, thereby concentrating on the particular areas that are critically evaluated by the proposed model in its endeavor to accurately detect the presence of tumor regions. By elucidating these pertinent areas of focus, the Grad CAM model not only bolsters the confidence in the predictions made by the CNN but also provides invaluable insights into the underlying mechanisms that govern the model's decision-making process in the context of medical imaging. This systematic approach to visualizing model predictions not only fosters a deeper understanding of CNN's operational dynamics but also aids in validating the trust ability and efficacy of such advanced diagnostic techniques in clinical settings.

Table 5. Comparison of previous research utilizing same dataset

Author	Model	Accuracy	Precision	Recall
Kumar RL, Kakarla J, Isunuri BV, Singh M (2021) [29]	ResNet-50 and GAP layer	98	97.82	98.06
Verma and Singh (2022) [30]	DenseNet201 and Transfer Learning	98.22	97.69	98.01
Athisayamani S, Antonyswamy RS, et al. (2023) [31]	ResNet-152	98.85	97	95
Alhassan AM, Zainon WMNW (2021) [32]	CNN with Swish based Relu	98.6	99.6	98.6
Rasheed Z, Ma Y-K, Ullah I et al. (2023) [33]	Convolutional neural network	98.04	98	98
Malla PP, Sahu S, Alutaibi AI (2023) [34]	Pre trained VGG16 and GAP	98.93	99.11	98.63
Khalid M. Hosny <sup>1</sup> , Mahmoud A. Mohammed et al. (2024) [35]	DenseNet121 and InceptionNetV3	99.02	98.75	98.98
Proposed Approach	EfficientNet B1 +any optimizer	<b>100</b>	<b>100</b>	<b>100</b>

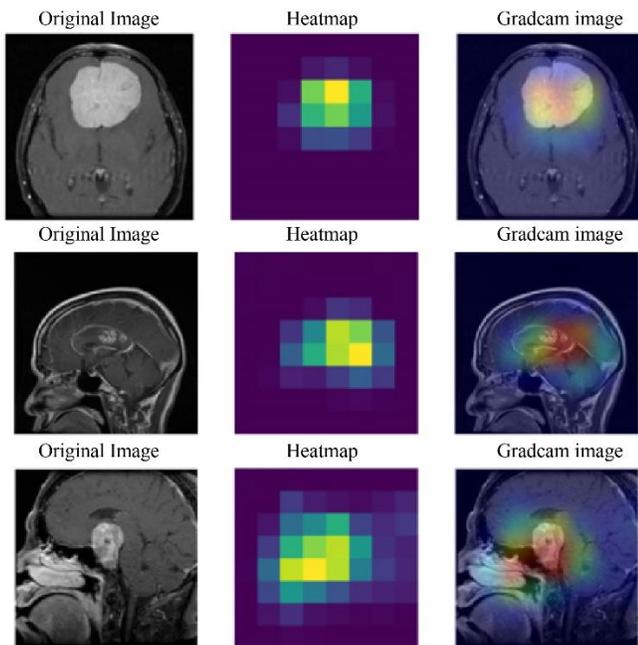


Fig. 7 Explainability indicated by Grad CAM model for EfficientNet

## 5. Conclusion

This research confirmed the effectiveness of deep learning techniques using transfer learning for brain neoplasm classification from MRI scans. Pre-trained neural networks ResNet50 and EfficientNetB1, along with other models, show successful results in accurate tumor classification according to distinct classes. Research showed that EfficientNetB1 partnered with any optimizer and provided exceptional performance metrics, so the evaluation reached perfect values for all accuracy and precision and F1-score and recall categories. The implementation of Grad CAM and other explainable artificial intelligence methods helps medical professionals better understand the operational logic of their models at a time when they need more clarity about decisions made by AI systems for clinical applications. The research demonstrates how correct model architecture combined with optimal optimization methods allows medical imaging diagnostic systems to achieve superior performance outcomes. The findings from this study simultaneously advance our understanding of brain tumor classification and establish guidelines for future projects that optimize deep learning models within medicine.

## References

- [1] Manon Dubol, I. Sundström-Poromaa, and Erika Comasco, "P.275 Severity of Premenstrual Dysphoric Disorder Symptoms Correlates with Grey Matter Volume and Brain Surface Morphology," *European Neuropsychopharmacology*, vol. 40, pp. S157-S158, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [2] Glioma, Johns Hopkins Medicine, 2023. [Online]. Available: <https://www.hopkinsmedicine.org/health/conditions-and-diseases/gliomas/>
- [3] M. Nagabushanam et al., "Detection and Localization of Brain Tumors Using Fractional Hartley Transform and Adaptive Neuro-Fuzzy Inference System Classification Methods," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, pp. 8851-8858, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [4] Brain Tumor Facts, National Brain Tumor Society, 2022. [Online]. Available: <https://braintumor.org/brain-tumors/about-brain-tumors/brain-tumor-facts/>
- [5] Suzan Uysal, *Functional Neuroanatomy and Clinical Neuroscience: Foundations for Understanding Disorders of Cognition and Behavior*, Oxford University Press, pp. 1-464, 2023. [[Google Scholar](#)] [[Publisher Link](#)]
- [6] Sakshi Ahuja et al., "Deep Learning-based Computer-aided Diagnosis Tool for Brain Tumor Classification," *11<sup>th</sup> International Conference on Cloud Computing, Data Science & Engineering*, Noida, India, pp. 854-859, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]

- [7] Paul Summers et al., "Whole-body Magnetic Resonance Imaging: Technique, Guidelines and Key Applications," *Ecancer Medical Science*, pp. 1-16, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [8] S. Deepak, and P.M. Ameer, "Brain Tumor Classification using Deep CNN Features via Transfer Learning," *Computers in Biology and Medicine*, vol. 111, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [9] Hossam H. Sultan, Nancy M. Salem, and Walid Al-Atabany, "Multi-Classification of Brain Tumor Images Using Deep Neural Network," *IEEE Access*, vol. 7, pp. 69215-69225, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [10] Wadhah Ayadi et al., "Brain Tumor Classification Based on Hybrid Approach," *The Visual Computer*, vol. 38, pp. 107-117, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [11] Javaria Amin et al., "Brain Tumor Classification based on DWT Fusion of MRI Sequences using Convolutional Neural Network," *Pattern Recognition Letters*, vol. 129, pp. 115-122, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [12] Maram Fahaad Almufareh et al., "Automated Brain Tumor Segmentation and Classification in MRI Using YOLO-Based Deep Learning," *IEEE Access*, vol. 12, pp. 16189-16207, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [13] Ayesha Jabbar et al., "Brain Tumor Detection and Multi-Grade Segmentation Through Hybrid Caps-VGGNet Model," *IEEE Access*, vol. 11, pp. 72518-72536, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [14] Ali Farzamnia et al., "MRI Brain Tumor Detection Methods Using Contourlet Transform Based on Time Adaptive Self-Organizing Map," *IEEE Access*, vol. 11, pp. 113480-113492, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [15] Zubair Atha, and Jyotismita Chaki, "SSBTCNet: Semi-Supervised Brain Tumor Classification Network," *IEEE Access*, vol. 11, pp. 141485-141499, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [16] Ankit Vidyarthi et al., "Machine Learning Assisted Methodology for Multiclass Classification of Malignant Brain Tumors," *IEEE Access*, vol. 10, pp. 50624-60640, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [17] Praveen Kumar Ramtekkar, Anjana Pandey, and Mahesh Kumar Pawar, "Accurate Detection of Brain Tumor using Optimized Feature Selection based on Deep Learning Techniques," *Multimedia Tools and Applications*, vol. 82, pp. 44623-44653, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [18] Ardendu Sekhar et al., "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. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [19] Shahriar Hossain et al., "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. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [20] Mingyang Xu, Limei Guo, and Hsiao-Chun Wu, "Novel Robust Automatic Brain-Tumor Detection and Segmentation Using Magnetic Resonance Imaging," *IEEE Sensors Journal*, vol. 24, no. 7, pp. 10957-10964, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [21] Hansen Wijanarko et al., "Tri-VAE: Triplet Variational Autoencoder for Unsupervised Anomaly Detection in Brain Tumor MRI," *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, Seattle, WA, USA, pp. 3930-3939, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [22] Wenxin Wang et al., "A Two-Stage Generative Model with CycleGAN and Joint Diffusion for MRI-based Brain Tumor Detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 6, pp. 3534-3544, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [23] Ashly Mathew, and A Christy, "A Comprehensive Study on Various Preprocessing Techniques for Brain MRI Image," *3<sup>rd</sup> International Conference on Electronics and Sustainable Communication Systems*, Coimbatore, India, pp. 1538-1541, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [24] Mingxing Tan, and Quoc Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," *Proceedings of the 36<sup>th</sup> International Conference on Machine Learning*, Long Beach, California, pp. 1-10, 2020. [[Google Scholar](#)] [[Publisher Link](#)]
- [25] Asmaul Hosna et al., "Transfer Learning: A Friendly Introduction," *Journal of Big Data*, vol. 9, pp. 1-19, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [26] Linan Fan, and Zaitian Long, "Optimization of Nadam Algorithm for Image Denoising based on Convolutional Neural Network," *7<sup>th</sup> International Conference on Information Science and Control Engineering*, Changsha, China, pp. 957-961, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [27] Ramprasaath R. Selvaraju et al., "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," *IEEE International Conference on Computer Vision*, Venice, Italy, pp. 618-626, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [28] Md Imran Nazir et al., "Utilizing customized CNN for Brain Tumor Prediction with Explainable AI," *Heliyon*, vol. 10, no. 20, pp. 1-16, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [29] R. Lokesh Kumar et al., "Multi-class Brain Tumor Classification using Residual Network and Global Average Pooling," *Multimedia Tools and Applications*, vol. 80, no. 9, pp. 13429-13438, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]

- [30] Aman Verma, and Vibhav Prakash Singh, "Design, Analysis and Implementation of Efficient Deep Learning Frameworks for Brain Tumor Classification," *Multimedia Tools and Applications*, vol. 81, pp. 37541-37567, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [31] Suganya Athisayamani et al., "Feature Extraction Using a Residual Deep Convolutional Neural Network (ResNet-152) and Optimized Feature Dimension Reduction for MRI Brain Tumor Classification," *Diagnostics*, vol. 13, no. 4, pp. 1-20, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [32] Afnan M. Alhassan, and Wan Mohd Nazmee Wan Zainon, "Brain Tumor Classification in Magnetic Resonance Image using Hard Swish-based RELU Activation Function-Convolutional Neural Network," *Neural Computing and Applications*, vol. 33, pp. 9075-9087, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [33] Zahid Rasheed et al., "Automated Classification of Brain Tumors from Magnetic Resonance Imaging using Deep Learning," *Brain Sciences*, vol. 13, no. 4, pp. 1-18, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [34] Prince Priya Malla, Sudhakar Sahu, and Ahmed I. Alutaibi, "Classification of Tumor in Brain MR Images Using Deep Convolutional Neural Network and Global Average Pooling," *Processes*, vol. 11, no. 3, pp. 1-17, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [35] Khalid M. Hosny et al., "Explainable Ensemble Deep Learning-based Model for Brain Tumor Detection and Classification," *Neural Computing and Applications*, vol. 37, pp. 1289-1306, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]