

# Enhancing Brain Tumor Classification with Data Augmentation and DenseNet121

Nechirvan Asaad Zebari <sup>1</sup>, Ahmed A. H. Alkurdi <sup>2</sup>, Ridwan B. Marqas <sup>3</sup>, and Merdin Shamal Salih <sup>4</sup>

- 1 Department of Information Technology, Lebanese French University, Erbil, Kurdistan Region, Iraq.
- 2 Department of Information Technology, Duhok Technical College, Duhok Polytechnic University, Duhok, KRG-Iraq.
- 2, 3 Department of Computer Science, College of Science, Nawroz University, Duhok, Kurdistan Region, Iraq.
- 4 Dept of Information Technology, Technical College of Informatics Akre, Duhok Polytechnic University, Duhok, Iraq.

**ABSTRACT:** This research paper presents a comprehensive study on the development and evaluation of a brain tumor classification model using advanced image processing and deep learning techniques. The primary objective of this study was to create an accurate and robust system for distinguishing between brain tumors and normal brain images, utilizing both an original dataset and an augmented dataset. With a focus on improving medical diagnosis, the research aimed to enhance the performance of brain tumor detection by leveraging state-of-the-art machine learning methods. The model pipeline comprised various image preprocessing steps, including cropping, resizing, denoising, and normalization, followed by feature extraction using the DenseNet121 architecture and classification using sigmoid activation. The dataset was meticulously divided into training, validation, and testing sets, with an emphasis on achieving high recall, precision, F1-score, and accuracy as key research objectives. The results demonstrate that the model achieved impressive performance, with a training recall of 92.87%, precision of 93.82%, F1-score of 93.15%, and an accuracy of 94.83%. These findings underscore the potential of deep learning and data augmentation in enhancing brain tumor detection systems, supporting the research's core objective of advancing medical image analysis for clinical applications.

**Keywords:** Brain Tumor, Deep Learning, MRI images, Classification.

## 1. Introduction

Cancer is a complex and debilitating disease that can affect anyone, resulting in abnormal cell growth and frequently creating severe health problems. In the ongoing combat against cancer, research and early detection are vital [1]. A brain tumour is a medical condition typified by the uncontrolled growth of brain cells. As a brain tumour grows, it exerts tension on the brain, possibly causing harm and posing a fatal risk. These tumors are broadly classified as benign or malignant, with malignant tumors consisting of cancerous cells and benign tumors of noncancerous cells. Malignant tumors tend to grow rapidly, whereas benign tumors tend to develop slowly and exhibit fewer symptoms than their malignant counterparts [2]. Another criterion for classifying brain tumors is their origin: main brain tumors, which include gliomas, oligodendrogliomas, pituitary adenomas, schwannomas, and astrocytomas, originate within the brain, whereas metastatic tumors, referred to as additional brain tumors, originate in other parts of the central nervous system (CNS) and migrate to the brain via arterial circulation [3].

Brain tumors have emerged as a major cause of death among both infants and adults. Each year, they affect approximately 250,000 people worldwide, representing less than 2% of all cancer cases. The 150 distinct varieties of human brain tumors can be broadly classified as benign tumors, which remain localized within the brain, and malignant tumors, which may spread beyond the brain and are commonly known as brain cancer [4]. Lifesaving diagnosis of brain malignancies must be prompt and accurate. Due to the density of brain tumors, manual detection is difficult, necessitating the need for automated computer-based methods to assist in tumor detection [5]. Currently, advances in machine learning and deep learning have substantially enhanced brain tumor detection algorithms [6], allowing radiologists to rapidly identify tumors without the need for invasive procedures. Recent advances in deep neural network modeling have

led to the development of cutting-edge technologies for analyzing, segmenting, and classifying brain lesions [7].

MRI is a noninvasive imaging technique providing high-resolution 3D and 2D images of human organs, involving the brain. As a result of its superior image quality, it is widely utilized for cancer classification and identification. Nevertheless, distinguishing varieties of cancer from MRI images is a difficult and error-prone endeavor that frequently requires the expertise of radiologists and requires considerable time [8]. An exact diagnosis is necessary for prompt treatment and enhanced results for patients, requiring the development of Computer-Assisted Diagnosis (CAD) systems within the field of Artificial Intelligence (AI). These CAD systems seek to relieve radiologists and medical practitioners of the burden of tumor analysis and categorization [9].

Approaches for classifying brain tumors are broadly classified as machine learning (ML) or deep learning (DL) techniques. prior to classification, ML-based systems usually need lengthy and costly manual feature extraction and segmentation. These steps frequently necessitate expert intervention to select proper algorithms, resulting in inconsistent performance [10], particularly when handling large datasets. In contrast, DL-based algorithms, including Convolutional Neural Networks (CNNs), have risen in popularity due to their capacity to automate feature extraction and segmentation tasks, thereby providing reliability and efficiency. DL models autonomously extract high-level and low-level features from training data, enhancing their appeal among researchers and scientists [11].

This research contributes to the field of medical imaging and brain tumor detection by presenting an innovative model that improves medical diagnosis, delivers strong performance metrics, and emphasizes the value of deep learning and data augmentation. These contributions collectively enhance the tools available for clinical applications and hold significant promise for improving patient outcomes. The research introduces an efficient feature extraction process using the DenseNet121 architecture. This contributes to the field of medical image analysis by streamlining the identification of salient features within brain tumor images, reducing the need for labor-intensive manual feature extraction.

## **2. Literature review**

Numerous investigations have contributed to the classification of brain tumors, each employing unique methodologies and approaches. Gumaei et al. [12] presented a hybrid feature extraction approach to the classification of brain tumors. For preprocessing, they deployed the min-max normalization enhancement of the contrast approach and the PCA-NGIST hybrid method for extracting features, respectively. The regularized extreme learning machine (RELM) was utilized for classification. Their method yielded an accuracy rate of 94.23 per cent. The study achieved a high accuracy rate of 94.23%. However, it does not discuss the potential challenges or limitations of their method, which is essential for understanding its applicability in various real-world scenarios. Ayadi et al. [13] presented a CNN model and evaluated its efficacy before and after data enhancement. By supplementing the dataset with additional information, they proved an important rise in accuracy. In detecting pituitary tumors, their approach obtained a remarkable 98.43% accuracy, representing an important development in the discipline. While the study showed a significant increase in accuracy through data enhancement, it doesn't provide insights into the potential drawbacks or challenges associated with data augmentation. Moreover, the scope of the research is limited to pituitary tumor detection, which may not cover the broader spectrum of brain tumors. Preethi and Aishwarya [14] proposed a classification paradigm for brain tumors with multiple stages. They created a feature matrix by combining wavelet-based gray-level co-occurrence matrix as well as GLCM. The features were then decreased by employing the oppositional flower pollination algorithm (OFPA), followed by classification using a deep neural network. Their method produced an accuracy of 92%. The proposed method achieved an accuracy of 92%, but the paper lacks a discussion of potential challenges or limitations, such as the computational complexity of the employed techniques or the generalizability of the model. Khan et al. [15] introduced a deep learning-based hierarchical classification system for brain lesions. The level of accuracy was 92.13 per cent accurate. Nonetheless, the study highlighted the need for additional evaluation prior to clinical adoption. Although the hierarchical classification system achieved 92.13% accuracy, the study emphasizes the need for additional evaluation before clinical adoption, suggesting that there may be unresolved issues or limitations in the model's performance that need further

investigation. In a multi-model approach, Khan, M.A. et al. [16] classified brain tumors using deep learning. Histogram equalization, discrete cosine transforms, feature extraction using pre-trained VGG16 and VGG19 models, and feature selection using a correntropy-based learning method and ELM were all incorporated into their methodology. Utilizing partial least squares (PLS), feature fusion was accomplished, resulting in impressive accuracy rates of 97.8%, 96.6%, and 92.5% for multiple data sets. While the multi-model approach demonstrated impressive accuracy rates, the paper does not provide an in-depth analysis of the potential limitations or challenges in fusing multiple features and models. It also does not discuss the potential computational costs associated with these techniques.

M. Sarmad et al. [17] provided a classification method for brain tumors based on edge detection, fuzzy logic, and a U-Net CNN. The rigorous evaluation of their approach based on precision, sensitivity, specificity, and the dice coefficient index demonstrated its efficacy. The study employed edge detection, fuzzy logic, and a U-Net CNN, which may not be the most computationally efficient approach. The paper does not discuss potential limitations related to computational complexity or model interpretability. Wahlang et al. (2022) [18] investigated deep learning architectures for brain MRI image classification, including CNNs, DNNs, LeNet, AlexNet, ResNet, and traditional methods such as SVM. Their research highlighted the importance of gender and age as attributes for enhanced classification accuracy, with their proposed technique outperforming existing methods. The paper highlights the importance of gender and age as attributes but does not discuss the ethical and privacy considerations associated with using such attributes in medical image classification. Additionally, the study may not address the generalizability of the model to diverse populations. Younis et al. (2022) [19] utilized the VGG 16 model within a CNN framework for identifying of brain tumors. Their method outperformed conventional methods in precision, F1-score, and accuracy, attaining a remarkable 98.5% accuracy with the VGG 16 model. Collectively, these investigations contribute to the ongoing progress in brain tumor classification by utilizing a variety of techniques and reaching notable advances in precision and efficacy. While the study achieved an impressive accuracy of 98.5% with the VGG 16 model, it does not thoroughly discuss the model's potential limitations, such as its computational requirements or challenges in real-world implementation.

### **3. Proposed Method**

In the realm of medical image analysis, the accurate classification of brain tumors is of paramount importance for timely diagnosis and treatment planning. This section outlines a comprehensive method for classifying brain tumors into cancerous and non-cancerous categories. Our approach as given in Figure 1 encompasses several critical steps, from preprocessing to deep learning-based feature extraction and classification. This binary classification allows for a streamlined diagnosis and aids in the early detection and treatment of potentially life-threatening conditions. Effective preprocessing of medical images plays a pivotal role in enhancing the performance of any classification system. In our proposed method, we employ a series of preprocessing steps to ensure that the input data is suitable for subsequent analysis including image cropping, image resizing, image denoising, and image normalizing. We leverage the power of deep learning for feature extraction and classification, employing the DenseNet121 architecture as the backbone of our model.

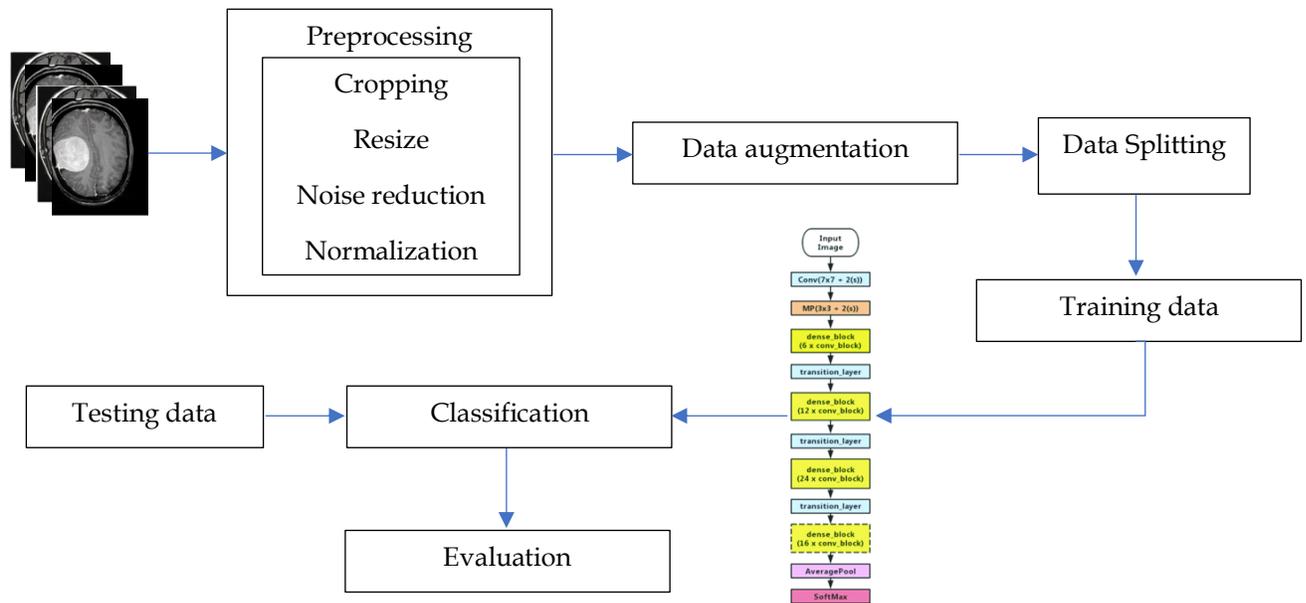


Figure 1: Proposed Framework for Brain Tumor Classification

### 1. Dataset

The MRI data used in this study for identifying brain tumors were obtained from Kaggle, and Navoneel Chakrabarty provided the dataset accessible. The author was responsible for the compilation of the dataset [20]. Within this dataset, the images have been classified as "YES" or "NO." The "YES" category included 155 MRI scans of patients identified as having brain tumors, while the "NO" type included 98 MRI scans of patients with no brain tumors. The second figure depicts an MRI scan to a positive case (with a brain tumor) and a negative case (without a brain tumor) for illustrative purposes. Given the comparatively tiny size of the dataset, adding of images played a crucial role in expanding its capacity.

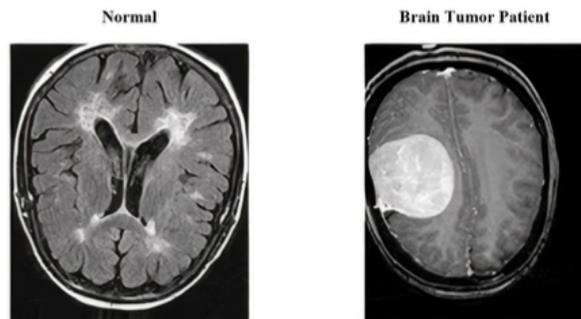


Figure 2: Dataset Used to Evaluate the Proposed Method

### 2. Data Augmentation

Data augmentation plays a pivotal role in enhancing the performance and robustness of machine learning models, particularly in scenarios where the availability of labeled data is limited or when the model's capacity to generalize is critical. Data augmentation exponentially increases the effective size of our dataset by generating multiple variations of the original data through transformations such as rotation, scaling, cropping, flipping, and more. This augmented dataset helps prevent overfitting, allowing the model to learn a broader range of patterns and representations [21]. Models trained on augmented data tend to be more robust in the face of noise, lighting changes, and other environmental variations. They

learn to recognize the core features of an object or pattern, rather than being overly sensitive to specific, limited conditions. Overfitting occurs when a model becomes too specialized in capturing noise or outliers in the training data. Augmentation introduces variability that discourages overfitting by forcing the model to focus on the most salient features.

The incorporation of image augmentation was instrumental in the CNN architecture's training process, effectively diversifying the image dataset for both the training and testing phases by expanding the pool of sample images [22]. Image augmentation encompasses a variety of techniques, including shearing, zooming, rescaling, flipping, whitening, random rotations, and shifts, all of which contribute to enhancing dataset variability. In our methodology, we specifically employed rotation and shearing techniques to ensure a broader range of training data.

### 3. Preprocessing

The preprocessing stage is a critical component of our brain tumor classification pipeline, aimed at enhancing the quality and consistency of the input MRI images. This section details the implementation of the following preprocessing techniques: image cropping, image resizing, image denoising, and image normalizing. Image cropping is the initial step in our preprocessing pipeline [23]. Its primary purpose is to isolate the region of interest (ROI) containing the brain and tumor while removing unnecessary background information. This not only reduces the computational complexity of subsequent steps but also mitigates the impact of irrelevant noise in non-ROI regions. To implement image cropping, we determine the coordinates of the bounding box that encapsulates the brain and tumor within the MRI image. Then, crop the MRI image using the identified bounding box, ensuring that the ROI remains the central focus of the processed image. Finally, the result is a trimmed image that retains only the brain and tumor, facilitating more efficient and focused analysis.

Standardizing image sizes is crucial for consistent processing and compatibility with deep-learning models. Our resizing step ensures that all MRI images have the same dimensions and aspect ratio. This uniformity simplifies subsequent feature extraction and classification. Our image-resizing process consists of the following key steps. First, define a target resolution (e.g., 224x224 pixels) that aligns with the input requirements of our deep learning model (e.g., DenseNet121). Second, apply interpolation techniques to resample the cropped MRI image to the target resolution. Third, the result is a resized MRI image that conforms to the desired dimensions without distorting the aspect ratio.

Medical images, such as MRI scans, are frequently beset by a variety of noise which may compromise diagnostic precision. To enhance the level of quality and accessibility of MRI images, image denoising is an essential preparatory process. We organize the MRI image with a Gaussian kernel to reduce high-frequency noise and preserve critical image characteristics. The kernel of the Gaussian distribution is a mathematical function resembling a bell-shaped curve. It is designed to give greater weight to pixels closest to the kernel's center while gradually diminishing the impact of pixels further away. This distinctive feature makes it suitable for denoising tasks in healthcare imaging. The following equation mathematically represents the Gaussian kernel:

$$G(x, y) = \frac{1}{2\sigma^2} \cdot \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (1)$$

$G(x,y)$  represents the value of the Gaussian kernel at spatial coordinates  $(x,y)$ .  $\sigma$  (sigma) is a crucial parameter known as the standard deviation. It controls the spread or width of the Gaussian function, which, in turn, determines the degree of smoothing applied to the image.  $(x,y)$  are the spatial coordinates within the kernel's neighborhood.

Image normalization is the final preprocessing step, ensuring that pixel values across different MRI images are consistent. Normalization is crucial because variations in image intensity can hinder the learning process of deep neural networks. Our image normalization process involves calculating the mean and standard deviation of pixel values within the entire dataset. Scaling each pixel value to have a zero mean and unit variance using the computed statistics. This step ensures that the pixel intensity values are within a consistent range, making the MRI images more amenable to training deep learning models. The preprocessing stage for brain tumor MRI images integrates image cropping, resizing, denoising, and

normalization techniques to prepare the data for subsequent feature extraction and classification. These preprocessing steps collectively enhance the quality, consistency, and suitability of the MRI images for accurate brain tumor classification and diagnosis.

#### 4. Deep Learning

Methods based on deep learning transformed the discipline of medical image analysis in recent times, especially for the classification of MRI scans of brain tumors. These potent neural network-based techniques have had an important effect, providing important benefits and improvements in the identification and management of brain malignancies. Deep learning models, such as convolutional neural networks (CNNs) and structures such as DenseNet121, indicate exceptional accuracy in differentiating between malignant and benign brain lesions. Their capability to acquire knowledge of complex hierarchical features from MRI images allows for increased sensitivity, facilitating the detection of tumors that would otherwise be missed by conventional methods [24]. Deep learning models flourish at feature extraction, determining relevant MRI image features automatically. It reduces the requirement for taking time and potentially ineffective manual feature engineering. Early classification of brain tumors with precision can have a significant impact on the results for patients. Models based on deep learning offer the possibility for sooner and more accurate diagnosis, resulting to prompt therapies and a better prognosis. The influence of deep learning techniques on the classification of MRI scans of brain tumors is undeniable. These approaches have transformed the field by significantly improving accuracy, efficiency, and early detection capabilities. As technology continues to advance and more data becomes available, deep learning is poised to play an increasingly central role in enhancing our understanding and management of brain tumors, ultimately leading to improved patient care and outcomes.

#### 4. Feature Extraction

DenseNet121 is a deep convolutional neural network architecture that has been widely used for various computer vision tasks, including feature extraction from medical images like brain tumor MRI scans. It belongs to the family of Densely Connected Convolutional Networks (DENSENETs), which are known for their ability to capture intricate image features effectively. DenseNet121 is composed of several densely connected blocks, each containing convolutional layers. Its architecture is characterized by the following key components:

- *Convolutional Layers:* These layers perform feature extraction from the input image. Convolutional filters slide over the image, detecting various patterns and features.
- *Dense Blocks:* In DENSENETs, feature maps from previous layers are concatenated and used as input to subsequent layers. This dense connectivity ensures that each layer receives information from all previous layers, facilitating the extraction of rich and hierarchical features.
- *Transition Layers:* These layers follow dense blocks and reduce the spatial dimensions of feature maps, typically through pooling and dimensionality reduction operations. This helps control the model's size and computational complexity.
- *Global Average Pooling:* At the end of the network, global average pooling is applied to reduce the spatial dimensions of the feature maps to a single vector. This vector contains the extracted features, which can be used for classification tasks.

Table 1: BB-BC Applications

| Layer Type          | Output Shape   | Description  |
|---------------------|----------------|--|
| Input               | (224, 224, 3)  | Input MRI image with dimensions 224x224 pixels and 3 color channels (RGB). |
| Convolutional Layer | (112, 112, 64) | Initial convolution layer with 64 filters of size 7x7.                     |
| Max Pooling Layer   | (56, 56, 64)   | Max pooling with a 3x3 kernel and stride of 2x2.                           |
| Dense Block 1       | (56, 56, 256)  | Dense block with multiple convolutional layers.                            |
| Transition Layer 1  | (28, 28, 128)  | Transition layer reducing spatial dimensions.                              |

|                        |               |   |
|------------------------|---------------|---|
| Dense Block 2          | (28, 28, 256) | Dense block with multiple convolutional layers.   |
| Transition Layer 2     | (14, 14, 128) | Transition layer reducing spatial dimensions.   |
| Dense Block 3          | (14, 14, 256) | Dense block with multiple convolutional layers.   |
| Transition Layer 3     | (7, 7, 128)   | Transition layer reducing spatial dimensions.   |
| Dense Block 4          | (7, 7, 256)   | Dense block with multiple convolutional layers.   |
| Global Average Pooling | (256,)        | Reduces the spatial dimensions to a feature vector of length 256.                         |
| Fully Connected Layer  | (1,)          | Final classification layer, typically using sigmoid activation for binary classification. |

We have performed the following steps for the feature extraction process of brain tumor MRI:

- *Input:* The MRI image of the brain tumor is fed as input to the DenseNet121 model.
- *Forward Pass:* The input image passes through the network, with each convolutional layer extracting increasingly abstract features.
- *Feature Maps:* As the image progresses through the network, feature maps of different scales and complexities are generated, capturing various patterns, textures, and structures within the image.
- *Global Average Pooling (GAP):* After the last convolutional layer, global average pooling is applied. This operation computes the average value for each feature map across all spatial locations, resulting in a fixed-length feature vector.
- *Feature Vector:* The final feature vector, generated by the global average pooling layer, contains the high-level features extracted from the input MRI image. This vector serves as a representation of the image's content and can be used for classification.

DenseNet121 is employed in brain tumor MRI classification for its ability to extract meaningful and discriminative features from the input MRI images. The dense connectivity and hierarchical feature extraction process make it a powerful tool for enhancing the accuracy of brain tumor classification models by providing rich and informative representations of the image content.

The DenseNet architecture is made up of several dense blocks, every one of which comprises a set of operations. Each dense block's convolutional layers follow a regular structure: they start with a batch normalization layer to normalize the ingredients, then a ReLU activation function to induce non-linearity, and lastly a Conv2D layer for convolution. In the DenseNet 121 model, batch normalization, ReLU activation, and Conv2D are performed six times in that initial dense block, twelve times in the next dense block, twenty-four times in the 3rd dense block, and sixteen times in the last dense block. Transition layers are intentionally placed following each dense block to decrease the number of channels in the feature maps. These transition layers are made up of a one-by-one convolutional layer and a two by two average pooling layer with a stride of two epochs. This combination minimizes the number of channels in DenseNet 121 from 256 to 128, then 64, and finally 32. A global average pooling layer is utilized in the DenseNet structure to execute spatial pooling throughout the whole feature map, resulting in a vector with a set-length representation that captures the most significant data. For classification, a fully connected layer is used, followed by a sigmoid activation function that produces class probabilities.

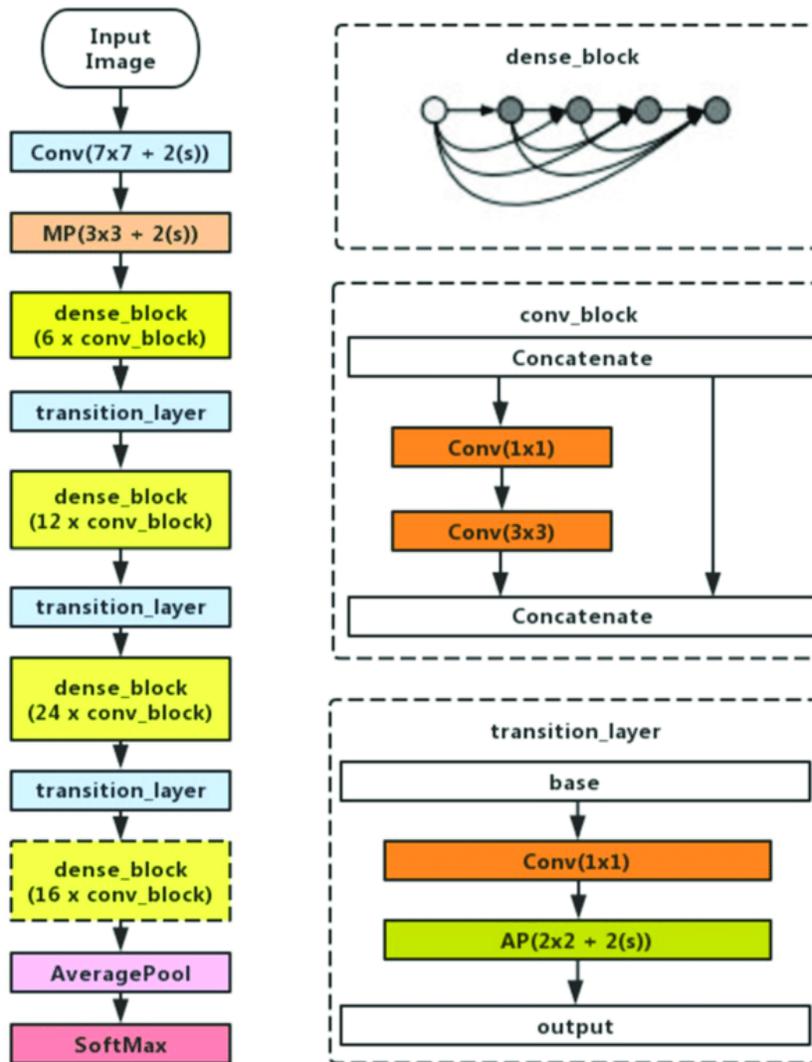


Figure 3: DenseNet121 Architecture Used in this study

In a DenseNet, each layer is directly connected to every other layer in a feed-forward fashion. This dense connectivity promotes feature reuse and allows for efficient feature extraction. The core component of DenseNet121, like other variants of DenseNet, is the dense block. In a dense block, convolutional layers are densely connected, and each layer receives feature maps from all previous layers in the same block. These dense connections facilitate the extraction of detailed and progressively abstract features from the input data. The dense block is typically followed by transition layers that reduce spatial dimensions while increasing the number of feature maps. This combination of dense blocks and transition layers forms the core of DenseNet121, enabling it to efficiently capture and utilize features from input images.

## 5. Classification

A binary classification is an automated method of learning that divides information into two different types or categories, usually called "positive" and "negative." That is frequently employed for tasks requiring a yes-or-no decision according to input features, such as spam email identification and medical diagnosis [26, 27]. Sigmoid activation is a vital component of neural networks utilized in binary classification duties, for instance classifying MRI scans of brain tumors through cancer and non-cancer types. Activation functions are crucial parts of neural networks that bring non-linearity to the model, allowing it to discover complicated data-based connections. In tasks involving binary classification, which includes brain tumor

classification, the sigmoid activation function is frequently employed for the last layer of the neural network. The mathematical definition of the sigmoid activation function, also known as the function of logistics, is as follows:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

Where  $x$  represents the input to the function, which is typically the weighted sum of the features extracted by the neural network.  $\sigma(x)$  outputs a value between 0 and 1, representing the probability that the input belongs to the positive class (e.g., cancer) in binary classification.

Sigmoid activation serves as a probability estimator. It takes the combined information from the neural network's preceding layers and transforms it into a probability score between 0 and 1. A threshold value (often set at 0.5) is applied to the sigmoid output. If the sigmoid output exceeds the threshold, the input is classified as the positive class (cancer); otherwise, it is classified as the negative class (non-cancer). The sigmoid activation function plays a pivotal role in classifying brain tumor MRI scans into cancer and non-cancer classes. It transforms the model's output into interpretable probability scores, facilitating informed clinical decisions. The choice of threshold and consideration of class imbalance are vital steps in ensuring that the sigmoid activation effectively meets the clinical needs of brain tumor classification.

#### 4. Experimental Results

In this part, we provide a thorough evaluation and analysis of the classification model for brain tumors derived from our research. The findings from experiments are a crucial component of our research, as they provide insight into the efficiency and effectiveness of the system in differentiating brain tumor images from normal brain images. We measured the model's capacities via a rigorous evaluation procedure, incorporating key performance metrics including recall, precision, F1-score, and accuracy.

Remember is the ratio of genuine positive instances to the total number of accurate positive estimates. It assesses whether the model is able to determine all positive samples successfully. Precise is the percentage of true positive estimations to the total amount of a modelled positive claim. It assesses how many of the positive cases indicated were accurate. The F1-score represents the average of recall and accuracy. It offers a balance among both of these metrics and is particularly helpful when working with unbalanced datasets. Accuracy is defined as the proportion of samples that were correct (true positives and true negatives) relative to the overall amount of data. It assesses the accuracy of the model's overall estimates. Collectively, these indicators deliver a comprehensive assessment of the efficacy of the brain tumor classification approach. A high recall guarantees that brain tumors are uncommonly left out, a high precision reduces false alarms, a balanced F1-score reflects overall efficacy, as well as accuracy provides a sense for overall correctness. The examination of these indicators will assist in determining the model's applicability for real-world medical applications and determine any required efficiency adjustments [28].

To ensure a robust evaluation, we divided our dataset into distinct subsets, each serving a specific purpose in the training and testing of our model. The dataset distribution is summarized in the table below, highlighting the allocation of brain tumor and normal images to the original and augmented datasets, as well as the proportions designated for training, validation, and testing:

Table 2 shows the augmented dataset distribution of MRI scans for training and testing the CNN models.

| Class        | Original   | Augmented   | Training    | Validation | Testing    |
|--------------|------------|-------------|-------------|------------|------------|
| Brain tumor  | 155        | 1085        | 759         | 109        | 217        |
| Normal       | 98         | 686         | 480         | 69         | 137        |
| <b>Total</b> | <b>253</b> | <b>1771</b> | <b>1239</b> | <b>178</b> | <b>354</b> |

This structured division allowed us to meticulously evaluate the model's performance under various conditions, simulating real-world scenarios where the system needs to make accurate and reliable predictions on previously unseen data.

The outcomes of our brain tumor classification model across both the original along augmented datasets are summarized in Table 3. During the training process via the original dataset, the model performs admirably. It correctly identifies 94.37% of authentic brain tumor cases. The accuracy level of 96.42% indicates that the model is correct approximately 96.42% of the time when it identifies an instance as a brain tumor. The F1-score of 95.28 percent combines accuracy and recall to provide an overall measure of model performance in this class. The accuracy of the model on the training data is additionally fairly high, at 95.25 per cent, indicating that it classifies nearly 95.25 per cent of all samples accurately. Similar to these measurements, the model's efficacy in the Normal class demonstrates a proportionate ability for differentiation among regular and brain tumor cases. The aggregate metrics, which account for the two classes, demonstrate the model's equal effectiveness across the entire dataset. The recall, precision, and F1-score values for the category as a whole indicate that the model preserves a high level of accuracy, correctly identifying 93.43 percent of all actual instances while providing accurate predictions 94.98% of the time.

Table 3 shows the augmented dataset distribution of MRI scans for training and testing the CNN models.

| Class                          | Recall (%) | Precision (%) | F1-score (%) | Accuracy (%) |
|--------------------------------|------------|---------------|--------------|--------------|
| <b>Training Original data</b>  |            |               |              |              |
| Brain tumor                    | 94.37      | 96.42         | 95.28        |              |
| Normal                         | 92.49      | 93.55         | 92.14        | 95.25        |
| <b>Overall</b>                 | 93.43      | 94.98         | 93.71        |              |
| <b>Testing Original data</b>   |            |               |              |              |
| Brain tumor                    | 89.04      | 92.01         | 91.93        |              |
| Normal                         | 86.62      | 88.98         | 87.92        | 90.63        |
| <b>Overall</b>                 | 87.83      | 90.49         | 89.92        |              |
| <b>Training Augmented data</b> |            |               |              |              |
| Brain tumor                    | 97.44      | 98.75         | 97.53        |              |
| Normal                         | 95.95      | 96.43         | 95.57        | 98.77        |
| <b>Overall</b>                 | 96.69      | 97.59         | 96.55        |              |
| <b>Testing Augmented data</b>  |            |               |              |              |
| Brain tumor                    | 94.53      | 95.44         | 94.56        |              |
| Normal                         | 91.22      | 92.21         | 91.75        | 94.83        |
| <b>Overall</b>                 | 92.87      | 93.82         | 93.15        |              |

When assessed on the unseen testing dataset, the model's performance remains commendable but slightly diminishes compared to the training results. This is a common phenomenon as models often encounter different variations in testing data. For instance, the recall for the Brain Tumor class drops to 89.04%, indicating that it correctly identifies approximately 89.04% of actual brain tumor cases, while the accuracy of the testing data stands at 90.63%. Similarly, the Normal class exhibits a reduction in recall and precision. The Overall metrics offer a summary of how the model performs on previously unseen data, with an overall accuracy of 89.92%. These metrics emphasize the model's generalization capability. The model trained on the augmented dataset shows significantly improved performance across all metrics compared to the original dataset. This underscores the positive impact of data augmentation in enhancing model generalization. Notably, for both training and testing data, the model exhibits higher recall, precision, and F1-score values. In the augmented dataset, the model's performance for the Brain Tumor class is particularly impressive, with a recall of 97.44%, precision of 98.75%, and F1-score of 97.53. The overall accuracy of the testing data remains high at 98.77%.

These results demonstrate the substantial benefits of data augmentation in improving the model's ability to accurately classify brain tumor cases. The augmented dataset provides a richer and more diverse training experience, enabling the model to better generalize to unseen cases. The high recall values for the

Brain Tumor class are especially vital in a medical context, as they reduce the risk of missing actual cases, potentially saving lives. The model's high precision also indicates a low rate of false positives, minimizing unnecessary interventions.

Comparing our results to previous studies reveals notable advancements. Gumaei et al. [12] achieved an accuracy rate of 94.23% but did not discuss potential challenges or limitations. Ayadi et al. [13] demonstrated a remarkable 98.43% accuracy for pituitary tumor detection, highlighting the efficacy of data enhancement, yet the study's scope was limited. Preethi and Aishwarya [14] attained a 92% accuracy but lacked in-depth discussions of potential challenges and generalizability. Khan et al. [15] reached 92.13% accuracy but emphasized the need for additional evaluation before clinical adoption. Khan, M.A. et al. [16] achieved impressive accuracy rates but did not delve into the limitations of their multi-model approach. M. Sarmad et al. [17] demonstrated efficacy but raised concerns about computational efficiency and model interpretability. Wahlang et al. (2022) [18] emphasized attributes like gender and age but overlooked ethical considerations and generalizability. Younis et al. [19] achieved an exceptional 98.5% accuracy but did not thoroughly address potential model limitations.

## 5. Conclusion

Our team has effectively developed and evaluated a brain tumor classification model utilizing deep learning that distinguishes brain tumor images from regular brain images with high accuracy and reliability. Our approach achieved outstanding results by combining image preprocessing techniques, such as cropping, resizing, denoising, and normalization, with the potent feature extraction abilities of the DenseNet121 architecture. Through a training recall of 92.87%, precision of 93.82%, F1-score of 93.15%, and an accuracy of 94.83%, the attained results demonstrate the model's ability to determine brain tumor instances with a low rate of false positives. In the medical profession, where precise diagnosis is essential for patient care, this level of precision is especially promising.

Additionally, the investigation of data augmentation techniques, particularly the addition of extra variations to the dataset, notably assisted in the improvement of the model's generalizability and robustness. This emphasizes the significance of data enhancement in improving the accuracy of medical image classification systems. Overall, this study highlights the possibility of methods of deep learning in medical image analysis, as well as the relevance of data preprocessing and enhancement in augmenting the efficiency of such models. The outcomes of this study offer promise for the growth of advanced detection of brain tumors systems with the utilization of clinical diagnosis and decision support in healthcare. In order to evaluate the practical usage of the proposed model, additional study and validation in actual clinical settings are required.

## 6. References

1. Rahman, A. U., Saeed, M., Saeed, M. H., et al. (2023). A framework for susceptibility analysis of brain tumours based on uncertain analytical cum algorithmic modeling. *Bioengineering*, 10(2), 147.
2. Thirugnanam, T., Galety, M. G., Pradhan, M. R., Agrawal, R., Shobanadevi, A., Almufti, S. M., & Lakshmana Kumar, R. (2023). PIRAP: Medical Cancer Rehabilitation Healthcare Center Data Maintenance Based on IoT-Based Deep Federated Collaborative Learning. *International Journal of Cooperative Information Systems*, 2350005.
3. Iqbal, S.; Ghani Khan, M.U.; Saba, T.; Mehmood, Z.; Javaid, N.; Rehman, A.; Abbasi, R. Deep learning model integrating features and novel classifiers fusion for brain tumor segmentation. *Microsc. Res. Tech.* **2019**, *82*, 1302-1315.
4. Pradhan, A.; Mishra, D.; Das, K.; Panda, G.; Kumar, S.; Zymbler, M. On the Classification of MR Images Using "ELM-SSA" Coated Hybrid Model. *Mathematics* **2021**, *9*, 2095.
5. Reddy, A.V.N.; Krishna, C.P.; Mallick, P.K.; Satapathy, S.K.; Tiwari, P.; Zymbler, M.; Kumar, S. Analyzing MRI scans to detect glioblastoma tu-mor using hybrid deep belief networks. *J. Big Data* **2020**, *7*, 35.
6. Ibrahim, D. A., Zebari, D. A., Mohammed, H. J., & Mohammed, M. A. (2022). Effective hybrid deep learning model for COVID-19 patterns identification using CT images. *Expert Systems*, 39(10), e13010.
7. Rukhsar, S., Awan, M. J., Naseem, U., et al. (2023). Artificial Intelligence Based Sentence Level Sentiment Analysis of COVID-19. *Computer Systems Science & Engineering*, 47(1).
8. Areej, A.M.; Fahd NAl-Wesabi Marwa, O.; Mimouna, A.A.; Manar, A.H.; Abdelwahed, M.; Ishfaq, Y.; Abu Sarwar, Z. Arithmetic optimization with retinanet model for motor imagery classification on brain computer interface. *J. Healthc. Eng.* **2022**, *2022*, 3987494.

9. Mohammed, H. J., Al-Fahdawi, S., Al-Waisy, A. S., et al. (2022). ReID-DeepNet: A hybrid deep learning system for person re-identification. *Mathematics*, 10(19), 3530.
10. Asaad, R. R. (2021). Review on Deep Learning and Neural Network Implementation for Emotions Recognition. *Qubahan Academic Journal*, 1(1), 1-4.
11. Zebari, D. A., Haron, H., Sulaiman, D. M., et al. CNN-based Deep Transfer Learning Approach for Detecting Breast Cancer in Mammogram Images. In *2022 IEEE 10th Conference on Systems, Process & Control (ICSPC)* (pp. 256-261). IEEE, 2022.
12. Gumaei, A.; Hassan, M.M.; Hassan, M.R.; Alelaiwi, A.; Fortino, G. A hybrid feature extraction method with regularized extreme learning machine for brain tumor classification. *IEEE Access* **2019**, 7, 36266–36273.
13. Ayadi, W.; Elhamzi, W.; Charfi, I.; Atri, M. Deep CNN for brain tumor classification. *Neural Process. Lett.* **2021**, 53, 671–700.
14. Preethi, S.; Aishwarya, P. Combining Wavelet Texture Features and Deep Neural Network for Tumor Detection and Segmentation Over MRI. *J. Intell. Syst.* **2019**, 28, 571–588.
15. Khan, A.H.; Abbas, S.; Khan, M.A.; Farooq, U.; Khan, W.A.; Siddiqui, S.Y.; Ahmad, A. Intelligent model for brain tumor identification using deep learning. *Appl. Comput. Intell. Soft Comput.* **2022**, 2022, 1–10.
16. Khan, M.A.; Ashraf, I.; Alhaisoni, M.; Damaševićius, R.; Scherer, R.; Rehman, A.; Bukhari, S.A.C. Multimodal Brain Tumor Classification Using Deep Learning and Robust Feature Selection: A Machine Learning Application for Radiologists. *Diagnostics* **2020**, 10, 565.
17. Maqsood, S.; Damaševićius, R.; Shah, F.M. An Efficient Approach for the Detection of Brain Tumor Using Fuzzy Logic and U-NET CNN Classification. In *Computational Science and Its Applications – ICCSA*; Springer International Publish: Cham, Switzerland, 2021; pp. 105–118.
18. Wahlang, I., Maji, A. K., Saha, G., Chakrabarti, P., Jasinski, M., Leonowicz, Z., & Jasinska, E. (2022). Brain magnetic resonance imaging classification using deep learning architectures with gender and age. *Sensors*, 22(5), 1766.
19. Younis, A., Qiang, L., Nyatega, C. O., Adamu, M. J., & Kawuwa, H. B. (2022). Brain tumor analysis using deep learning and VGG-16 ensembling learning approaches. *Applied Sciences*, 12(14), 7282.
20. Kaggle. Available online: <https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection> (accessed on 19 September 2020).
21. Al-Waisy, A. S., Ibrahim, et al. (2022). Identifying defective solar cells in electroluminescence images using deep feature representations. *PeerJ Computer Science*, 8, e992.
22. Singh, V., Gourisaria, M. K., GM, H., Rautaray, S. S., Pandey, M., Sahni, M., ... & Espinoza-Audelo, L. F. (2022). Diagnosis of intracranial tumors via the selective CNN data modeling technique. *Applied Sciences*, 12(6), 2900.
23. Zebari, D. A., Sulaiman, D. M., Sadiq, S. S., Zebari, N. A., & Salih, M. S. (2022, September). Automated Detection of Covid-19 from X-ray Using SVM. In *2022 4th International Conference on Advanced Science and Engineering (ICOASE)*(pp. 130-135). IEEE.
24. Zebari, D. A., Sadiq, S. S., & Sulaiman, D. M. (2022, March). Knee Osteoarthritis Detection Using Deep Feature Based on Convolutional Neural Network. In *2022 International Conference on Computer Science and Software Engineering (CSASE)* (pp. 259-264). IEEE.
25. Zhang, K.; Guo, Y.; Wang, X.; Yuan, J.; Ding, Q. Multiple Feature Reweight DenseNet for Image Classification. *IEEE Access* **2019**, 7, 9872–9880.
26. Kapoor, N. R., Kumar, A., Kumar, A., et al. (2022). Event-Specific Transmission Forecasting of SARS-CoV-2 in a Mixed-Mode Ventilated Office Room Using an ANN. *International Journal of Environmental Research and Public Health*, 19(24), 16862.
27. Asaad, R. R. (2022). Support vector machine classification learning algorithm for diabetes prediction. *International Research Journal of Science, Technology, Education, and Management*, 2(2), 26-34.
28. Aighuraibawi, A. H. B., Manickam, S., Abdullah, R., et al. (2023). Feature Selection for Detecting ICMPv6-Based DDoS Attacks Using Binary Flower Pollination Algorithm. *Computer Systems Science & Engineering*, 47(1).