Brain Tumor Classification Using Machine Learning

Group Members:

Muhammad Awais	20L-2188
Akbar Khan	20L-2094
Fatima Munir	20L-2133

National University Of Computer and Emerging Sciences

Department of Computer Science

Lahore, Pakistan

Table of Contents

Abstract

Brain tumors are one of the most important causes of death among cancer types. Timely and precise diagnosis is pivotal for successful treatment outcomes. Contemporary advancements in technology aim to increase the success rate of neurosurgery and mitigate associated complications. Among these, the Magnetic Resonance Imaging (MRI) technique stands out as a widely employed method for examining brain tumor images. Various techniques and algorithms exist for image classification, with the overarching goal of automatic learning from training data to make informed decisions with high accuracy. In the classification process, the statistical properties of the input images were analyzed and the data were systematically divided into various categories. Benefits encompass empowering healthcare professionals, streamlining treatment decisions, and potentially reducing healthcare costs. The proposed work is divided into three parts: preprocessing, and then classification is done using a machine learning algorithm.

Problem

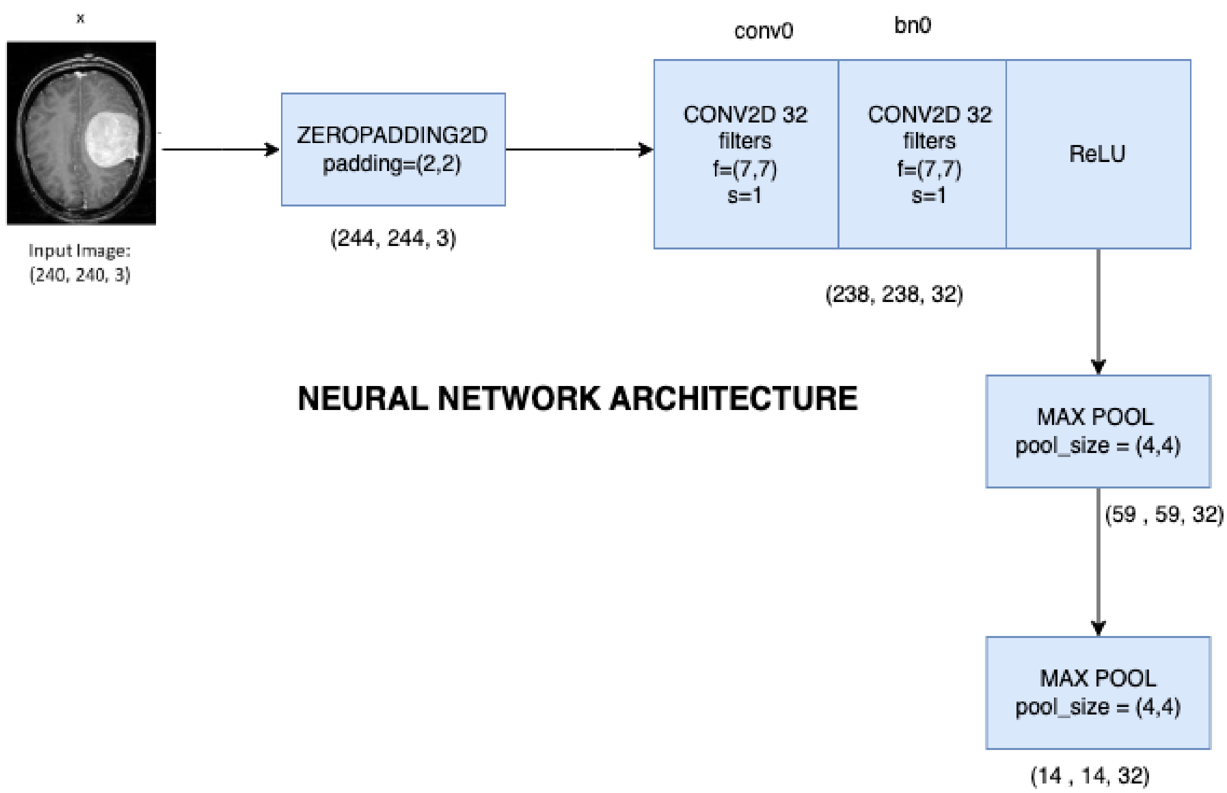
The detection and classification of brain tumors is a significant challenge in medical practices. Existing diagnostic methods are often time consuming and subject to potential inaccuracies, leading to delays in crucial treatment decisions. Additionally, the diverse nature of brain tumors requires a more sophisticated classification system for precise identification. Traditional approaches struggle to efficiently differentiate between tumor types due to their complexity. Recognizing these limitations, our research addresses the need for improved diagnostic precision and timely interventions. Using the capabilities of machine learning, we aim to revolutionize brain tumor classification. By training computers to recognize patterns in large datasets, we aspire to develop a more efficient and accurate system for categorizing various brain tumor types. This initiative seeks to enhance medical decision-making, making treatment faster and more tailored treatment strategies for individuals facing the challenges of brain tumors.

Introduction

The vitality of brain health is undeniably critical as it exerts a profound influence on overall well-being. In this research, we embark on a mission to use various brain scan datasets to accurately classify brain tumors, employing cutting-edge machine learning models and techniques for the precise categorization of brain scan images. In addressing the challenge of accurately classifying brain tumors, this research uses the power of machine learning models. Our methodology involves utilizing a comprehensive dataset comprising both brain scans with tumors and those representing healthy brains. To enhance the effectiveness of our model, we employ data augmentation techniques, introducing variations to the dataset for improved learning. Subsequently, we perform data preprocessing using diverse techniques to refine the information for optimal model training. The data is strategically divided into training and testing sets, helping the development of a robust machine learning model. Through rigorous training, our model becomes adept at distinguishing between brain tumor classifications. The summary of this process is the evaluation of our model's accuracy, providing valuable insights into its efficacy for addressing the complex task of brain tumor classification.

Implementation

Each input image, with dimensions (240, 240, 3), is processed through a neural network that comprises several layers, each serving a specific purpose. The initial step involves applying a Zero Padding layer with a pool size of (2, 2), which pads the image with zeroes to preserve spatial information during convolution. Subsequently, a convolutional layer with 32 filters is employed, utilizing a filter size of (7, 7) and a stride equal to 1. To accelerate computations, a batch normalization layer is introduced to normalize pixel values. This is followed by a Rectified Linear Unit (ReLU) activation layer, enhancing the model's ability to capture complex patterns. Max Pooling layers with a filter size (f) of 4 and a stride (s) of 4 are then applied twice, downsampling the spatial dimensions. A flatten layer is strategically placed to convert the 3-dimensional matrix into a one-dimensional vector. Finally, a Dense layer with one neuron, utilizing a sigmoid activation function, serves as the output unit for the binary classification task, facilitating the model's decision-making process.



Methodology

This project adopts a modular approach, consisting of Data Input, Pre-processing, and Classification modules, to achieve its goal of accurately diagnosing brain health conditions.

Data Input

The classification system begins its process by taking in brain scans as input, where each image is formatted with dimensions (240, 240, 3). As the neural network unfolds, the input image undergoes a sequence of operations.

Data Pre-processing

Raw scan data often contains noise that can hinder accurate classification. The Data Pre-processing module addresses this challenge using techniques like image normalization, filtering, and enhancement. This module employs standardization of image sizes, noise reduction through filters, and enhancement of relevant features. Traditional image processing and deep learning methodologies are combined to optimize data quality.

Classification

In the brain tumor classification model, a Convolutional Neural Network (CNN) architecture is employed to discern patterns and features within brain scan images. The model takes as input brain scans with dimensions (240, 240, 3). The initial layer, a Zero Padding2D, ensures the preservation of spatial information. Subsequently, a convolutional layer with 32 filters and a filter size of (7, 7) is applied, capturing details within the images. Batch normalization enhances computational efficiency by normalizing pixel values, followed by a Rectified Linear Unit (ReLU) activation layer to introduce non-linearity. The network employs two MaxPooling layers with distinct filter sizes, downsampling the data and facilitating hierarchical feature extraction. The flattened output is then connected to a Dense layer with a single neuron, utilizing a sigmoid activation function for binary classification. This architecture, encapsulated within the "BrainDetectionModel," strives to effectively classify brain scans, specifically focusing on the presence or absence of tumors.

Results

The results obtained from our most effective model, which was selected based on its superior validation accuracy, demonstrate its proficiency in detecting brain tumors. This model exhibits a remarkable accuracy of **88.7%** on the **test set**, highlighting its ability to correctly identify the presence or absence of brain tumors in a significant majority of cases. Additionally, it achieved an **F1 score** of **0.88** on the same set, reflecting a balanced and high performance in terms of both precision and recall. Such a high F1 score is particularly notable as it indicates the model's strength in accurately classifying true positives while minimizing false positives and negatives.

The efficacy of the model is underscored by the fact that these results were achieved using a balanced dataset, which presents an equal representation of both classes (tumors and no tumors). This balance in data helps in ensuring that the model's performance is not skewed towards any specific class, further emphasizing the reliability and robustness of these results.

	Validation set	**Test set**
Accuracy	91%	89%
F1 score	0.91	0.88

Literature Reviews

Brain Tumor Identification and Classification of MRI images using deep learning techniques [1]

Summary of the research item

The research paper introduces a novel approach called FAHS-SVM for the automated identification and segmentation of brain tumors. The method aims to achieve accuracy levels comparable to manual segmentation, addressing the challenge of inter observer variability. By leveraging both image structure hierarchy and statistical classification, the technique effectively identifies spatially small and consistent tumor areas, serving as a robust guide for subsequent segmentation. Experimental results on multiparametric Magnetic Resonance images demonstrate the method's promise in accurately and quickly locating brain tumors, with an impressive 98.51 percent accuracy in distinguishing abnormal and normal tissues. Overall, this research holds significant potential for aiding radiologists and clinical experts in diagnosing and supporting clinical decision systems in the field of MRI brain tumor detection.

Critical analysis of the research item

Strengths:

- High reported accuracy (98.51 percent).
- Focus on full automation for tumor detection.
- Combination of image structure and classification techniques.
- Emphasis on clinical relevance.

Weaknesses:

- Lack of technical details about the method.
- Limited information on the dataset and validation.
- Integration into clinical workflows not discussed.

Relationship to the proposed research work

The project shares similarities with the research work on “Brain Tumor Identification and Classification of MRI images using deep learning techniques.” Both projects involve the application of machine learning and deep learning techniques for the detection and classification of brain-related medical conditions. While the specific focus of the research work is on the brain.

Transfer Learning for Automatic Brain Tumor Classification Using MRI Images [2]

Summary of the research item

Brain tumor segmentation and classification using MRI through deep learning techniques is very challenging. Also, manual classification with human assistance can lead to improper prediction and diagno-sis. So, the study proposes a deep learning model for brain tumor classification, employing CNNs based on transfer learning. Three pre-trained CNN architectures, called ResNet, Xception, and MobileNet-V2, are discussed. Among these, MobileNet-V2 achieves the best results with an accuracy of 98.24 percent. Transfer learning is used to leverage pre-trained models for this task, addressing the challenge of limited available medical data.

Critical analysis of the research item

Strengths:

- This paper used a very brilliant strategy of transfer learning to address the lack of medical data.
- The model used data augmentation, normalization, and resizing techniques for brain tumor classification, ensuring the robustness of the system.

Weaknesses:

- The model only classifies brain tumors and its scope is limited.
- We cannot classify other brain diseases as the model trained through transfer learning restricts our ability to classify other brain diseases.

Relationship to the proposed research work

In healthcare, there’s a lack of medical data for training advanced deep learning models. To solve this, the study uses transfer learning, using pre-trained models for new tasks, and solving data scarcity problems. We can use this technique for better classification of diseases with limited dataset.

An Efficient Classification of MRI Brain Images [3]

Summary of the research item

The research highlights a significant advancement in medical imaging through the utilization of MRI for diagnosing brain disorders and injuries. This paper aims to develop an efficient method for classifying MRI brain images. The research proposes a simple yet efficient method for classifying MRI brain images as normal or abnormal, brain conditions like brain tumors. The methodology comprises four key stages pre-processing involving scalp and skull removal, feature extraction, feature reduction, and classification. The study employs individual classifiers called FF-ANN alongside hybrid classifiers called RSwithRF and RSwithBN, resulting in impressively accurate classifications 95.83 percent, 97.14 percent and 95.71 percent respectively. These outcomes highlight the method's potential in medical image classification.

Critical analysis of the research item

Strengths:

- The study suggested a new mechanism to differentiate MRI brain images into normal and abnormal using individual and hybrid classifiers.
- The model is based on a four-step process, data pre-processing, feature extraction, feature reduction, and classification for precise MRI brain image categorization.
- Accuracy is 97.14 percent which is far better than existing techniques.

Weaknesses:

- The evaluation involved only one single individual and two hybrid classifiers.
- Testing was conducted on just 70 images from a single dataset.
- Analysis focused on only three statistical features.
- The study does not compare its findings with the latest research in the field, particularly those based on deep learning techniques.

Relationship to the proposed research work

This study deals with the classification of MRI brain images, which aligns closely with our focus on utilizing MRI images for brain diseases detection. Understanding their approach to MRI image classification, even at a high level, might provide inspiration or points of comparison for our machine learning models. Considering the similarities in the use of MRI images for classification purposes, analyzing the methods presented in the paper could enhance the foundation of our research. It is important to critically observe their techniques, compare their results, and identify gaps in the context of brain cancer detection which potentially lead to more robust and accurate machine learning models for brain disease detection.

Classification of MRI Brain Images using k-Nearest Neighbor and Artificial Neural Network [4]

Summary of the research item

The research introduces an automated diagnosis approach on classification of MRI commonly utilized in outlining soft tissue. The method involves two main stages: feature extraction and classification. During the first stage, DWT is applied to MRI images which enables the extraction of frequency space information from non-stationary signals. The obtained features are then reduced to essential components using principal component analysis. In the classification stage, two classifiers are used: a FP-ANN and a k-NN classifier. The results demonstrate a classification success rate of 90 percent and 99 percent for FP-ANN and k-NN respectively, indicating the effectiveness of the proposed technique in automated diagnosis.

Critical analysis of the research item
Strengths:

- The research explores a various set of methodologies, enhancing the depth of analysis and comparison by employing both k-NN and ANN,
- The combination of traditional machine learning (k-NN) and deep learning (ANN) techniques allows for a broad evaluation.
- The study helps medical professionals to classify brain conditions with accuracy of 99 percent.

Weaknesses:
- The paper lacks an in-depth exploration of the limitations of the chosen classification methods.
- Limited information provided regarding the dataset used, such as its size, diversity, and sources.

Relationship to the proposed research work
One of the end goals of our project is to classify whether the brain scan is normal or not. This project uses a K-NN classification model and has the accuracy of 99 percent. So we can learn from the classification techniques used in the model preparation to increase the accuracy of our project.

A Hybrid CNN-SVM Threshold Segmentation Approach for Tumor Detection and Classification of MRI Brain Images [5]

Summary of the research item
The main thing discussed in this research paper is that it presents a wide view on detection of brain tumor and its classification using MRI images. The two technologies or methodologies used in this are CNN and SVM. The main reason for this research was to timely treat the brain tumors and so they could be detected timely. The accuracy which was obtained from this by using the public data set (BRATs2015) was 98.4959 percent.

Critical analysis of the research item
Strengths:
- The strength of this model was the combined use of CNN and SVM.
- This hybrid approach resulted in higher accuracy as they achieved an accuracy of 98 percent.
- This research paper provides all the pre-processing steps and the statistical functions which are used by the model.

Weaknesses:
- The weakness of this paper is that it does not discuss real world clinical validation
- There is an assumption of preprocessed data.

Relationship to the proposed research work
The hybrid approach in which CNN and SVM are working together can be used in our research and it can lead to improved accuracy. By studying this paper, we can make a model which can be more accurate than this and it will be very useful.

Brain Tumor Detection and Classification Using Convolutional Neural Network and Deep Neural Network [6]

Summary of the research item
This research paper mainly focuses on the concept of machine learning and CNN for the detection of brain tumors and the automated detection of brain tumors using MRI images. This research combines the neural networks and the CNN in order to achieve a high accuracy. The accuracy which was achieved by this research and its model was 96.03 percent.

Critical analysis of the research item
Strengths:
- The model was successful with a very high accuracy of 96 percent and it gives a deep concept of the CNN which was used to detect the brain tumor.
- It successfully differentiates between a healthy individual and a person with brain tumor.
- It detects the brain tumor early which can save lives of people.

Weaknesses:
- The research does not mention the main elements of the data set which are its size and the other main elements of the data set.
- There was no proper differentiation between the training data and the validation data.

Relationship to the proposed research work
This research is related to our research work as it gives a very higher accuracy in detecting the brain tumor and our work is also to classify between various brain diseases and it will be of great help in our own research.

Deep Learning Based Brain Tumor Detection and Classification [7]

Summary of the research item
The main thing discussed in this research is the detection of MRI. It mainly is the research of the development of two deep learning based approaches. There is a fast AI deep learning library. The study is conducted using a subset of the BRATS 2018 data set, containing 1,992 Brain MRI scans. The first model achieved the accuracy of 85.95 percent and the second model achieved the accuracy of 95.78 percent.
Critical analysis of the research item
Strengths:

- The concepts of deep learning are used which have shown very precise and accurate results.
- This research gives a comparison of two deep learning models and one is more accurate but we can say that both models achieved a very high accuracy.
- The Data set which was used is well known and publicly available.

Weaknesses:

- The research mentions the use of google Colab and tesla p100 GPU which is not available in all healthcare platforms.
- The Data set is a subset of BRATS 2018 Data set and because of that the accuracy is less.
- The research also does not show any Pre-processing steps.

Relationship to the proposed research work

This research paper is related to our research as our research is also about differentiation between various brain diseases and this model shows a high accuracy in differentiating between a healthy brain and the one with brain tumor so this research is related to our research.

REFERENCES

[1] Z. Jia and D. Chen, “Brain tumor identification and classification of mri images using deep learning techniques,” IEEE Access, 2020.

[2] M. Arbane, R. Benlamri, Y. Brik, and M. Djerioui, “Transfer learning for automatic brain tumor classification using mri images,” in 2020 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH), pp. 210–214, IEEE, 2021.

[3] M. Assam, H. Kanwal, U. Farooq, S. K. Shah, A. Mehmood, and G. S. Choi, “An efficient classification of mri brain images,” IEEE Access, vol. 9, pp. 33313–33322, 2021.

[4] N. H. Rajini and R. Bhavani, “Classification of mri brain images using k-nearest neighbor and artificial neural network,” in 2011 International conference on recent trends in information technology (ICRTIT), pp. 563–568, IEEE, 2011.

[5] M. O. Khairandish, M. Sharma, V. Jain, J. M. Chatterjee, and N. Jhanjhi, “A hybrid cnn-svm threshold segmentation approach for tumor detection and classification of mri brain images,” Irbm, vol. 43, no. 4, pp. 290–299, 2022.

[6] C. L. Choudhury, C. Mahanty, R. Kumar, and B. K. Mishra, “Brain tumor detection and classification using convolutional neural network and deep neural network,” in 2020 international conference on computer science, engineering and applications (ICCSEA), pp. 1–4, IEEE, 2020.

[7] N. M. Dipu, S. A. Shohan, and K. Salam, “Deep learning based brain tumor detection and classification,” in 2021 International conference on intelligent technologies (CONIT), pp. 1–6, IEEE, 2021.