Jurnal Infotel
Vol. 16, No. 3, August 2024, pp. 598–614.
doi:10.20895/infotel.v16i3.1175

Research Article

MRI-Based Brain Tumor Classification using ResNet-50 and Optimized Softmax Regression

Muhammad Nazeer Musa[1,*], Muhammad Bashar Sanusi[2], Phillip Odion[3] and Saifullahi Sadi Shitu[4]

[1,4]Cyber Security Department, Nigerian Defence Academy, Kaduna, Nigeria
[2,3]Computer Science Department, Nigerian Defence Academy, Kaduna, Nigeria

*Corresponding email: muhammadmusa2502@nda.edu.ng

Received: April 29, 2024; Revised: June 26, 2024; Accepted: June 27, 2024.

Abstract: Brain tumors represent a significant public health concern due to their high morbidity and mortality rates. Early and accurate detection is paramount for optimizing treatment efficacy. However, manual analysis of large-scale medical imaging datasets presents a substantial challenge, characterized by both time constraints and inherent susceptibility to human error. This has spurred the development of automated brain tumor classification systems using machine learning and deep learning techniques, but achieving high accuracy remains a persistent issue. This study addresses this challenge by proposing a hybrid brain tumor classification model that extracts hierarchical features from MRI scans using a pre-trained ResNet-50 and fuses into an optimized Logistic Regression classifier to enhance classification accuracy. Data augmentation is employed during training to improve generalization. The model is evaluated on an independent test set using metrics like accuracy, precision, recall, and F1-score. The deep feature fusion with the optimized Softmax regression technique on the test dataset achieved an overall classification accuracy of 96.3%, outperforming individual deep learning and machine learning models. Incorporating data augmentation further boosted performance, with 98.47% accuracy and 99% F1 score. The study demonstrates the potential of synergistically combining deep representation learning for accurate brain tumor classification from MRI scans. This framework shows promise for clinical deployment to support improved computer-aided diagnosis and treatment planning.

Keywords: brain Tumor, classification, hybrid learning, MRI, ResNet-50, Softmax Regression

---

1 Introduction

The pathology of brain tumors is one of the most common death problems and is considered an important priority for the healthcare society [1]. A brain tumor is a condition that causes poor brain function as a result of abnormal tissue growth in the brain [2]. Brain tumors come in various forms, with gliomas, meningiomas, and pituitary tumors being prime examples. Gliomas arise from supportive tissues within the brain, excluding nerves and blood vessels. In contrast, meningiomas form from the protective membranes surrounding the brain and spinal cord. Pituitary tumors, on the other hand, develop within the skull [3]. A crucial distinction lies in their malignancy: while meningiomas are typically benign, gliomas are often malignant. Even benign pituitary tumors can cause health issues. Accurate differentiation between these tumors throughout diagnosis and beyond is essential for optimal patient care [4]. The position of such tumors is determined using Magnetic Resonance Image (MRI). Accurate diagnosis of brain tumor type is crucial for crafting a recovery plan that minimizes life-threatening risks. This task traditionally involves visual analysis by trained professionals, but it's a time-consuming process and prone to human error [5, 6]. Machine learning-based automated image analysis offers a promising alternative to these concerns. Recent studies explored the use of machine learning methods like Support Vector Machines, K-Nearest Neighbors, and Random Forests for classifying medical images. These methods usually begin with a pre-processing stage to extract key features before performing classification [7].

Different feature extraction techniques exist, with the Gray-Level Co-occurrence Matrix (GLCM) being a common choice for capturing low-level features from images [8]. However, pinpointing the single most effective feature for peak accuracy remains a challenge. Deep Learning (DL) algorithms, particularly Convolutional Neural Networks (CNNs), are revolutionizing the field. Their popularity stems from their ability to handle massive datasets by passing data through multiple layers. Each layer progressively extracts features and transmits them to the next, creating a powerful feature hierarchy [9, 10]. Deep learning stands out for its ability to automatically learn features from data, eliminating the need for manual feature engineering [11]. However, DL algorithms require vast amounts of labeled training data, which is difficult to obtain. To address these difficulties, recent studies have used transfer learning to create hybrid models that combine the strength of pre-trained deep learning for feature extraction while requiring less data and applying machine learning for classification. The application of this method has proven to be efficient and to produce greater accuracy in brain tumor classification [12].

Transfer learning by feature extraction is effective when the data is small, as in the brain tumor dataset. However, even though several ML algorithms have been used on the extracted features for classification, none have explored the power of logistic or Softmax regression. According to [13], combining optimized logistic regression with techniques inspired by Newton's methods offers a compelling classification approach, particularly for datasets of moderate size. Studies comparing this method to Support Vector Machines (SVM) on twelve public datasets revealed comparable accuracy while achieving significantly faster processing times. The Logistic Regression (LR) and Softmax Regression (SR) approach provides direct prediction probabilities as well as rapid active learning of extracted images [14, 15]. Previous research in medical image analysis has faced limitations in terms of its reliance on manual feature engineering, which is time-consuming and prone to errors [7, 8]. Additionally, while deep learning approaches like CNNs show promise, they often require large datasets that are difficult to obtain in the medical field [11]. While transfer learning has partially addressed the data scarcity issue, many studies still employ basic classifiers for the final classification stage, leading to suboptimal performances [12]. Therefore, there is a need for methods that can effectively leverage pre-trained networks and optimize the final classification process to achieve high accuracy. This study proposes a novel approach that leverages the ResNet-50 architecture for feature extraction and an optimized Softmax regression model for classification. The knowledge transferred from large-scale image datasets makes ResNet-50 effectively address the data limitations inherent in the medical domain. Additionally, the optimized Softmax regression model is specifically tailored to the medical imaging context, enhancing classification accuracy beyond that of basic classifiers. The notable contributions of the proposed study include:

1. An empirically pre-trained ResNet-50 architecture for deep feature extraction and optimized Softmax regression for classification is presented to outperform previous state-of-the-art CNN architectures in terms of efficacy and efficiency.
2. The proposed hybrid model is trained on both augmented and unaugmented images and tested on four different classes of brain MRI images, namely glioma, meningioma, pituitary, and no tumor images, concluding that augmentation influences classifier accuracy.
3. Furthermore, a detailed comparative analysis is presented to evaluate the accuracy and efficiency of the proposed research study with previous state-of-the-art hybrid models on brain tumors.

2 Literature Review

Initially, automatic brain tumor classification relied on traditional machine learning techniques, which included pre-processing and feature extraction before making classifications. Traditional machine learning algorithms heavily rely on feature extraction for image classification tasks like brain tumor detection. These features can be categorized as either low-level (global) or high-level (local). Low-level features capture basic image properties like texture, intensity, and statistical measures (mean, standard deviation, etc.). Common examples include gray-level co-occurrence matrices, wavelength transforms, and Gabor features. High-level features, on the other hand, focus on specific details within an image, such as Fisher Vectors (FV), Scale-Invariant Feature Transforms (SIFT), and Bag-of-Words (BoW) [12].

Several studies have explored classical machine learning methods for brain tumor classification. For instance, one study used a Support Vector Machine (SVM) with first and second-order statistics to distinguish normal and abnormal brain MRIs [16]. Another employed Gray-Level Co-occurrence Matrix (GLCM) and wavelet transform-based approaches for tumor identification and classification [17]. While these low-level features effectively capture basic image characteristics, their ability to differentiate between tumors can be limited due to similarities in texture, shape, and size [18]. Other studies have used artificial neural networks and genetic algorithms for feature extraction and classification, but these approaches often require manual intervention and prior knowledge about the tumor's location [18, 19]. A key challenge with traditional machine learning for brain tumor classification lies in its focus on either local or global features. Additionally, these methods rely on manually designed features, which can be time-consuming and require domain expertise. This increases the likelihood of human error. When compared to traditional ML, deep learning, particularly the Convolutional Neural Network (CNN), eliminated the problem of handcrafted features that machine learning had [20,21]. This automatic feature extraction in CNN eliminates the need for manual feature engineering, a crucial step in traditional machine learning methods for tasks like image segmentation, classification, and object recognition [22,23].

In recent years, automatic brain tumor classification has gained popularity due to its automatic representation of pertinent characteristics and has outperformed other approaches in various medical imaging tasks [24]. Several studies employed artificial neural networks (ANNs) to distinguish between normal and malignant tumors. For instance, one study utilized a three-layer feed-forward ANN with a single hidden layer containing 50 nodes for classification [25]. More recent work has explored deep learning architectures for classifying brain tumors into different types. One such study focused on classifying meningioma, glioma, and pituitary tumors using T1-weighted contrast-enhanced MRI (CE-MRI) images [26]. To mitigate potential confusion from analyzing multiple image planes, the study used 989 axial slices from 191 patients and achieved an accuracy of 91.43% after applying data augmentation techniques. To address challenges like vanishing gradients and overfitting in multi-tumor classification, another study proposed a deep neural network model based on ResNet-50 with global average pooling [27]. This approach achieved a mean accuracy of 97.08% with data augmentation and 97.48% without, utilizing MRI images from three distinct brain tumor types. Meanwhile, [28] developed a three-class MRI brain tumor classification method using Levenberg-Marquardt for network construction, as well as a neural network with an effective training function, which yielded a 95.4% accuracy rate.

In a different study, a deep learning network was trained by [4] using 3-class brain MRI scans, employing a 10-fold cross-validation technique, achieving a classification accuracy of 95.40% for the original dataset and 96.56% for the augmented images dataset. To address the resource limitations and small size of medical image databases, transfer learning of pre-trained networks was used by later researchers. While [29] improved the MRI reconstruction by using the alternative direction method of multiplier method based on variable splitting technique, deep learning requires a significant amount of resources. This, together with the small size of medical image databases, poses a problem. Later researchers used the transfer learning of pre-trained networks to overcome this difficulty. [30] utilized a pre-trained network to classify the images into their appropriate categories. The fine-tuning method was used to assess the accuracy of the AlexNet, GoogleNet, and VGGNet models. For the original datasets, accuracies of 95.86%, 95.61%, and 95.42% were reported with training times of 42 minutes 36 seconds, 79 minutes 25 seconds, and 89 minutes 30 seconds, respectively. The researchers employed a five-fold cross-validation technique to thoroughly assess the performance of their integrated system. This evaluation demonstrated that the proposed model surpassed existing leading methods, achieving an impressive overall classification accuracy of 95.82%. [31] reported three alternative CNN models: (i) for detecting tumors, (ii) for classifying tumor types, and (iii) for determining tumor stages. The accuracy achieved in classifying tumors into three categories was reported as 92.66%. Another study [32] used CNN to extract features from brain MRI scans and combined them with SVM to classify medical images. The method was tested on an open dataset from Figshare that contained MRI images of three brain tumor types. The proposed model outperformed the previous state-of-the-art method with an overall classification accuracy of

95.82%, which was tested and evaluated using a five-fold cross-validation technique. Subsequent methods were able to add a "no tumor" class to the classification. [20] proposed employing a Multi-Scale CNN (MSCNN) architecture to scale deep learning in order to construct a robust classification model for brain tumor diagnosis. The suggested model was a multi-class classification solution for MRIs that categorizes them as glioma, meningioma, pituitary, or non-tumor. According to the experimental results, MCNN achieved 91.2% accuracy and a 91% F1-score, which were much higher than AlexNet and ResNet approaches. Some investigation proposed a series of three CNN models: one for tumor detection, another for classifying tumor types, and a third for determining the stages of different tumors [31]. This particular model achieved an accuracy of 92.66% in classifying tumors into three distinct categories. Another study combined CNN-extracted features with Support Vector Machines (SVM) for brain tumor classification [32]. This method, tested on a publicly available dataset of MRI images with three tumor types, surpassed previous methods with an overall accuracy of 95.82% using a five-fold cross-validation technique. Building upon this progress, subsequent research incorporated a "no tumor" class into the classification process [20]. This study proposed a Multi-Scale CNN (MSCNN) architecture to achieve robust classification for brain tumor diagnosis. The model categorized MRIs as glioma, meningioma, pituitary, or non-tumor, achieving an accuracy of 91.2% and an F1-score of 91%, surpassing the performance of AlexNet and ResNet architectures. [12] proposed the potential of combining deep features extracted from multiple CNN architectures (DenseNet-169, Shufflenet, and MnasNet) with a machine learning classifier (SVM with Radial Basis Function kernel) for brain tumor classification. This ensemble approach aimed to classify brain tumors into four categories (glioma, meningioma, pituitary, and no tumor) and achieved an accuracy of 93.7%. [11] investigated a transfer learning-based hybrid model for brain tumor classification. This method leveraged a fine-tuned GoogleNet architecture for feature extraction followed by pattern classification. The researchers evaluated the model's performance using a dataset containing a substantial number of MRI images categorized as glioma, meningioma, pituitary tumor, and normal brain tissue. The results showed that using Google-Net for feature extraction and adding SVM for classification improved the recognition accuracy to 98.1%.

Despite recent advancements in deep learning for brain tumor classification [11, 12, 20], significant challenges persist. While transfer learning has shown promise [30, 31], the full potential of combining fine-tuned state-of-the-art architectures like ResNet-50 with optimized classifiers remains unexplored, particularly for improving classification accuracy across multiple tumor types. Current approaches often struggle with limited medical imaging datasets [29, 32], and the impact of data augmentation on classification performance is not fully understood [4, 26]. Notably, previous research has largely overlooked the potential of logistic/softmax regression for this task [11, 20, 32], despite its simplicity and ability to effectively capture clear intrinsic features when combined with powerful feature extractors [14, 15]. This oversight represents a missed opportunity to leverage a computationally efficient classifier that could potentially enhance accuracy when paired with advanced feature extraction techniques. Therefore, this study aims to address these gaps by proposing a novel hybrid approach combining a fine-tuned ResNet-50 for feature extraction with an optimized Softmax Regression classifier. Our primary goal is to significantly enhance classification accuracy across various brain tumor types, even with limited data, while maintaining computational efficiency suitable for clinical applications.

3 Research Method

In this study, an MRI-based brain tumor classification model is developed using transfer learning by feature extraction technique to differentiate four different brain images. The research study followed a normal image recognition pipeline with preset phases, comparable to most classic recognition applications. The proposed technique applies ResNet-50 with optimized Softmax regression for the classification. Figure 1 depicts the hybrid process model employed in this study.

3.1 Dataset

In this study, we used a dataset that was publicly available on Kaggle, originally provided by [33]. The dataset included a total of 3264 MRI images categorized into four groups: glioma, meningioma, pituitary, and non-tumor, as depicted in Figure 2. The images used in this study were a combination of T1, T2, and Flair images. T1 images are generated by using short Time to Echo (TE) and long Repetition Time (RT), while T2 images are produced using long TE and RT. Flair images are similar to T2 images but have longer TE and RT times. These different image sequences generate contrast and brightness in the MRI scans based on the properties of the tissue being imaged. The dataset was randomly split into training and testing sets, with 80% of the data used for training and the remaining 20% used for testing. Each folder in the dataset was further divided into four subfolders based on the categories of brain tumors.

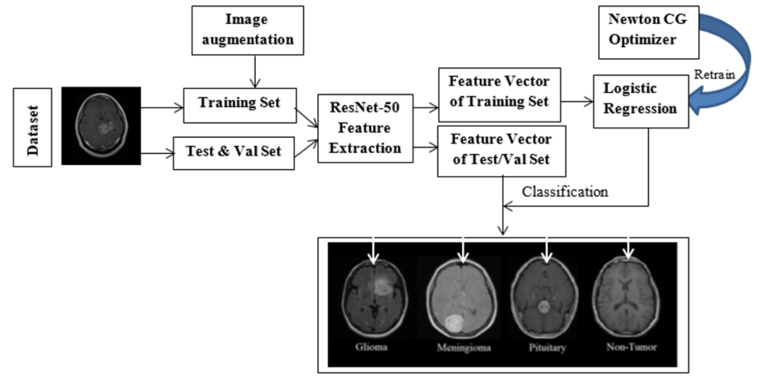

Figure 1: Proposed hybrid process model.

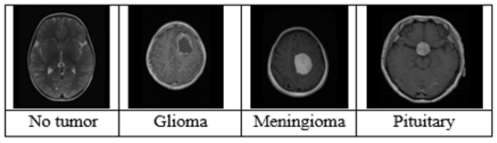

Figure 2: Class divisions of brain tumors from MRIs.

These folders contain MRIs of various tumor classes. Table 1 depicts the dataset's class distribution. 926 MRIs are from gliomas, 937 are from meningiomas, 500 are from non-tumor patients, and the remaining 901 are from pituitary patients.

Table 1: Class distribution of MRI brain images used

S/N	Class	Number of Images
1	Glioma	926
2	Meningioma	937
3	Non-tumor	500
4	Pituitary	901
	Total	3264

3.2 Image Augmentation

To increase the size of the training dataset, image augmentation was used due to the limited size of the dataset. Image augmentation is a technique that creates an artificial dataset by modifying the original images. Various image enhancement techniques such as shearing, contrast adjustment, horizontal flipping, rotation, zooming, and blurring were used for this purpose as in [34]. In this study, we used auto augmentation [35], which automates augmentation procedures for effective data augmentation by applying reinforcement learning as a search method to determine the optimum selections and ordering of these operations such that training a neural network provides the best validation accuracy. The generated dataset is then resized to 224×224 pixels, transformed to array format, and the pixel intensities in the input image are scaled to the $[-1, 1]$ range (via the convenience preprocess input function). Figure 3 depicts the axial view of the meningioma MRI image before and after augmentation.

3.3 Transfer Learning by Feature Extraction Using Pre-trained ResNet-50

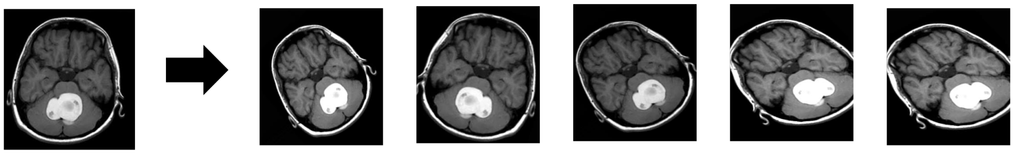

Figure 3: Axial view of meningioma tumor before and after augmentation.

Transfer learning refers to a technique in which a pre-trained model on a large benchmark dataset, such as ImageNet [36], can be used as a starting point for a different task with a smaller dataset, such as MRI data. This approach allows for a significant reduction in the training time required for deep learning models, as well as the need for a large dataset for model training. Instead of starting from scratch, transfer learning leverages the existing knowledge learned by the pre-trained model on large datasets, which can be beneficial for the target task [37]. In this particular study, we used a state-of-the-art network called ResNet-50, which was pre-trained on the ImageNet dataset. The ResNet-50 architecture was designed to overcome the vanishing gradient problem that was previously experienced by deep neural networks, making it well-suited for image recognition challenges. We chose to use ResNet-50 because of its advantages in image classification problems, as demonstrated experimentally by its superior performance on the COCO dataset [38]. By utilizing transfer learning and ResNet-50, the researchers were able to leverage the pre-trained knowledge of the model and fine-tune it for the specific task of classifying brain tumor MRI images. To match the input criteria of ResNet-50, the MRI images were resized and scaled down to 244×244. The network was then used as a feature extractor, where propagation was halted at a pre-defined layer, specifically the pooling layer before the fully connected layers. The resulting values were treated as a feature vector with an output shape of $7 \times 7 \times 2048$. These values were then flattened, resulting in a list of 100,352 values that served as our feature vector in the last conv block as shown inFigure 4. This process was repeated for the entire image dataset, generating a list of N vectors, each with 100,352 dimensions. The resulting batch of feature vectors was then saved to a CSV file, with the class as the first entry in each row. This process was repeated for all batches within each split of the dataset and finally, the label encoder was saved to disk. With these feature vectors, the Softmax regression model was optimized and trained to recognize new classes of MRI images. This technique proved to be particularly useful when dealing with small datasets, as CNN models generally perform better with larger datasets. Transfer learning enables the utilization of pre-trained models, reducing the need for large training datasets and minimizing training time.

3.4 Optimized Softmax Regression for Classification

LR is a supervised learning approach that uses the Sigmoid function on the input feature to identify target values. This Sigmoid function calculates the likelihood of making a classification judgment. The LR algorithm is most commonly used for binary classification, although it can also be adjusted for multiclass classification tasks. The improved version, called "Softmax regression," is based on the Softmax function, which provides a better approximation of the expected target class [39].

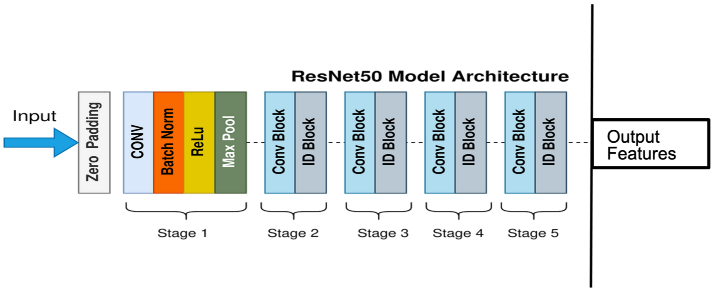

Figure 4: ResNet-50 for feature extraction.

SR (also known as multinomial LR) extends logistic regression to handle many classes. We assumed that the labels in LR were binary: $y^{(i)} \in 0, 1$ and we use such a classifier to distinguish between two classes. In the case of Softmax regression, it allows us to handle $y^{(i)} \in 1, \ldots, n$, where N is the number of classes. Recall that in logistic regression, we had a training set $(x^{(1)}, y^{(1)}), \cdots, (x^{(m)}, y^{(m)})$ of m labeled examples, where the input features are $x^{(i)} \in \mathbb{R}^k$. In the binary classification setting, the labels were $y^{(i)} \in 0, 1$. So our equation took the form of 1:

$$h_\theta(x) = \frac{1}{1 + \exp -\theta^{T_x}} \tag{1}$$

Also, the model parameters were trained to minimize the cost function as in **??**

$$labeleq : 2J(\theta) = -\left[\sum_{i=1}^{m} y^{(i)} \log h_\theta(x^{(i)}) + (1 - y^{(i)}) \log(1 - h_\theta(x^{(i)})\right] \tag{2}$$

The focus extends beyond binary classification rather, multi-class classification problems in the context of Softmax regression models. In such cases, the label variable y is not constrained to two values but can assume multiple discrete values across a broader range. Thus our training set $(x^{(1)}, y^{(1)}), \ldots, (x^{(m)}, y^{(m)})$ now have $y^{(i)} \in 1, 2, \cdots, N$. For example, in this study classification, we would have $N = 4$ different classes, i.e. there may exist 4 different classes that the input data could belong to. Given a previously unseen test input x, we want our hypothesis to estimate the probability that $P(y = n|x)$ for each value of $n = 1, \cdots, N$. i.e., the probability of the class label taking on each of the N different possible values. Thus, our equation will output a N-dimensional vector (whose elements sum to 1) giving N estimated probabilities. There is a loss function that expresses how close the classifier's output is to the target label.

The aim is to minimize the difference between the expected and actual output, which is represented by the loss function. However, if the selected features are not appropriate, the model may overfit the data, leading to an increase in the value of the corresponding coefficient. To avoid overfitting, regularization is used to penalize excessive weights. L2 (also known as Ridge) regularization was employed for the study, which adds a scaled sum of the squares of the coefficients to the objective function. A mathematical model known as a solver is used to minimize the loss function by combining regularization. Newton-cg, Liblinear, LBFGS, SAG, and SAGA are the solvers, with each having its advantages and disadvantages. Newton-cg was chosen because it applies to nonlinear classification. The solver's inverse is set as the C parameter, and while fine-tuning, the ideal value for C was chosen as 0.1.

The choice for LR in this paper is because of LR inherent strength in binary classification and its natural extension to multi-class problems through SR makes it a well-suited choice for the four-class brain tumor classification task. Additionally, Logistic Regression's computational efficiency and straightforward optimization are advantageous when dealing with high-dimensional features extracted by ResNet-50, particularly in time-sensitive medical applications. Furthermore, SR when optimized thrives with well-separated features, and the fine-tuned ResNet-50 model is designed to extract precisely such features from MRI images.

4 Results

This section delves into the experimental analysis conducted to assess the effectiveness of the proposed model. We first describe the computational environment used for the experiments. Next, we explored the various evaluation metrics employed to measure the model's performance. Following this, we presented the experimental results achieved by the model, considering both scenarios with and without data augmentation. Finally, to gain deeper insights into the model's effectiveness, we conducted a comparative analysis against other leading techniques discussed in previous literature. Through this comprehensive evaluation process, we aim to provide a detailed account of the experimental methodology and outcomes of our study.

4.1 Experimental Environment

A Python programming language is used to carry out the experiment. Libraries used are Keras and TensorFlow for feature extraction and PyTorch for data augmentation. Other tools and specifications are outlined in Table 2.

Table 2: System environment used for the model

System Specification	Description
Operating System	Windows 10
RAM	8 Gb
Processor	Core i5
Tool	Anaconda
Language	Python

4.2 Performance Evaluation

As indicated in Table 3, four performance evaluation metrics were employed to assess the performance of the proposed model: accuracy, precision, recall, and F1-measure. Classification Accuracy reflects the percentage of correctly classified examples. Precision indicates the proportion of true positives among the model's positive predictions. Conversely, Recall measures the model's ability to identify all positive instances within the dataset. The F1-Score (also known as F1-measure) combines precision and recall into a single metric, providing a balanced view of the model's performance when dealing with imbalanced datasets, where the ratio between positive and negative examples might be uneven. The confusion matrix is used to calculate these measures.

Table 3: Performance metrics used

S/N	Performance Metrics	Formula
1	Accuracy	$\frac{\text{(Total Number of Correct prediction)}}{\text{(Actual test data)}} \times 100$ or $\frac{(\mathrm{TP}+\mathrm{TN})}{(\mathrm{TP}+\mathrm{TNFP}+\mathrm{FN})}$
2	Precision	$\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}$
3	Recall	$\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}$
4	F1-Measure	$\frac{(2 \times \text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})}$

4.3 Experimental Results and Discussion

During the evaluation phase, the models that were developed during the training phase are used in making predictions on data from the evaluation set. The classification algorithm utilizes the feature vectors that were extracted from the CSV file to perform classification. To assess the model's effectiveness, performance metrics such as accuracy, F1 score, recall, and precision are used. The results are analyzed and presented using a classification report and a confusion matrix. Specifically, the classification report displays various performance metrics while the confusion matrix details the correctly classified instances alongside the misclassified examples for each class. Figure 5 and Figure 6 depict the model's output confusion matrix without and with data augmentation respectively.

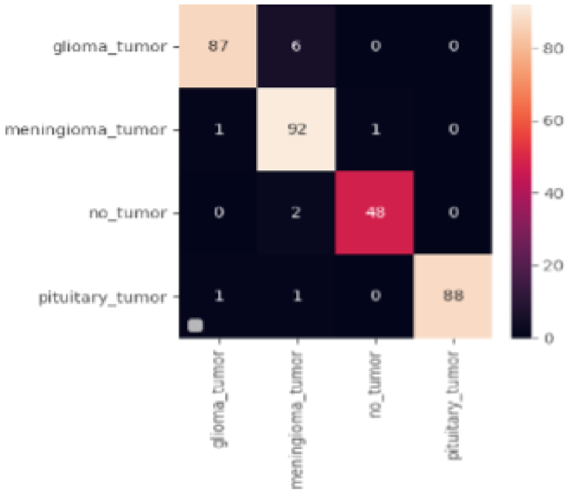

Figure 5: Confusion matrix of the proposed model without data augmentation.

The classification report in Table 4, presents the performance measures used in the study. The results indicated that using image augmentation improves the model's performance significantly. The accuracy, precision, recall, and F1-score obtained with data augmentation were 98.47%, 99%, 98%, and 99%, respectively. These values are superior to those obtained without data augmentation, which were 96.3%, 97%, 96%, and 96%, respectively.

This improvement is particularly evident in the reduction of misclassified tumor instances, especially for the glioma type, as observed in the confusion matrices in Figure 5 and Figure 6.

5 Discussion

The experimental results demonstrate the effectiveness of the proposed deep learning model for brain tumor classification from MRI images. Without any data augmentation, the model achieved an overall accuracy of 96.3%, precision of 97%, recall of 96%, and F1-score of 96%. These performance metrics indicate the model's strong ability to correctly classify different brain tumor types from the input image data.

Table 4: Performance comparison of optimized logistic regression with and without data augmentation

Metric	Class	With Augmentation	Without Augmentation
Accuracy	Overall	98.47%	96.33%
Precision	Glioma tumor	0.99	0.97
	Meningioma tumor	0.96	0.94
	No tumor	1.00	0.97
	Pituitary tumor	1.00	0.99
Recall	Glioma tumor	0.99	0.98
	Meningioma tumor	0.99	0.93
	No tumor	0.98	0.98
	Pituitary tumor	0.98	0.97
F1-Score	Glioma tumor	0.99	0.97
	Meningioma tumor	0.97	0.94
	No tumor	0.99	0.97
	Pituitary tumor	0.99	0.98

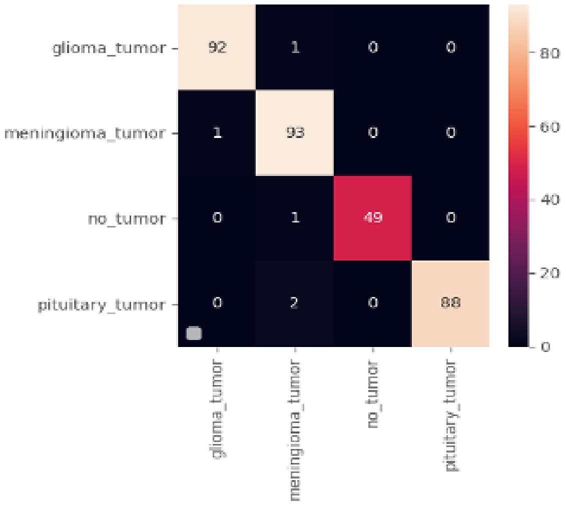

Figure 6: Confusion matrix of the proposed model with data augmentation.

However, the results were further improved by incorporating data augmentation techniques during training. With augmented data, the model's accuracy increased to 98.47%, with 99% precision, 98% recall, and 99% F1-score. Data augmentation helps expose the model to more variations of the training data. This increase in data diversity acts as a regularizer to reduce overfitting and improve the model's generalization capability.

Analyzing the confusion matrices provides additional insights. Without augmentation (Fig. 4b), the model struggled the most with glioma tumors, incorrectly classifying 6 cases. After data augmentation (Fig. 5b), the total number of misclassified cases dropped from 12 to only 5, significantly reducing errors across all tumor categories. This reinforces how augmentation helped the model learn more robust and discriminative features.

To evaluate the impact of our optimization approach, we conducted a comparative analysis between standard Softmax Regression and our optimized version. The standard Softmax Regression without optimization achieved an accuracy of 97.8%, with two misclassifications in the ′no tumor′ category. In contrast, our optimized Softmax Regression model achieved a higher accuracy of 98.4%. This improvement, while seemingly modest, is statistically significant and crucial in clinical applications where misdiagnoses can have severe consequences.

While the proposed deep feature fusion on an optimized Softmax regression approach achieved state-of-the-art performance for brain tumor classification, certain limitations should be acknowledged.

First, the current work focused solely on tumor type classification. However, for comprehensive decision support, an integrated system incorporating additional clinically relevant tasks like tumor segmentation, grading, and survival prediction would be highly valuable. Extending the framework to these complementary tasks is an important direction for future research. Secondly, while data augmentation improved performance, more advanced augmentation strategies tailored for medical imaging could potentially extract further gains in accuracy and robustness. Techniques like elastic deformations, noise injections, and blended augmentations aimed at simulating real-world variations may be beneficial. Finally, the deep learning model used off-the-shelf architectures pre-trained on natural images. Exploring neural architecture search customized for medical imaging data or training models from scratch could uncover more optimal representations specific to the brain tumor domain.

Despite these limitations, the current study provides a strong baseline demonstrating the synergies between deep learning and machine learning for this high-stakes clinical application. Addressing the outlined limitations through subsequent research can pave the way for real-world deployment and impact.

5.1 Comparative Analysis with Previous Study

Table 5 presents a comparison of the accuracy of the proposed approach with that of other similar methods. The table shows the accuracy of four different methods in addition to the proposed method. The results show that the suggested model outperforms the other four models in terms of accuracy, with GoogleNet + SVM being the second-best model. The superior performance of the proposed model can be attributed to the fine-tuned ResNet-50′s ability to extract relevant features and the ability to optimize the SR algorithm with appropriate parameters for high-dimensional classification tasks. Therefore, the proposed model is highly efficient in terms of performance.

Table 5: Comparison with Previous Study

Ref	Proposed Method	Accuracy
Yazdan et al. (2022)	Multi-Scale CNN (MSCNN)	91.2
Kang et al. (2021)	(DenseNet-169 + Shufflenet + MnasNet) feature + SVM RBF	93.7
Rasool et al. (2022)	GoogleNet + Fine tuning	93.1
Rasool et al. (2022)	GoogleNet + SVM	98.1
Proposed Method without Optimization	ResNet50 + SR	97.88
Proposed Method	ResNet50 + Optimized SR	98.4

6 Conclusion

This study introduced a novel hybrid deep-learning classification method that utilizes transfer learning through feature extraction. This methodology leverages transfer learning by utilizing a pre-trained ResNet-50 model for feature extraction with optimized SR to identify and categorize brain tumor types in MRI images. The proposed approach automates the classification process of MRI brain image datasets into four classes, including meningioma, glioma, pituitary tumor, and no tumor. The experimental findings indicated that the suggested method outperforms existing techniques achieving an impressive accuracy of 98.4%. These findings suggest that this approach has the potential to be a valuable diagnostic tool for radiologists, assisting them in identifying suspicious regions within the brain. While the initial performance is encouraging, further research is necessary to reduce the model size. This optimization would be crucial for deploying the model in real-time medical diagnosis systems.

References

[1] S. Abbasi and F. Tajeripour, "Detection of brain tumor in 3d mri images using local binary patterns and histogram orientation gradient," *Neurocomputing*, vol. 219, pp. 526–535, 2017.

[2] M. K. Abd-Ellah, A. I. Awad, A. A. Khalaf, and H. F. Hamed, "Two-phase multi-model automatic brain tumour diagnosis system from magnetic resonance images using convolutional neural networks," *EURASIP Journal on Image and Video Processing*, vol. 2018, no. 1, pp. 1–10, 2018.

[3] D. N. Louis, A. Perry, G. Reifenberger, A. Von Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues, and D. W. Ellison, "The 2016 world health organization classification of tumors of the central nervous system: a summary," *Acta neuropathologica*, vol. 131, pp. 803–820, 2016.

[4] M. M. Badža and M. Č. Barjaktarović, "Classification of brain tumors from mri images using a convolutional neural network," *Applied Sciences*, vol. 10, no. 6, p. 1999, 2020.

[5] N. Alalwan, A. Abozeid, A. A. ElHabshy, and A. Alzahrani, "Efficient 3d deep learning model for medical image semantic segmentation," *Alexandria Engineering Journal*, vol. 60, no. 1, pp. 1231–1239, 2021.

[6] P. M. de Azevedo-Marques and J. R. Ferreira, "Medical image analyst: A radiology career focused on comprehensive quantitative imaging analytics to improve healthcare," *Academic Radiology*, vol. 29, no. 1, p. 170, 2022.

[7] A. Sawner, S. Shukla, V. Roy, S. Shukla, and S. Kapoor, "4 way classification of brain tumor using shallow convolution neural network.." EasyChair Preprint 7084, EasyChair, 2021.

[8] Q. T. Ostrom, H. Gittleman, P. M. De Blank, J. L. Finlay, J. G. Gurney, R. McKean-Cowdin, D. S. Stearns, J. E. Wolff, M. Liu, Y. Wolinsky, *et al.*, "American brain tumor

association adolescent and young adult primary brain and central nervous system tumors diagnosed in the united states in 2008-2012," *Neuro-oncology*, vol. 18, no. suppl_1, pp. i1–i50, 2016.

[9] W. Yafooz, A. Alsaeedi, R. Alluhaibi, and M. E. Abdel-Hamid, "Enhancing multi-class web video categorization model using machine and deep learning approaches," *Int. J. Electr. Comput. Eng*, vol. 12, p. 3176, 2022.

[10] S. Ben Atitallah, M. Driss, W. Boulila, A. Koubaa, and H. Ben Ghezala, "Fusion of convolutional neural networks based on dempster–shafer theory for automatic pneumonia detection from chest x-ray images," *International Journal of Imaging Systems and Technology*, vol. 32, no. 2, pp. 658–672, 2022.

[11] M. Rasool, N. A. Ismail, W. Boulila, A. Ammar, H. Samma, W. M. Yafooz, and A.-H. M. Emara, "A hybrid deep learning model for brain tumour classification," *Entropy*, vol. 24, no. 6, p. 799, 2022.

[12] J. Kang, Z. Ullah, and J. Gwak, "Mri-based brain tumor classification using ensemble of deep features and machine learning classifiers," *Sensors*, vol. 21, no. 6, p. 2222, 2021.

[13] M. Maalouf, T. B. Trafalis, and I. Adrianto, "Kernel logistic regression using truncated newton method," *Computational management science*, vol. 8, pp. 415–428, 2011.

[14] Y. Jiao, J. Yuan, Y. Qiang, and S. Fei, "Deep embeddings and logistic regression for rapid active learning in histopathological images," *Computer Methods and Programs in Biomedicine*, vol. 212, p. 106464, 2021.

[15] C.-M. L. Eduardo, V.-P. N. Enrique, and M.-D. R. Javier, "Heterogeneous processing for estimating the parameters of a binary logistic regression model using the fisher scoring algorithm," *IAENG Internaitonal Journal of Computer Science*, vol. 48, no. 2, pp. 1035–1039, 2021.

[16] H. Selvaraj, S. T. Selvi, D. Selvathi, and L. Gewali, "Brain mri slices classification using least squares support vector machine," *International journal of intelligent computing in medical sciences & image processing*, vol. 1, no. 1, pp. 21–33, 2007.

[17] P. John *et al.*, "Brain tumor classification using wavelet and texture based neural network," *International Journal of Scientific & Engineering Research*, vol. 3, no. 10, pp. 1–7, 2012.

[18] Z. Ullah, M. U. Farooq, S.-H. Lee, and D. An, "A hybrid image enhancement based brain mri images classification technique," *Medical hypotheses*, vol. 143, p. 109922, 2020.

[19] M. Sharif, U. Tanvir, E. U. Munir, M. A. Khan, and M. Yasmin, "Brain tumor segmentation and classification by improved binomial thresholding and multi-features selection," *Journal of ambient intelligence and humanized computing*, pp. 1–20, 2024.

[20] S. A. Yazdan, R. Ahmad, N. Iqbal, A. Rizwan, A. N. Khan, and D.-H. Kim, "An efficient multi-scale convolutional neural network based multi-class brain mri classification for samd," *Tomography*, vol. 8, no. 4, pp. 1905–1927, 2022.

[21] M. Turuk, R. Sreemathy, S. Kadiyala, S. Kotecha, and V. Kulkarni, “Cnn based deep learning approach for automatic malaria parasite detection,” *IAENG Int. J. Comput. Sci*, vol. 49, no. 3, 2022.

[22] D. A. Forsyth, J. L. Mundy, V. di Gesú, R. Cipolla, Y. LeCun, P. Haffner, L. Bottou, and Y. Bengio, “Object recognition with gradient-based learning,” *Shape, contour and grouping in computer vision*, pp. 319–345, 1999.

[23] S. Wang, D. Liu, Z. Yang, C. Feng, and R. Yao, “A convolutional neural network image classification based on extreme learning machine,” *IAENG Int. J. Comput. Sci*, vol. 48, no. 3, pp. 799–803, 2021.

[24] S. A. Agnes, J. Anitha, S. I. A. Pandian, and J. D. Peter, “Classification of mammogram images using multiscale all convolutional neural network (ma-cnn),” *Journal of medical systems*, vol. 44, no. 1, p. 30, 2020.

[25] T. Rajesh and R. S. M. Malar, “Rough set theory and feed forward neural network based brain tumor detection in magnetic resonance images,” in *International Conference on Advanced Nanomaterials & Emerging Engineering Technologies*, pp. 240–244, IEEE, 2013.

[26] B. S. Gimi and A. Krol, “Biomedical applications in molecular, structural, and functional imaging,” in *Proc. of SPIE Vol*, vol. 12036, pp. 1203601–1, 2022.

[27] I. Wahlang, A. K. Maji, G. Saha, P. Chakrabarti, M. Jasinski, Z. Leonowicz, and E. Jasinska, “Brain magnetic resonance imaging classification using deep learning architectures with gender and age,” *Sensors*, vol. 22, no. 5, p. 1766, 2022.

[28] A. Biswas and M. S. Islam, “Brain tumor types classification using k-means clustering and ann approach,” in *2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)*, pp. 654–658, IEEE, 2021.

[29] Z. Luo, Z. Zhu, and B. Zhang, “An scadtv nonconvex regularization approach for magnetic resonance imaging,” *IAENG International Journal of Computer Science*, vol. 48, no. 4, pp. 1005–1012, 2021.

[30] A. Rehman, S. Naz, M. I. Razzak, F. Akram, and M. Imran, “A deep learning-based framework for automatic brain tumors classification using transfer learning,” *Circuits, Systems, and Signal Processing*, vol. 39, no. 2, pp. 757–775, 2020.

[31] E. Irmak, “Multi-classification of brain tumor mri images using deep convolutional neural network with fully optimized framework,” *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, vol. 45, no. 3, pp. 1015–1036, 2021.

[32] S. Deepak and P. Ameer, “Automated categorization of brain tumor from mri using cnn features and svm,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 8, pp. 8357–8369, 2021.

[33] S. Bhuvaji, A. Kadam, P. Bhumkar, S. Dedge, and S. Kanchan, “Brain tumor classification (mri) | kaggle,” 2020.

[34] M. Musa, S. Sengupta, and Y. Chen, "Mri-compatible soft robotic sensing pad for head motion detection," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 3632–3639, 2022.

[35] E. D. Cubuk, B. Zoph, D. Mane, V. Vasudevan, and Q. V. Le, "Autoaugment: Learning augmentation strategies from data," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 113–123, 2019.

[36] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[37] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[39] H. Sharma and A. S. Jalal, "An improved attention and hybrid optimization technique for visual question answering," *Neural Processing Letters*, vol. 54, no. 1, pp. 709–730, 2022.