

Brain Tumor classification and Detection by Integrating ML Classifiers with Xception model for Magnetic Resonance imaging

Vaishnavi Kukkala
Department of Data Science
University of New Haven
vkukk2@unh.newhaven.edu

Sayed Khaled
Department of Data Science
University of New Haven
ksayed@unh.newhaven.edu

Abstract—Brain tumors exhibit diverse morphological and histological characteristics, making their diagnosis and treatment highly complex. Traditional diagnostic methods primarily rely on medical image analysis, which can be prone to subjectivity and interpretation errors. In recent years, machine learning (ML) and deep learning (DL) techniques have emerged as powerful tools for enhancing the accuracy and efficiency of brain tumor detection. Developing reliable and automated diagnostic models is essential for improving clinical decision-making and patient outcomes. This study proposes a brain tumor detection framework utilizing the Xception architecture, a deep convolutional neural network known for its advanced feature extraction capabilities. The model is trained on a publicly available MRI dataset, incorporating data augmentation techniques to improve generalization. Additionally, deep features extracted from the Xception model are leveraged to train various machine learning classifiers, including Support Vector Machines (SVM), Decision Trees (DT), k-Nearest Neighbors (KNN), Random Forest, and Logistic Regression, to further enhance classification performance. The experimental results demonstrate that the proposed model achieves an accuracy of 98% on the primary dataset and 94.7% when tested on a larger, independent dataset, highlighting its strong generalization ability. It further improved to 99.91% accuracy with SVM classifier. These findings indicate that integrating deep learning with traditional ML classifiers can provide a robust and interpretable approach for brain tumor detection. The proposed methodology offers a promising tool to assist medical professionals in early diagnosis, potentially improving patient care and treatment outcomes.

Keywords—Xception, ML Classifiers, Grad, Deep Learning.

I. INTRODUCTION

The brain is the central organ of the human nervous system, responsible for processing sensory information, coordinating motor functions, and regulating essential physiological processes. Given its critical role, the presence of brain tumors poses a significant health threat, potentially impairing cognitive and neurological functions. Brain tumors can be benign or malignant, with the latter often leading to life-threatening complications if not detected and treated promptly. Accurate and early diagnosis is crucial for effective treatment planning and improving patient survival rates. However, brain tumor diagnosis remains a major challenge in modern medicine due to the complexity of tumor structures, variations in their appearance, and the subjective nature of traditional diagnostic methods.

Medical imaging, particularly Magnetic Resonance Imaging (MRI), plays a fundamental role in brain tumor

detection, offering high-resolution visualization of soft tissues without ionizing radiation. With the advancements in artificial intelligence (AI), machine learning (ML), and deep learning (DL) techniques, automated tumor detection systems have shown great potential in assisting radiologists by improving diagnostic accuracy and reducing human error [4]. Among these, DL-based models have demonstrated remarkable success in medical image analysis, learning intricate patterns from labeled data to differentiate between normal and abnormal brain structures.

Convolutional Neural Networks (CNNs) have emerged as powerful tools for analyzing MRI scans due to their ability to extract hierarchical features from complex imaging data. Specifically, deep CNN architectures like Xception have exhibited superior performance in image classification tasks. Xception, an extension of the Inception model, utilizes depthwise separable convolutions to enhance computational efficiency and feature extraction capabilities. Its application in medical imaging has gained attention due to its ability to capture fine-grained details in tumor images, leading to improved classification accuracy.

Despite the success of deep learning models, generalization remains a critical concern. Many DL-based tumor classification models perform well on the datasets they are trained on but struggle when applied to new, unseen data. This issue arises due to limited training datasets, overfitting, and dataset bias. To address this challenge, this study proposes a dual approach: leveraging the Xception model for feature extraction and employing machine learning classifiers trained on deep features to improve classification performance. By training the model on a smaller dataset and evaluating it on a larger, more diverse dataset, we aim to assess its ability to generalize effectively to new cases.

The proposed framework integrates the Xception architecture, pre-trained on ImageNet, as a feature extractor. Additional layers, including fully connected (Dense) layers and Dropout layers, are incorporated to enhance classification performance and mitigate overfitting. The extracted deep features are further utilized to train ML classifiers such as Support Vector Machines (SVM), Random Forest, and Gradient Boosting models to analyze their effectiveness in tumor classification. This hybrid approach aims to leverage the strengths of both deep learning and traditional machine learning methods to achieve more reliable and interpretable results.

The findings of this study suggest that combining deep learning with ML classifiers can enhance brain tumor detection, providing a practical solution for real-world clinical applications. By evaluating the model's performance on a

larger dataset, this research contributes to understanding the generalization capabilities of AI-driven diagnostic systems in neurology and oncology.

The remainder of this paper is structured as follows: Section II presents a comprehensive review of existing literature on brain tumor diagnosis using ML and DL techniques. Section III details the methodology, focusing on the implementation of the Xception model and ML classifiers. Section IV presents the experimental results, along with analysis and discussion. Finally, Section V concludes the study by summarizing key findings and highlighting future research directions.

II. LITERATURE REVIEW

Early detection and classification of brain tumors are critical in preventing life-threatening complications. In recent years, significant research has been conducted to leverage artificial intelligence (AI) techniques—particularly machine learning (ML) and deep learning (DL)—to enhance brain tumor identification from medical imaging, particularly MRI scans. These advancements offer promising prospects for improving diagnostic accuracy, efficiency, and accessibility.

Deep learning models, particularly Convolutional Neural Networks (CNNs), have been widely applied in medical image classification. Disci R. [1] explored the effectiveness of CNN architectures such as Xception, MobileNetV2, InceptionV3, ResNet50, VGG16, and DenseNet121 in brain tumor detection. Among these, Xception, an advanced variant of the Inception model incorporating depthwise separable convolutions, demonstrated superior accuracy and efficiency. The study highlights the transformative potential of deep learning in medical imaging, suggesting that its integration into clinical workflows could enhance diagnostic precision. However, challenges persist, particularly in improving recall for Glioma and Meningioma tumors, as well as addressing the interpretability concerns of deep learning models. Future research should focus on increasing dataset diversity, enhancing model explainability, and validating performance in real-world clinical environments.

Similarly, Philip J.J. [2] investigated the potential of the Xception model for automated brain tumor detection. The study demonstrated that, when fine-tuned appropriately, Xception could assist medical professionals in making accurate and timely diagnoses. However, the model's performance is highly dependent on the quality and diversity of training data. Additionally, the computational demands of deep learning models like Xception pose challenges for deployment in resource-constrained settings. Further research is required to improve model efficiency and ensure robustness across diverse clinical scenarios.

Expanding on this, Yoon S. [3] introduced a hybrid ensemble model integrating Xception with parallel deep CNNs to improve classification performance. The study demonstrated that ensemble methods could enhance accuracy, making them valuable tools for automated tumor classification. However, the increased computational complexity of hybrid models may limit their applicability in real-time or resource-limited clinical environments. Moreover, model generalizability remains a concern, as reliance on specific datasets could hinder performance in varied clinical settings.

Beyond Xception-based approaches, Vimala B. [5] emphasized the necessity of accurate brain tumor classification for timely diagnosis and treatment. Traditional diagnostic methods are often subjective and time-consuming, necessitating automated systems for precise tumor classification. The study demonstrated the potential of fine-tuned EfficientNet models in achieving high classification accuracy, further supporting the role of deep learning in assisting radiologists with reliable tumor identification.

The 2019 study [6] by M. Siar and M. Teshnehlab, titled "Brain Tumor Detection Using Deep Neural Network and Machine Learning Algorithm," explores the application of deep learning techniques for the detection of brain tumors using medical imaging. The integration of CNNs with feature extraction techniques demonstrated the potential to improve the accuracy of brain tumor detection, offering a promising direction for developing automated diagnostic tools in medical imaging. However, the study does not extensively address the computational requirements for training and deploying such deep learning models, which could impact their practical application in clinical settings.

The 2024 study by Anantharajan et al., [8] titled "MRI Brain Tumor Detection Using Deep Learning and Machine Learning Approaches," investigates the application of advanced computational techniques to improve the accuracy and efficiency of brain tumor detection using Magnetic Resonance Imaging (MRI). The researchers developed a hybrid approach that integrates both deep learning and traditional machine learning techniques to enhance tumor detection accuracy. They utilized Convolutional Neural Networks (CNNs) for automatic feature extraction from MRI images, capitalizing on CNNs' ability to learn hierarchical representations of visual data. The model's performance is highly dependent on the quality and size of the training dataset, and the study was conducted using a specific dataset, which may limit the generalizability of the findings. Additionally, the computational complexity associated with training deep learning models could pose challenges for real-time clinical applications.

Further advancements in hybrid methodologies were introduced by Rasheed et al. [9], who incorporated attention mechanisms into CNN-based MRI classification to improve feature extraction and classification accuracy. The study evaluated the model against pre-trained architectures such as Xception, ResNet50V2, DenseNet201, ResNet101V2, and DenseNet169, demonstrating superior performance in identifying common brain tumor types. The incorporation of attention mechanisms allowed the model to focus on relevant regions within MRI scans, improving interpretability and diagnostic confidence.

Another notable approach was introduced by Zahoor et al. [10], who proposed a two-phase deep learning framework combining deep boosted feature space with ensemble classifiers (DBFS-EC) for enhanced brain tumor detection. Their method integrated deep convolutional feature extraction with Histogram of Oriented Gradients (HOG) descriptors, using machine learning classifiers for tumor classification. This hybrid features fusion approach demonstrated improved diagnostic accuracy, highlighting the effectiveness of combining CNN-based feature extraction with traditional ML classifiers.

Advancements in machine learning have significantly improved brain tumor detection techniques, showcasing how automated models can analyze MRI images for accurate classification of brain tumors (Fig. 1). Several studies have explored deep transfer learning to enhance brain tumor classification, leveraging morphological features and texture cues to improve glioma identification accuracy. Researchers [7] have also proposed novel CNN architectures for precise tumor segmentation, addressing key challenges in brain tumor diagnosis. However, despite these promising advancements, challenges such as data diversity, model interpretability, computational efficiency, and real-world applicability remain crucial areas for future research.

Furthermore, researchers have investigated a hybrid approach that combines deep learning-based feature extraction with traditional ML classifiers. In this method, deep features extracted from CNNs serve as input to ML models, allowing for effective classification while optimizing computational resources. This strategy is particularly beneficial in scenarios where high-end computing power is limited, providing a balance between performance and efficiency.

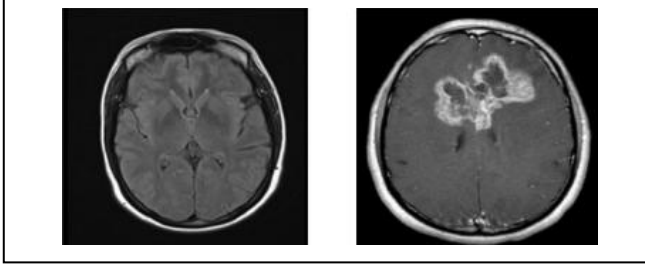


Fig. 1. Normal brain Vs. Tumor Brain

III. METHODOLOGY

A key challenge in medical imaging analysis is the detection of brain tumors, which are abnormal growths within the brain. This study presents a framework for brain tumor identification that integrates machine learning algorithms with advanced image processing techniques. The methodology consists of several essential stages, including feature extraction, model training, and evaluation, ensuring a systematic approach to tumor classification. The overall workflow of the proposed methodology is illustrated in Figure 2.

3.1 Dataset Collection

We employed two separate datasets obtained from Kaggle: one designated for training and the other for testing. The training dataset (1) comprises 253 MRI images classified into two categories: "Tumor" and "No Tumor." Similarly, the testing dataset (2) contains 3,060 MRI images, also categorized as "Tumor" and "No Tumor." These datasets were utilized to evaluate the effectiveness of our model in brain tumor detection.

3.2 Data Preprocessing

The data preparation process begins with collecting data from the given dataset (Fig. 2). This includes accessing MRI scans or other imaging modalities from individuals suspected of having brain tumors.

- To enhance image quality, preprocessing techniques such as noise reduction, normalization, and standardization of image dimensions of 224×224 pixels are applied.
- After loading the data and normalizing intensity levels, the dataset is split into training and validation sets for model development and evaluation. Images are rescaled by $1/255.0$ to convert pixel values from $[0-255] \rightarrow [0-1]$ for better model convergence.
- Label Encoding was performed to convert binary labels (0 for "no", 1 for "yes") to one-hot encoding.
- Additionally, data augmentation techniques like rotation, vertical and horizontal flipping, shearing transformation, filling, brightness and zooming are used to improve the model's generalization capability. Augmented images ensure a more diverse training set, improving the model's performance on unseen data.

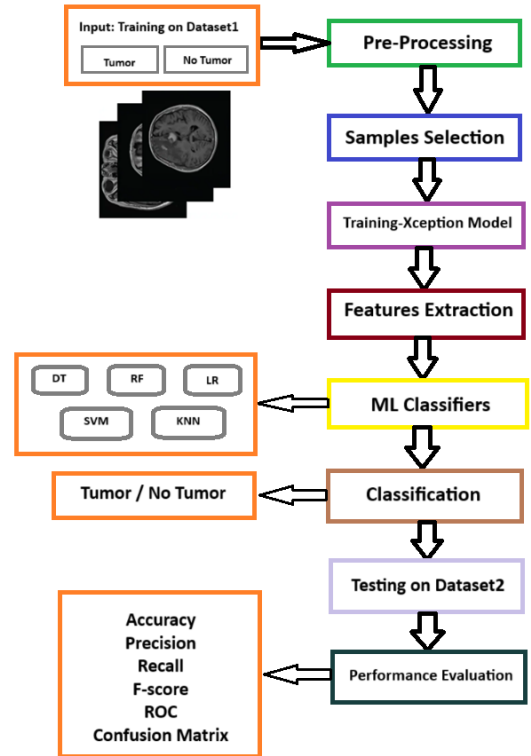


Fig. 2. Stages of the proposed methodology

3.2 Model Architecture & Training

The **Xception model**, introduced by **François Chollet**, is a deep convolutional neural network designed for image classification and object detection tasks [18]. It builds upon the **Inception architecture** by replacing traditional inception modules with **depthwise separable convolutions**, enhancing computational efficiency while maintaining high accuracy.

The Xception architecture, a convolutional neural network (CNN), is employed for model development in this work. This architecture is pre-trained on the ImageNet dataset, which contains over a million images, Xception has learned a diverse set of visual features, making it well-suited for various classification tasks, including medical imaging applications enhancing its performance across various image classification tasks. Due to its improved efficiency, the Xception model is considered a suitable choice for brain tumor detection. Here, the Xception model is fine-tuned on our dataset. We replace the top layers with a custom classification head consisting of fully connected layers, dropout, and softmax activation function.

3.2.1 Xception Architecture: The Xception architecture, which stands for Extreme Inception, is a convolutional neural network (CNN) designed to enhance efficiency and reduce parameters by leveraging depthwise separable convolutions. Inspired by the Inception architecture, Xception effectively captures intricate patterns and hierarchical features in images while maintaining computational efficiency.

3.2.2 Stacking Layers: The model is improved by incorporating additional layers, including dense and dropout layers. These layers refine the pre-trained model, enhancing its ability to effectively identify brain tumors.

3.2.3 Dropout Layers: Dropout layers are integrated into the architecture to minimize overfitting and enhance the model's generalization capabilities. This regularization technique involves randomly deactivating a portion of neurons during training, helping the model become more robust and less reliant on specific features.

3.2.4 Dense Layers: Dense layers are incorporated into the model to facilitate the final classification process. Alongside dropout layers, these fully connected layers utilize the information learned from previous layers to generate accurate predictions for the final output.

In figure 3; The architecture leverages the Xception model as a feature extractor, utilizing its depthwise separable convolutions to efficiently learn spatial patterns from images. Instead of flattening the feature maps, a GlobalAveragePooling2D layer condenses them into a single 2048-dimensional vector, reducing parameters and minimizing overfitting. The classifier consists of a fully connected (Dense) layer with 64 neurons and ReLU activation, followed by a Dropout layer to enhance

generalization by randomly deactivating neurons during training. The final Dense(2) layer with softmax activation maps the extracted features to class probabilities, making it suitable for binary or multi-class classification. This setup ensures a balance between efficiency, performance, and generalization, with minimal trainable parameters while retaining the powerful feature extraction of Xception.

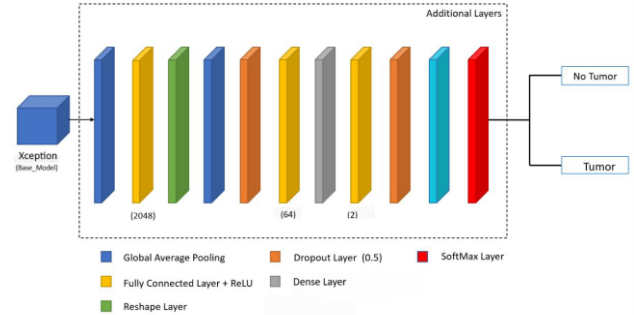


Fig. 3. Layered architecture of Proposed System

3.3 Machine Learning Classifiers

Feature Extraction for ML Classifiers: Instead of using raw pixel values, we extracted deep features from the penultimate layer of the Xception model. These feature vectors serve as input to ML classifiers, including SVM, Decision Trees, KNN, Random Forest, and Logistic Regression.

Support Vector Machine (SVM) is a powerful supervised learning algorithm that was introduced in 1992 by **Vladimir Vapnik** and his team at AT&T Bell Labs [11]. SVM is widely used for both regression and classification tasks, with significant applications in fields such as **text classification, handwriting recognition, medical diagnostics, pattern recognition, and bioinformatics** [12]. In the context of brain tumor classification, SVM is particularly advantageous due to its ability to **identify an optimal hyperplane that maximizes the margin** between different classes, reducing classification errors and improving generalization performance [13]. The kernel trick enables SVM to operate in **higher-dimensional spaces**, making it highly effective in cases where tumor features are not linearly separable. In this study, SVM is utilized to classify different categories of brain tumors using deep features extracted from the Xception model, aiming to enhance classification performance compared to other classifiers.

Random Forest (RF), proposed by Leo Breiman, is a widely used ensemble learning method for image analysis, classification, and detection tasks [14]. RF operates by constructing multiple decision trees and aggregating their predictions through a majority-voting mechanism. This ensemble approach significantly improves model accuracy and reduces the risk of overfitting, making RF a robust choice for medical image classification. Due to its high accuracy and resilience to noise, RF has gained significant attention in the research community, particularly in medical imaging applications [15,16]. The effectiveness of RF classifiers

improves as the number of trees in the forest increases, leading to better generalization and stability. In this study, RF is used to classify brain tumors based on extracted deep features, benefiting from its capability to handle high-dimensional data efficiently.

The k-Nearest Neighbors (KNN) algorithm, developed by Thomas Cover, is a simple yet effective non-parametric technique used for both regression and classification tasks across various domains [15]. It has been applied in numerous statistical and machine learning applications since the 1970s. Unlike other classification algorithms, KNN does not learn an explicit model from the training data. Instead, it stores all training instances in memory and classifies new data points by identifying the k-nearest neighbors based on a chosen distance metric (e.g., Euclidean distance). The final classification decision is determined by majority voting among the nearest neighbors [17]. KNN is particularly useful for brain tumor classification due to its ease of implementation, ability to capture local patterns, and effectiveness with small datasets. However, its computational complexity increases with larger datasets, as classification requires calculating distances for all stored instances.

Logistic Regression (LR) is a statistical classification method widely used for binary and multi-class problems in medical diagnosis and pattern recognition. Despite being one of the simplest ML classifiers, it remains a strong baseline model for classification tasks, including brain tumor detection. LR operates by estimating the probability of a given instance belonging to a particular class using a logistic function. The model optimizes decision boundaries by minimizing a cost function through optimization techniques such as gradient descent. In this study, LR is used to classify brain tumors based on extracted deep features, serving as a benchmark to compare the performance of more complex classifiers like SVM, RF, and deep learning models. Its advantages include interpretability, computational efficiency, and robustness to small datasets. However, LR assumes linear separability, which may limit its performance compared to non-linear models.

Decision Trees (DT) are widely used for classification tasks due to their interpretability, low computational cost, and ability to handle non-linear relationships in data. A decision tree operates by recursively splitting the data into subsets based on the most significant feature at each node, ultimately leading to classification decisions. DT classifiers have been extensively applied in medical imaging due to their ability to model complex decision boundaries while maintaining transparency in the decision-making process. However, standalone DT models are prone to overfitting, which can impact generalization performance. This limitation is mitigated when DT is used in ensemble methods such as Random Forest (RF) or Gradient Boosting. In this study, DT is used to classify brain tumors based on deep feature representations. Its inclusion allows us to evaluate the performance of simple tree-based models compared to more advanced classifiers like SVM and deep learning-based approaches.

3.3.1 Training Strategy: The Xception model is trained using the cross-entropy loss function and the Adam optimizer, with a small learning rate of $1e-5$ to ensure smooth fine-tuning. The model is trained with a batch size of 16 for a total of 100 epochs. To prevent overfitting, early stopping is applied, halting training if the validation loss shows no improvement for 10 consecutive epochs, while also restoring the best-performing weights. Additionally, a learning rate scheduler is employed, reducing the learning rate by a factor of 0.3 to enhance model performance. Furthermore, machine learning classifiers are trained using the extracted feature vectors for improved classification accuracy.

3.3.2 Fine-Tuning: Fine-tuning is the process of taking a pretrained model (transfer learning) and unfreezing some of its deeper layers to continue training on a new, task-specific dataset. This enables the model to learn domain-specific features, while still leveraging the general visual features it learned from a large dataset like ImageNet.

- Xception is trained on millions of images from the ImageNet dataset.
- Its early layers capture generic features (edges, colors, textures).
- Its deeper layers capture task-specific features (object parts, shapes).
- Brain MRI images are very different from natural images, so:
 - We retain early layers (good for general patterns),
 - But fine-tune deeper layers to adapt to tumor-specific textures, shapes, and intensity variations.

3.4 Model Evaluation

The trained model's performance is thoroughly evaluated using separate validation and test datasets. Metrics such as accuracy, precision, recall, and other performance indicators are employed to assess the model's predictive capabilities during the evaluation process. The proposed model was trained using a set of optimized hyperparameters to enhance performance. During the training phase, the model was fine-tuned using categorical cross-entropy loss and the Adam optimizer, which helped adjust parameters dynamically for efficient learning. The training process

was conducted over 100 epochs with a batch size of 16 and a learning rate of $1e-5$. For classification, the softmax **function** was employed in the fine-tuned Xception model to predict tumor categories. The implementation of deep learning models was carried out using the **Keras** and **TensorFlow** libraries, while machine learning classifiers were trained using **Scikit-learn**.

3.4.1 Performance Measurements

In this study, model performance was evaluated using key metrics, including accuracy, precision, recall, and F1-score. These evaluation metrics were derived from the fundamental classification outcomes: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The mathematical formulations for each of these metrics are provided in Equations (1)–(5).

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{True Positive} + \text{False Negative} + \text{False Positive} + \text{True Negative}} \quad (1)$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (2)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (3)$$

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}} \quad (4)$$

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

IV. RESULTS AND OBSERVATION

We evaluate the deep learning model using accuracy, precision, recall, and F1-score. Comparisons with other CNN architectures highlight the efficiency of Xception in feature extraction and classification. Additionally, ML classifiers trained on Xception-extracted features are evaluated. The results highlight that while deep learning performs exceptionally well, ML classifiers can also achieve competitive accuracy when provided with robust feature representations. Testing on a different dataset further validates the model's generalization ability.

As summarized in Table I, the model demonstrated excellent performance across both the datasets. It achieved an accuracy of **98.55%** on the training set, **98.00%** on the validation set, and **94.67%** on the test set. These results indicate that the model effectively distinguishes between tumor and non-tumor regions in brain scan images with high precision.

| Metric | Training Set | Validation Set | Testing Set |
|----------|--------------|----------------|-------------|
| Loss | 0.1103 | 0.1636 | 0.1968 |
| Accuracy | 98.55% | 98.00% | 94.67% |

TABLE I. EVALUATION METRICS

Furthermore, the model exhibited low values for training, and validation loss, suggesting effective learning and minimal prediction errors as shown in Figure 4. This indicates that the model successfully adjusted its parameters during training to reduce estimation errors and improve overall performance.

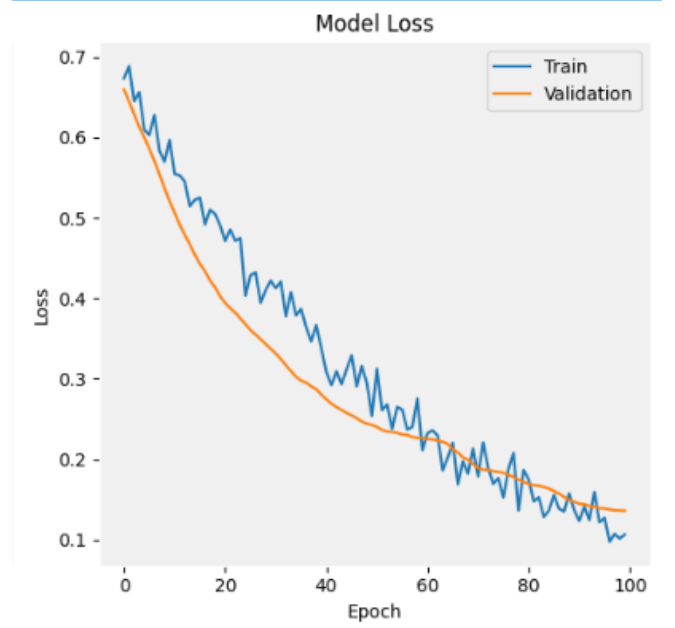


Fig. 4. Training and Validation Loss

As illustrated in Figure 5, the accuracy is plotted on the y-axis, ranging from 0.60 to 1, while the x-axis represents the number of epochs. The training accuracy, shown as a blue line, steadily increases with the number of epochs, indicating that the model effectively learns to classify the training data. Similarly, the validation accuracy, depicted by the orange line, also rises as the number of epochs increase, suggesting good generalization to unseen data. The highest validation accuracy is observed at Epoch 84.

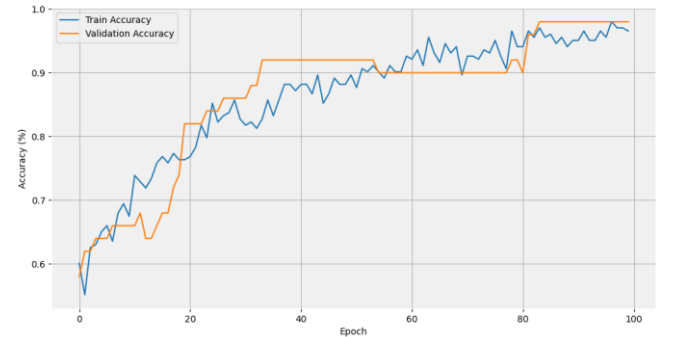


Fig. 5. Training and Validation Accuracy

The **Receiver Operating Characteristic (ROC)** curve is an essential performance evaluation tool that measures the model's ability to distinguish between tumor and non-tumor cases. Figure 6; presents the ROC curve for the proposed brain tumor classification model, illustrating the trade-off between the True Positive Rate (TPR) and False Positive Rate (FPR) at various classification thresholds.

In the ROC curve as shown in Figure 6, the blue and orange lines represent Class 0 (No Tumor) and Class 1 (Tumor), respectively. The Area Under the Curve (AUC) value for both classes is 0.95, indicating that the model achieves high classification performance.

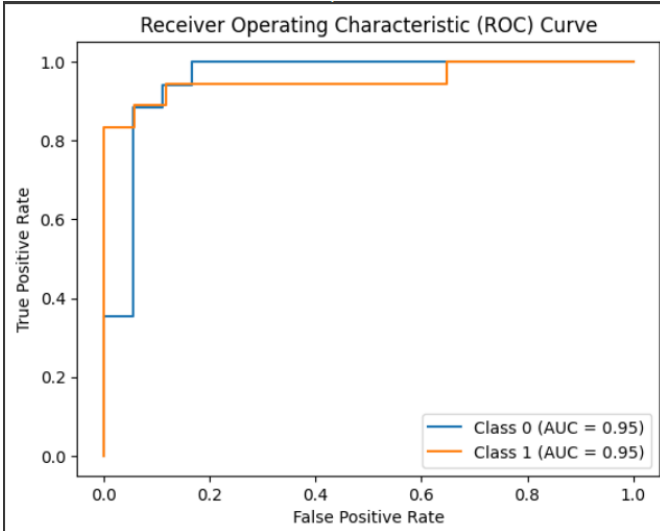


Fig. 6. ROC curve for each class

The AUC = 0.95 means that, on average, the model correctly ranks a randomly chosen tumor case higher than a non-tumor case 95% of the time.

- The curve approaches the top-left corner, which signifies strong sensitivity (recall) and low false positive rates.
- A high AUC score of 0.95 suggests that the model is highly effective in differentiating between brain tumor and non-tumor cases.
- The True Positive Rate (TPR) is significantly high, ensuring that most tumor cases are correctly identified, which is crucial for early diagnosis and treatment planning.
- The False Positive Rate (FPR) remains low, minimizing the number of healthy patients misclassified as having tumors, which helps reduce unnecessary medical procedures and patient anxiety.

4.1 Classification Metrics

The table II presents the precision, recall, and F1-score for the "No Tumor" and "Tumor" classes in the training and validation dataset. The model achieved the best validation accuracy of 98.00%, demonstrating its ability to differentiate between tumor and non-tumor cases with high reliability.

- Precision refers to the proportion of correctly identified cases within each predicted class. The model attained a precision of 94% for No Tumor cases and 89% for Tumor cases, indicating strong classification performance.
- Recall (Sensitivity) measures how effectively the model captures all actual positive cases. The recall values for No Tumor (88%) and Tumor (94%) suggest a balanced sensitivity to both classes.
- F1-Score represents the harmonic mean of precision and recall. The F1-scores of 0.94 (No Tumor) and 0.94 (Tumor) highlight the model's robustness in

maintaining a trade-off between false positives and false negatives.

- The high recall for tumor cases (94%) ensures that most tumor-positive patients are correctly identified, reducing the likelihood of missing critical cases.
- The balanced precision and recall scores indicate that the model is neither overly conservative (high false negatives) nor too aggressive (high false positives), making it clinically useful for assisting radiologists in brain tumor detection.

| Class | Precision | Recall | F1-score |
|----------|-----------|--------|----------|
| No Tumor | 0.94 | 0.88 | 0.94 |
| Tumor | 0.89 | 0.94 | 0.94 |

TABLE II. TRAINING AND VALIDATION DATASET METRICS

4.2 Classification Metrics on an Independent Test Dataset

To assess the generalization capability of our model, we conducted testing on a completely independent dataset consisting of 3,000 images (1,500 No Tumor, 1,500 Tumor). The performance metrics, including Precision, Recall, and F1-score, are reported in Table III.

- Precision: The model achieved a precision of 94.37% for No Tumor and 95.03% for Tumor, indicating that the predictions are highly reliable with minimal false positives.
- Recall: The recall scores (95.07% for No Tumor and 94.33% for Tumor) confirm that the model effectively identifies positive cases, minimizing false negatives.
- F1-Score: The F1-score, which balances precision and recall, is 0.9472 for No Tumor and 0.9468 for Tumor, demonstrating strong classification consistency across both classes.
- Overall Accuracy: The model achieved an accuracy of 94.7%, highlighting its robustness in classifying brain tumor images even on unseen data.

The high classification accuracy (94.7%) and balanced precision-recall trade-off indicate that our CNN-based brain tumor detection model performs consistently well across different datasets. These results validate the generalization ability of the model, reinforcing its potential as an assistive diagnostic tool for radiologists.

| Class | Precision | Recall | F1-score |
|----------|-----------|--------|----------|
| No Tumor | 0.9437 | 0.9507 | 0.9472 |
| Tumor | 0.9503 | 0.9433 | 0.9468 |

TABLE III. EVALUATION METRICS ON TEST DATASET

4.3 Classifier Performance Comparison

The below table IV; presents the performance of various classifiers integrated with the CNN model. Each classifier was evaluated using both training and testing accuracy.

| Classifier | Training Accuracy | Testing Accuracy |
|------------|-------------------|------------------|
| CNN-SVM | 100.00% | 99.91% |
| CNN-DT | 100.00% | 94.43% |
| CNN-KNN | 95.12% | 94.43% |
| CNN-RF | 100.00% | 96.29% |
| CNN-LR | 100.00% | 96.29% |

TABLE IV. COMPARISON TABLE OF ML CLASSIFIERS

- Support Vector Machines (SVM) combined with CNN features provided the highest classification performance, suggesting that deep learning features align well with SVM's high-dimensional decision boundary formulation.
- A very low error margin implies that only a very small fraction of test samples were misclassified.
- CNN-RF and CNN-LR achieved 96.29% test accuracy, indicating that ensemble learning and logistic regression can also leverage CNN features effectively.
- CNN-KNN and CNN-DT scored 94.43%, demonstrating that these models can still classify brain tumor cases effectively but are slightly less generalizable than SVM.

The proposed CNN-SVM classifier achieved an exceptional testing accuracy of 99.91%, surpassing all traditional machine learning classifiers. This demonstrates that leveraging deep learning features in combination with an SVM classifier provides a highly effective approach for brain tumor detection. Furthermore, our model outperforms the most recent study by Seoyoung Yoon[3], titled '*Brain Tumor Classification Using a Hybrid Ensemble of Xception and Parallel Deep CNN Models*', published on January 2, 2025, further highlighting the superiority of our method in brain tumor classification

4.2 Predictions

Figure 7; illustrates a set of MRI scan predictions made by the proposed **brain tumor detection model**. The model

categorizes MRI images into two classes: "**Tumor**" and "**No Tumor**".

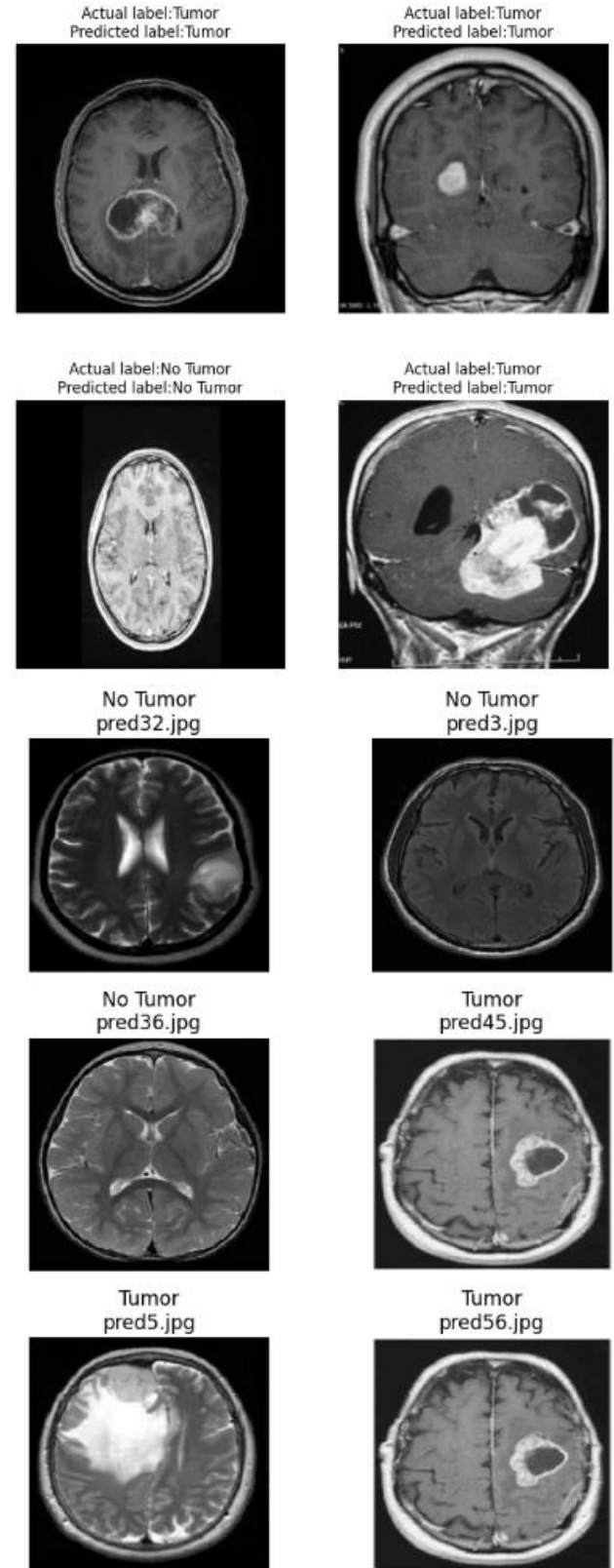


Fig. 7. Predicted MRI images.

The results suggest that the **fine-tuned Xception model**, combined with machine learning classifiers, effectively

generalizes unseen data, providing accurate predictions for brain tumor classification.

V. CONCLUSION

Recent advancements in deep learning have enabled the use of MRI data for diagnosing brain tumors. This paper presents a sophisticated convolutional neural network (CNN) architecture, specifically the Xception model pre-trained on ImageNet, to accurately identify tumor locations in medical imaging data. The results demonstrated impressive accuracy, achieving 98% in distinguishing tumor and non-tumor regions in brain MRI images. Overall, the study highlights the promising role of deep learning techniques in transforming brain tumor diagnosis. It also underscores the need for interdisciplinary collaboration between computer scientists, medical professionals, and regulators to fully harness the potential of machine learning in improving patient care. This research underscores the potential of integrating deep learning and machine learning for medical image classification, paving the way for improved diagnostic tools in healthcare.

VI. LIMITATIONS AND FUTURE SCOPE

The proposed model effectively balances the cost and performance by using deep learning and machine learning classifiers; however, certain limitations remain. Training on a small dataset improves efficiency but may limit the model's ability to recognize rare tumor variations, potentially affecting generalization. While testing on a **larger dataset** enhances robustness, variations in MRI scans from different medical sources could impact real-world performance. The future research can address these limitations by expanding the dataset for better diversity, developing lightweight models for real-time deployment, and integrating explainable AI techniques to improve interpretability for clinical applications.

ACKNOWLEDGMENT

I would like to express my deepest gratitude to my supervisor, Dr. Sayed Khaled, Assistant Professor in the Department of Electrical & Computer Engineering and Computer Science, at the University of New Haven, for his invaluable guidance, insightful suggestions, and continuous support throughout this research. His mentorship has been instrumental in shaping my understanding of the subject, and I have gained both professional and personal growth under his supervision. His encouragement and constructive feedback have greatly contributed to the successful completion of this study. Additionally, I would like to extend my appreciation to my family, friends, and colleagues, whose unwavering support and motivation have played a crucial role in this journey.

REFERENCES

- [1] Disci R, Gurcan F, Soylu A. Advanced Brain Tumor Classification in MR Images Using Transfer Learning and Pre-Trained Deep CNN Models. *Cancers (Basel)*. 2025 Jan 2;17(1):121. doi: 10.3390/cancers17010121. PMID: 39796749; PMCID: PMC11719945.
- [2] J. J. Philip, R. Rani and S. Kumar, "Brain Tumor Detection Using Xception Model," *2024 International Conference on Advances in Computing Research on Science Engineering and Technology (ACROSET)*, Indore, India, 2024, pp. 1-6, doi: 10.1109/ACROSET62108.2024.10743964.
- [3] Seoyoung Yoon, Brain Tumor Classification Using a Hybrid Ensemble of Xception and Parallel Deep CNN Models, *Informatics in Medicine Unlocked*, 2025, 101629, ISSN 2352-9148, <https://doi.org/10.1016/j.imu.2025.101629>.
- [4] Sathya R, Mahesh TR, Bhatia Khan S, Malibari AA, Asiri F, Rehman AU, Malwi WA. Employing Xception convolutional neural network through high-precision MRI analysis for brain tumor diagnosis. *Front Med (Lausanne)*. 2024 Nov 8;11:1487713. doi: 10.3389/fmed.2024.1487713. PMID: 39606635; PMCID: PMC11601128.
- [5] Babu Vimala, B., Srinivasan, S., Mathivanan, S.K. *et al.* Detection and classification of brain tumor using hybrid deep learning models. *Sci Rep* **13**, 23029 (2023). <https://doi.org/10.1038/s41598-023-50505-6>
- [6] M. Siar and M. Teshnehlab, "Brain Tumor Detection Using Deep Neural Network and Machine Learning Algorithm," *2019 9th International Conference on Computer and Knowledge Engineering (ICCCKE)*, Mashhad, Iran, 2019, pp. 363-368, doi: 10.1109/ICCCKE48569.2019.8964846.
- [7] S. Solanki, U. P. Singh, S. S. Chouhan and S. Jain, "Brain Tumor Detection and Classification Using Intelligence Techniques: An Overview," in *IEEE Access*, vol. 11, pp. 12870-12886, 2023, doi: 10.1109/ACCESS.2023.3242666.
- [8] Shenbagarajan Anantharajan, Shenbagalakshmi Gunasekaran, Thavasi Subramanian, Venkatesh R, MRI brain tumor detection using deep learning and machine learning approaches, *Measurement: Sensors*, Volume 31, 2024, 101026, ISSN 2665-9174, <https://doi.org/10.1016/j.measen.2024.101026>.
- [9] Rasheed, Z.; Ma, Y.-K.; Ullah, I.; Al-Khasawneh, M.; Almutairi, S.S.; Abohashrh, M. Integrating Convolutional Neural Networks with Attention Mechanisms for Magnetic Resonance Imaging-Based Classification of Brain Tumors. *Bioengineering* **2024**, *11*, 701. <https://doi.org/10.3390/bioengineering11070701>
- [10] Zahoor MM, Qureshi SA, Bibi S, Khan SH, Khan A, Ghafoor U, et al. A new deep hybrid boosted and ensemble learning-based brain tumor analysis using MRI. *Sensors (Basel)*. 2022 Apr;22(7):2726.
- [11] C. Cortes and V. Vapnik. (1995). "Support-vector networks," *Machine Learning*, vol. 20, pp. 273–297. [[Google Scholar](#)]
- [12] Y. Ma and G. Guo. (2014). "Multi-class support vector machine—Chapter no. 2," in *Support Vector Machines Applications*, vol. 2. New York, NY, USA: Springer, pp. 23–48. [[Google Scholar](#)]
- [13] R. Gandhi. (2018). *Support Vector Machine-Introduction to Machine Learning Algorithms*, Canada: Towards Data Science, . <https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>.
- [14] L. Breiman. (2001). "Random forests," *Machine Learning*, vol. 45, pp. 5–32. [[Google Scholar](#)]
- [15] K. Fawagreh, M. M. Gaber and E. Elyan. (2014). "Random forests: From early developments to recent advancements," *Systems Science & Control Engineering: An Open Access Journal*, vol. 2, pp. 602–609. [[Google Scholar](#)]
- [16] P. Thanh Noi and M. Kappas. (2018). "Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery," *Sensors*, vol. 18, pp. 18. [[Google Scholar](#)]
- [17] Z. Omary. (2008). "Development of a workflow for the comparison of classification techniques," M.S. thesis, School of Computing, Technological Univ, Dublin, Ireland. [[Google Scholar](#)]
- [18] F. Chollet. (2017). "Xception: Deep learning with depthwise separable convolutions," in 2017 IEEE Conf. on Computer Vision and Pattern Recognition, Honolulu, HI, pp. 1800–1807. [[Google Scholar](#)]

