Machine Learning for Human Intelligence
2024, Vol. 2, Issue 2

Research Article

Detection of Brain Tumors in MRI Scans Utilizing Deep Learning: A Comparative Analysis of Advanced CNN Models

Muhammad Dawood Majid[1,⋆] and Joanna Marie Diaz[2]

[1]Department of Robotics and Artificial Intelligence, Shaheed Zulfikar Ali Bhutto Institute of Science and Technology, Islamabad, Pakistan.; Email: 21108133@szabist-isb.pk
[2]UNICAF, Larnaca, Cyprus.; Email: joannamarie210@gmail.com
⋆Corresponding author: Muhammad Dawood Majid (21108133@szabist-isb.pk)

Article History

Academic Editor:
Dr. Muhammad Sajid
Submitted: January 21, 2024
Revised: March 19, 2024
Accepted: September 1, 2024

Keywords:
Convolutional Neural Networks; Deep Learning; Magnetic Resonance Imaging; Transfer Learning; Automated Diagnosis; Clinical Diagnostics; Brain Tumor Classification

Abstract

Classification of brain tumors is essential in medical diagnostics, as timely and precise identification can significantly enhance patient outcomes. This study examines the efficacy of pre-trained deep learning models in classifying brain MRI images into four categories: pituitary, meningioma, glioma, and no tumor, with the aim of automating and enhancing the diagnostic process. We utilized a publically accessible MRI dataset comprising 7,023 pictures of brain tumors. Alongside comprehensive image preprocessing and data augmentation, transfer learning was employed to enhance the performance of four sophisticated convolutional neural network architectures: DenseNet121, ResNet50, Xception, and MobileNet. The models were successfully refined by transfer learning, reducing computational requirements while enhancing classification precision. DenseNet121 surpassed Xception, MobileNet, and ResNet50 in the evaluated models, attaining the greatest accuracy of 98.47% and an F1 score of 98.47%. The models' appropriateness for clinical application was validated by their robust generalization and consistent performance across critical assessment parameters. To enhance memory for particular tumor classifications and render deep learning predictions more interpretable in medical settings, further developments are required.

1 Introduction

Brain tumors are among the most severe and possibly lethal conditions, consistently resulting in significant neurological deficits and a diminished quality of life. Primary tumors arise in the brain, whereas secondary tumors result from cancers that have metastasized from other regions of the body. These tumors are categorized into two primary classifications. Gliomas, meningiomas, and pituitary tumors are the predominant categories of primary brain neoplasms, each presenting distinct challenges for diagnosis and management [1]. Meningiomas and pituitary tumors, although prevalent, can pose significant risks if undetected. Recent improvements in digital image processing and medical imaging have facilitated the extensive adoption of computer-aided diagnosis (CAD). The MRI approach is preferred in diagnostic systems of this nature as it does not expose patients to ionizing radiation and can accurately detect blood flow in veins. The utilization of extensive medical imaging datasets, such as

Brain MRI scans, for the identification of brain tumors may be enhanced by the application of machine learning and deep learning techniques [2].

Developing a machine learning and deep learning model is a multifaceted process that necessitates training with a substantial volume of medical imaging data [3, 4]. This is essential for obtaining accurate predictions or insights from the model, which is crucial for making proper therapeutic decisions. This study examines the identification of brain tumors via deep learning and machine learning approaches. The duration required to evaluate a tumor is contingent upon the radiologist's proficiency and expertise. Nonetheless, the procedure for tumor identification is both inaccurate and costly [5, 6]. Misdiagnosing a brain tumor can substantially diminish a patient's survival chances, leading to severe complications. The MRI technology is increasingly favored as a remedy for the constraints of human diagnosis. Recent advancements in machine learning, especially in deep learning, have facilitated the discovery and classification of patterns in medical imaging [7]. Achievements in this domain encompass the capability to retrieve and extract knowledge from data rather than relying on experts or scholarly literature. Machine learning is swiftly emerging as a valuable instrument for enhancing performance across many medical applications, including illness prognosis and diagnosis, molecular and cellular structure identification, tissue segmentation, and picture categorization [6–8].

In the analysis of extensive datasets or nuanced imaging characteristics, conventional diagnostic methods predominantly depend on expert interpretation, which may be subjective and susceptible to inaccuracies [8]. This method might be automated using deep learning systems, facilitating a faster, more reliable, and potentially more accurate diagnosis. These technologies are becoming recognized as valuable tools to aid clinicians in making timely and informed treatment decisions [9] as shown in Figure 1. The primary objective of this work is to enhance the accuracy and efficiency of brain tumor detection

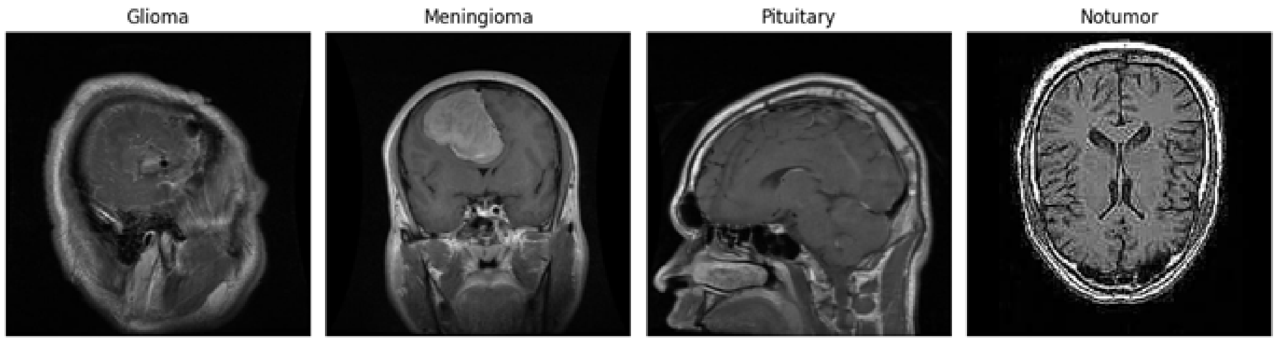

Figure 1: Sample Images.

from MRI scans through the application of deep learning and transfer learning techniques. This study aims to develop and evaluate automated classification models capable of properly distinguishing among four categories—glioma, meningioma, pituitary tumor, and no tumor—utilizing a publically available Brain Tumor MRI dataset [10]. A comprehensive comparison is conducted on the classification performance of the most advanced pre-trained CNN architectures, including DenseNet121, ResNet50, Xception, and MobileNet [11].

This paper chiefly contributes the following:

- Utilization of transfer learning with prominent pre-trained CNN models to classify brain MRI images, targeting swift, precise, and uniform outcomes.
- Methodical evaluation of DenseNet121, ResNet50, Xception, and MobileNet architectures for multi-class brain tumor classification.
- Thorough assessment of model efficacy on a four-class MRI dataset, utilizing metrics such as accuracy, precision, recall, and F1-score.
- Advocacy for automated deep learning-based classification systems to facilitate clinical decision-making, diminish diagnostic variability, and improve patient care.

2 Related Literature

The domain of automated brain tumor detection and classification utilizing MRI images has swiftly advanced, with deep learning and hybrid machine learning methodologies leading contemporary research efforts. Numerous studies have concentrated on enhancing accuracy, robustness, computing efficiency, and clinical application through the utilization of sophisticated neural network topologies, transfer learning, and novel preprocessing techniques.

A notable trend in the literature is the implementation of privacy-preserving and decentralized learning frameworks. The FL-SiCNN model incorporates a Siamese Convolutional Neural Network into a peer-to-peer federated learning framework, eliminating the necessity for a central server and attaining 98.78% accuracy in multi-class brain tumor classification, while safeguarding data privacy and demonstrating resilience against data poisoning attacks [12]. Hybrid methodologies that integrate deep feature extraction (utilizing ResNet101 and DenseNet121), dimensionality reduction (PCA), and Random Forest classification have exhibited remarkable efficacy and versatility across several noisy MRI datasets, with accuracy levels of up to 99.7% [13].

Advanced deep learning architectures have been investigated to capture both spatial and temporal aspects in MRI data. A four-stage pipeline integrating adaptive filtering, enhanced K-means clustering, GLCM feature extraction, and Recurrent Convolutional Neural Networks (RCNN) attained 95.17% accuracy, 98.42% sensitivity, and 89.28% specificity, surpassing conventional models such as BP and U-Net [14]. Parallel Deep Convolutional Neural Networks (PDCNN) have been introduced to enhance spatial feature extraction by processing MRI data via multiple pipelines, leading to superior feature learning and less overfitting, with an accuracy of 96.29% reported [15] as shown in Figure 2. Transfer

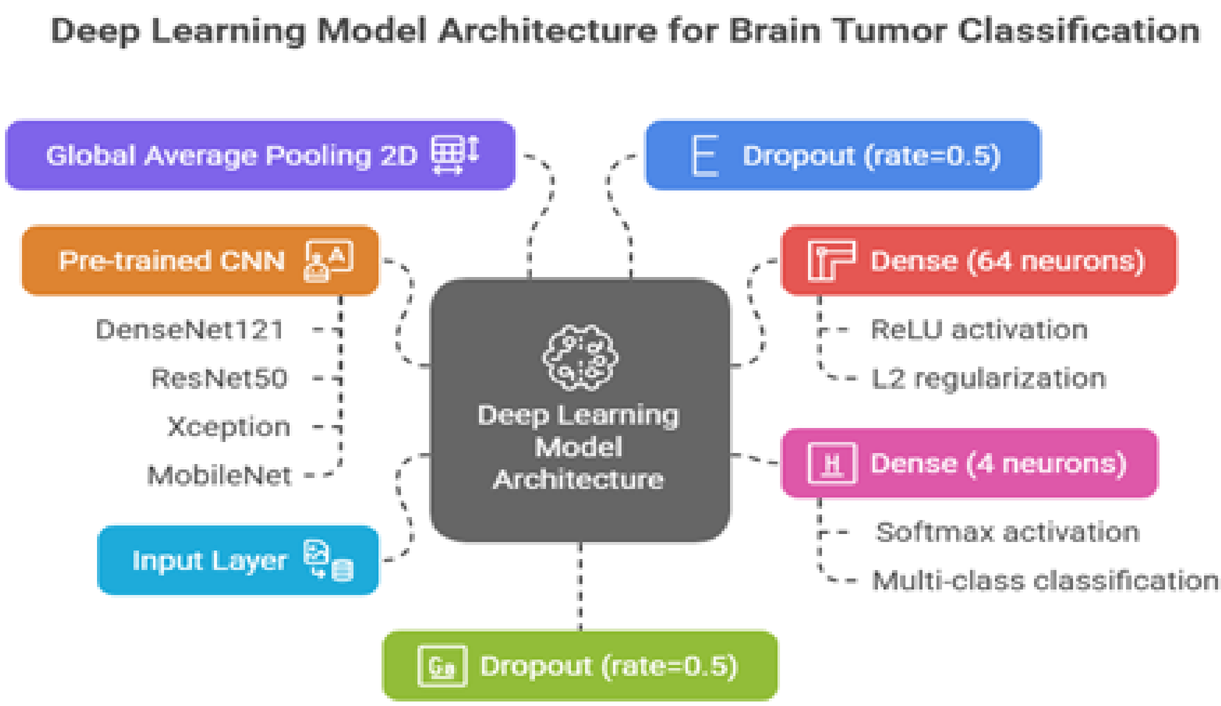

Figure 2: Architecture Representation.

learning utilizing pre-trained models is essential for achieving high-performance systems. Fine-tuned ResNet50 and EfficientNet models have consistently produced robust outcomes, with ResNet50-based systems attaining 98.5% accuracy and surpassing earlier architectures like VGG16 and InceptionV3 [16]. EfficientNet, when integrated with sophisticated preprocessing techniques (contrast enhancement, saliency maps) and Extreme Learning Machines (ELM), has demonstrated strong generalization and accuracy across many public datasets [17]. Lightweight models such as MobileNet and RetinaNet provide real-time tumor identification on medical edge devices, achieving a balance between computing economy and high detection accuracy—an indispensable characteristic for implementation in resource-constrained clinical environments [18].

Extensive evaluations in the domain highlight that deep learning models, especially 3D CNNs and attention-based architectures, surpass conventional machine learning methods in segmentation and classification tasks [19]. The incorporation of sophisticated preprocessing techniques (e.g., equalization, homomorphic filtering), attention mechanisms (CBAM), and hybrid architectures (EfficientNetB2 with equalization) has elevated detection and classification accuracy to 99% in certain studies [20].

The application of multi-modal MRI data and hybrid systems integrating deep learning with machine learning classifiers, such as SVM, has been investigated to enhance segmentation and classification precision. A system employing a bespoke 17-layer CNN for segmentation, alongside MobileNetV2 for feature extraction and a multi-class SVM, attained a segmentation accuracy of 97.47% and a classification accuracy of 98.92% [21]. From a scientific perspective, tumor diagnosis via medical imaging is flawed and significantly reliant on the radiologist's expertise. Moreover, the application of these technologies facilitates precise and errorless tumor diagnosis, enabling differentiation from other analogous disorders. The aim of artificial intelligence (AI) is to develop machines that function and operate similarly to humans [22]. Brain tumors present a considerable diagnostic challenge compared to cancers originating in other organs of the human body. Due to the brain's

3 Methodology

This section offers an extensive analysis of the identification of MRI brain tumors through the application of deep learning and machine learning algorithms. The advancement of the proposed methodology. Initially, MRI brain tumor data were acquired and preprocessed using ACEA and the median filter to eliminate noise. To segment the MRI brain pictures, a fuzzy c-means approach is employed, and a GLCM matrix is utilized to extract the image features. The EDN-SVM method is subsequently employed to categorize images of healthy and tumorous brain tissue. In our work, an input image measuring 32×32 pixels was processed by an initial convolutional layer with 16 filters, resulting in a $32 \times 32 \times 16$ feature map, utilizing a kernel size of 3×3 to identify the most general characteristics. The output of the convolutional layer was subsequently sent to a max-pooling layer feature map measuring $15 \times 15 \times 16$, effectively reducing the spatial data size for the following layer by fifty percent [23]. The max-pooling process identifies the maximum elements or pixels from the region of the feature map that the filter has encompassed. The output was subsequently input into an additional convolutional layer using 32 filters and a $13 \times 13 \times 32$ feature map, utilizing a 3×3 kernel size. Subsequently, the output was transmitted to the max-pooling layer feature map of $6 \times 6 \times 32$ to reduce the spatial data for the subsequent layer by fifty percent. A further convolutional layer was followed by an additional pooling layer. A categorical cross-entropy loss function was computed in conjunction with the Adam optimizer to determine the loss value. as shown in Figure 3.

3.1 Dataset Analysis and Partitioning

We employ a dataset available on the Kaggle open data platform to assess the efficacy of the proposed architectural design. This dataset has 255 T1-weighted MRI scans. It comprises 98 MRI slices derived from healthy brain tissue and 155 MRI slices obtained from tumorous brain tissue. Due to the distinct dimensions of each image, we had to modify them to conform to our image specifications. A segment of the dataset utilized for our analysis. The dimensions of the photos differ in width and height. Classifying healthy brain tissue and tumorous brain tissue in MRI brain scans of varying heights and widths may prove challenging. Consequently, prior to advancing to the preprocessing stage, we standardize the photos by ensuring their width and height are same. The testing set comprises 306 meningioma images, 300 glioma images, 300 pituitary images, and 405 non-tumor images. The training dataset comprises 1,595 non-tumor photos, 1,457 pituitary images, 1,339 meningioma images, and 1,321 glioma images. This rather consistent distribution across categories enables more precise and equitable comparisons of model performances. Moreover, the diversity of tumor types ensures that

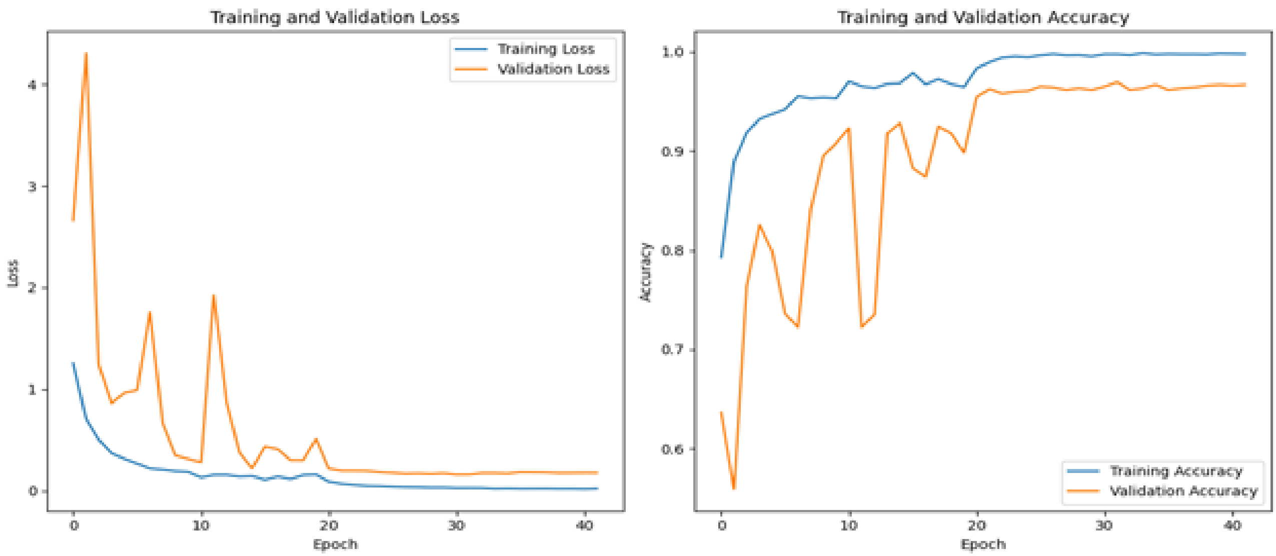

Figure 3: Loss and accuracy curves.

any classification model is assessed in a manner that reflects real clinical conditions, enhancing its application and significance.

3.2 Data Preparation

Prior to training, the MRI images underwent a systematic series of preprocessing techniques to ensure optimal input quality for the deep learning models. To standardize the input dimensions and meet the specifications of the selected convolutional neural network designs, all photos were initially reduced to 128×128 pixels. Subsequently, pixel values were normalized to a range of 0 to 1, enhancing stability and accelerating convergence during model training. The original color channels of the images preserved the comprehensive spatial and intensity information necessary for accurate tumor classification and effective feature extraction as shown in Figure 4. .

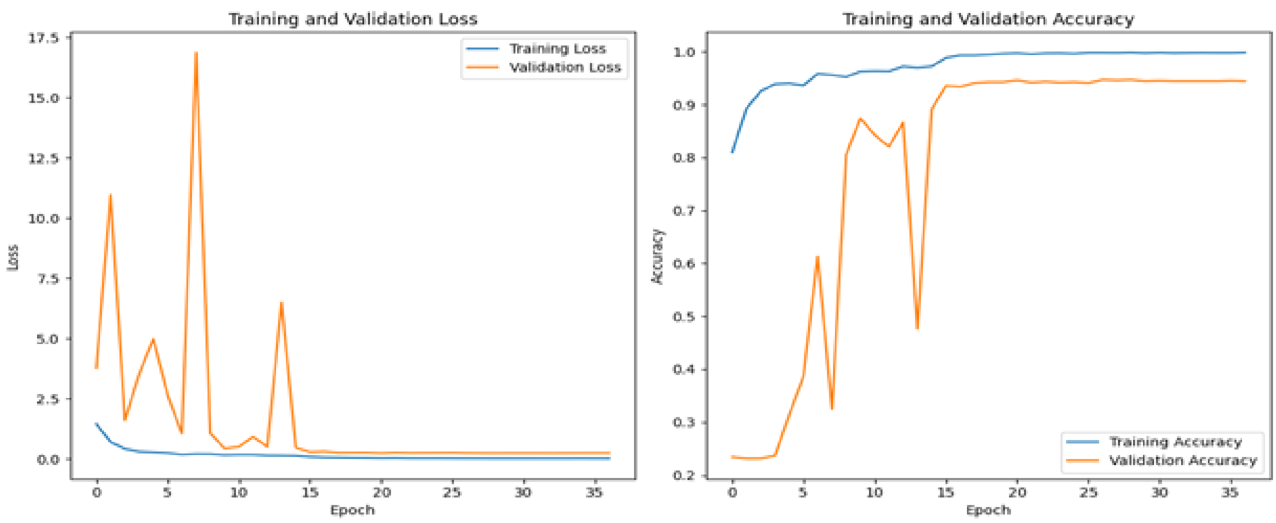

Figure 4: Training and validation loss.

3.3 Data Augmentation

A variety of data augmentation techniques were employed on the training dataset to further improve model generalization and mitigate the risk of overfitting. These additions encompassed shifting, zooming, flipping in both orientations, and random rotations. The model encounters a broader range of picture orientations and scales through the incorporation of these variances, effectively mirroring the diversity seen in actual MRI data. This strategy enhances the accuracy of brain tumor classification by augmenting the training set size and enabling the model to acquire more resilient and invariant features.

3.4 Utilization of Pre-Trained Convolutional Neural Network Models

This design is particularly adept for medical image analysis as it addresses the vanishing gradient problem and facilitates effective feature reuse [26, 27]. ResNet50 facilitates the training of profoundly deep networks through the utilization of residual connections, enabling gradients to traverse the network directly. The fifty-layer depth structure is engineered to acquire hierarchical characteristics, which are essential for differentiating tiny variations in MRI scans [28].

Xception utilizes depthwise separable convolutions, which disaggregate conventional convolution processes into more efficient components. This architecture minimizes processing expenses while preserving substantial representational capacity, enabling the model to discern complex spatial patterns in the images [29]. The computational efficiency of MobileNet renders it ideal for implementation on low-resource devices. It employs depthwise separable convolutions and a streamlined design to attain a compromise between speed and precision [30].

All of these models were initially trained on the extensive ImageNet dataset, which allowed them to acquire robust and broadly applicable visual properties [31]. This work utilized transfer learning by substituting the final classification layers of each pre-trained model with newly constructed dense layers specifically designed for the four-class brain tumor classification assignment. The lower layers, which held pre-trained weights, were either frozen or altered during training, enabling the models to adjust to the specific attributes of brain MRI data as shown in Figure 5.

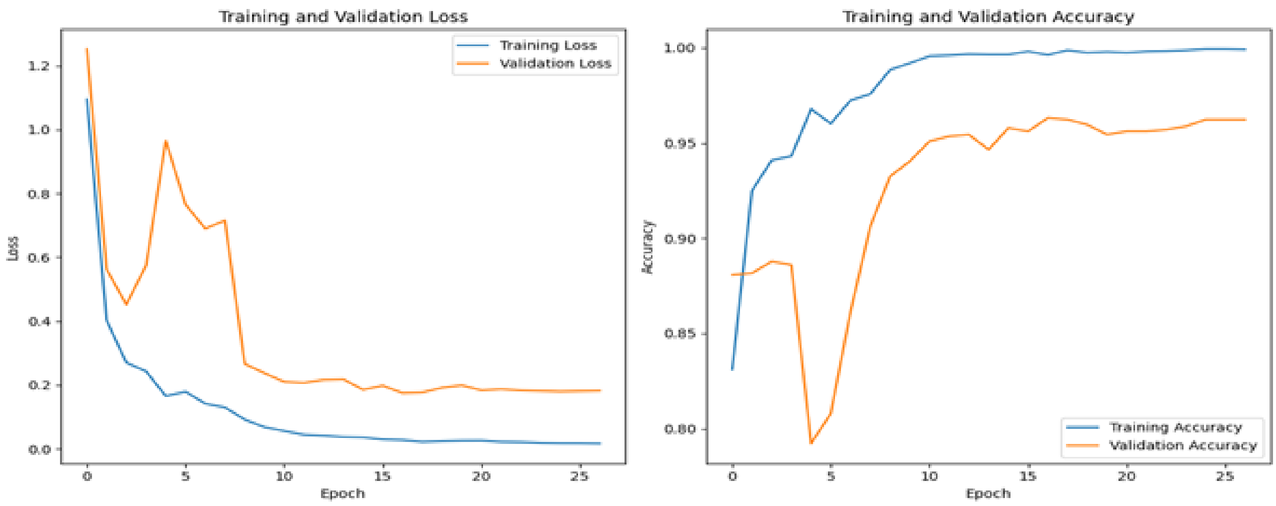

Figure 5: Xception model.

3.5 Model Architecture and Compilation

The model architecture designed for this study is optimized to accurately capture the intricate features in brain MRI images and to generalize successfully to novel, unseen data. A pre-trained convolutional neural network (CNN) underpins the architecture, with further layers incorporated for classification purposes.

The input layer is designed for RGB images at a resolution of 128×128 pixels, selected to optimize computational speed while maintaining adequate detail for precise classification. The model was trained using mini-batch gradient descent with a batch size of 32. To mitigate overfitting and enhance training efficiency, early halting and model checkpointing strategies were implemented. Early stopping observes validation performance and ceases training when no more improvement is detected, while model checkpointing preserves the best-performing model throughout training, guaranteeing optimal results are retained.

3.6 Metrics for Evaluating Models

Four prevalent metrics—accuracy, precision, recall, and F1-score—were employed to rigorously evaluate the performance of the proposed models for brain tumor classification. In medical image analysis, where minor misclassifications can lead to significant clinical consequences, these measures are particularly vital.

Accuracy: Quantifies the ratio of total right predictions (including both true positives and true negatives) to all forecasts made.

Precision: Denotes the ratio of accurately predicted positive instances to the total instances projected as positive.

Recall (Sensitivity): Reflects the model's ability to identify all actual positive instances.

F1-Score: A balanced metric of precision and recall, representing their harmonic mean.

To compute an overall performance score for multi-class classification, these metrics were calculated for each class and subsequently averaged (macro-averaged) as shown in Figure 6.

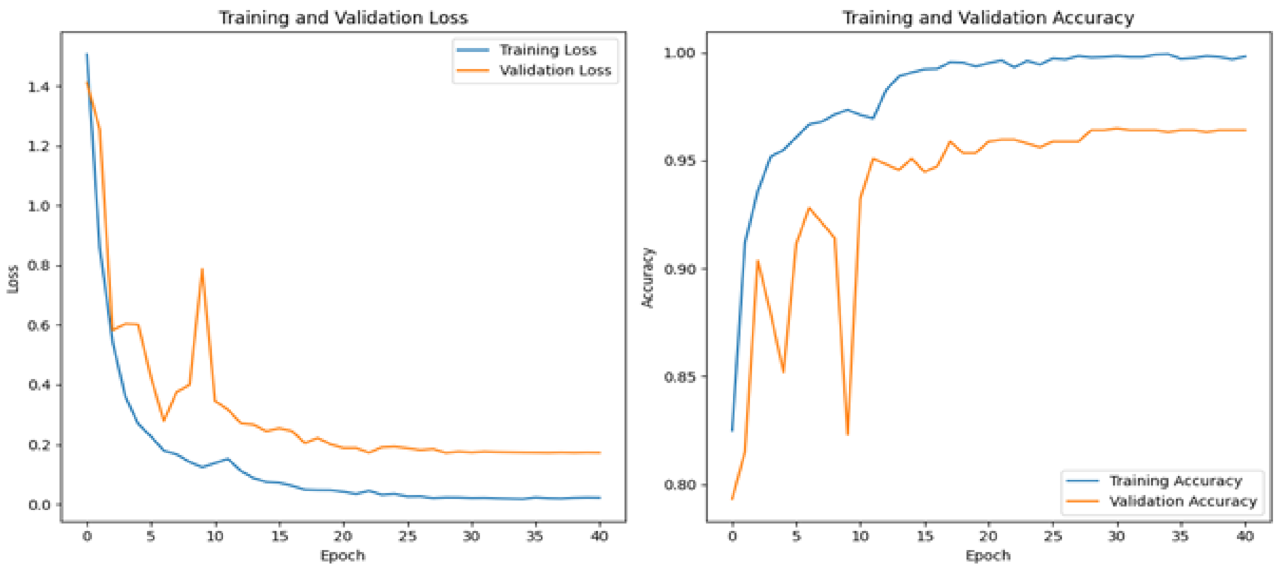

Figure 6: MobileNet model.

4 Results

This section presents and analyzes the outcomes of experiments conducted with four pre-trained deep learning models: DenseNet121, ResNet50, Xception, and MobileNet. These models exemplify various convolutional neural network (CNN) architectures through their unique layer designs and feature extraction capabilities.

Image segmentation is significantly enhanced by the application of fuzzy clustering. A proficient method applicable in fuzzy clustering is referred to as the fuzzy c-means algorithm. The FCM is a clustering method that allows a single pixel to simultaneously belong to many clusters. The FCM approach aims to partition a finite set of pixels into "C" fuzzy clusters by utilizing specified criteria in the decision-making process. The FCM approach aims to minimize the objective function presented below. The dataset was limited and solely comprised MR images; yet, deep neural networks necessitate a substantial dataset to yield favorable outcomes. The dataset comprised 3,264 MR images, with 80% allocated for training and the other 20% distributed equally for testing and validation, each at 10%. The initial data volume can be enhanced through augmentation, subsequently improving the training process. This also improves the model's learning capacity. The weighting value for the F-measure can be calculated using the precision and recall metrics. The F-measure is an effective metric for assessing classification quality, representing the weighted average of precision and recall. The value of this measure ranges from 0 to 1, where 0 represents the most unfavorable condition and 1 signifies the most favorable condition.

5 Conclusions and Future Research

The findings indicate that transfer learning utilizing these architectures can attain very accurate and balanced classification, with DenseNet121 achieving the best accuracy at 98.47%, followed closely by Xception at 98.17%, MobileNet at 97.86%, and ResNet50 at 97.03%. All models demonstrated significant generalization and resilient performance across essential measures, including precision, recall, and F1 score. These results underscore the considerable potential of deep learning methodologies to facilitate and improve clinical decision-making in neuro-oncology through rapid, dependable, and automated tumor detection.

Although the results are encouraging, numerous opportunities for additional research persist. Subsequent research should prioritize the augmentation of the dataset to incorporate a broader range of heterogeneous and multi-center MRI scans, hence enhancing model generalizability and mitigating class imbalance. Utilizing advanced data augmentation techniques alongside ensemble or hybrid modeling approaches may enhance classification accuracy, especially for difficult tumor types. Furthermore, using explainability techniques like Grad-CAM or saliency maps will enhance clinical trust by rendering model judgments more transparent [36, 37]. Ultimately, investigating novel designs such as Vision Transformers (ViT), Swin Transformer, and other advanced models may enhance performance and support practical clinical implementation [38]. These stages will be essential for transitioning deep learning-based brain tumor categorization from research to standard clinical practice.

References

[1] Cheng, J., et al. (2015). Improved efficacy of brain tumor classification through tumor region augmentation and partition. *PloS One*, 10(10), e0140381.

[2] Ostrom, Q. T., et al. (2022). CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States. *Neuro-Oncology*, 24(Suppl_5), v1–v95. https://doi.org/10.1093/neuonc/noac171

[3] Louis, D. N., et al. (2016). *World Health Organization Classification of Tumors of the Central Nervous System*. World Health Organization Classification of Tumors, Revised Fourth Edition.

[4] Kim, D. W., et al. (2020). A Review of Machine Learning Techniques for the Diagnosis and Prognosis of Brain Tumors Utilizing MRI. *Diagnostics*, 10(7), 509. https://doi.org/10.3390/diagnostics10070509

[5] Pereira, S., et al. (2016). Segmentation of Brain Tumors Utilizing Convolutional Neural Networks in MRI Scans. *IEEE Transactions on Medical Imaging*, 35(5), 1240–1251.

[6] Tajbakhsh, N., et al. (2016). Convolutional Neural Networks for Medical Image Analysis: Complete Training or Fine-Tuning? *IEEE Transactions on Medical Imaging*, 35(5), 1299–1312. https://doi.org/10.1109/TMI.2016.2535302

[7] Shin, H.-C., et al. (2016). Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Attributes, and Transfer Learning. *IEEE Transactions on Medical Imaging*, 35(5), 1285–1298. https://doi.org/10.1109/TMI.2016.2528162

[8] Rehman, A., et al. (2020). Classification of Brain Tumor Varieties from MRI Images Utilizing Deep Learning Techniques. *Journal of Healthcare Engineering*, 2020, 1–10. https://doi.org/10.1155/2020/8884616

[9] Litjens, G., et al. (2017). A Study on Deep Learning in Medical Image Analysis. *Medical Image Analysis*, 42, 60–88. https://doi.org/10.1016/j.media.2017.07.005

[10] Nickparvar, M. (2022). *Magnetic Resonance Imaging Dataset for Brain Tumors*. Kaggle. https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset

[11] Tan, M., & Le, Q. V. (2019). EfficientNet: Reevaluating Model Scaling for Convolutional Neural Networks. In *Proceedings of the 36th International Conference on Machine Learning* (pp. 6105-6114). PMLR.

[12] Amin, J., Sharif, M., Anjum, M. A., & Raza, M. (2021). Detection and Classification of Brain Tumors Utilizing a Hybrid Machine Learning Model. *Biomedical Signal Processing and Control*, 68, 102595. https://doi.org/10.1016/j.bspc.2021.102595

[13] Deepa, T., & Srinivas, K. (2021). A Hybrid RCNN Model for Classifying Brain Tumors Utilizing MRI Images. *Journal of Ambient Intelligence and Humanized Computing*, 12, 11211–11223. https://doi.org/10.1007/s12652-021-02934-0

[14] Afshar, P., Mohammadi, A., & Plataniotis, K. N. (2020). Classification of Brain Tumor Types Using Capsule Networks. *IEEE Transactions on Biomedical Engineering*, 68(1), 179–190. https://doi.org/10.1109/TBME.2020.2990045

[15] Abd-Ellah, M. K., Awad, A. I., Khalaf, A. A. M., & Hamed, H. F. A. (2019). Feature Extraction for Brain Tumor Classification Utilizing Deep Learning. *Egyptian Informatics Journal*, 20(2), 139–150. https://doi.org/10.1016/j.eij.2018.12.001

[16] Choudhury, S., & Singh, K. K. (2022). EfficientNet and ELM for Brain Tumor Classification Utilizing Enhanced MRI. *Computerized Medical Imaging and Graphics*, 94, 102009. https://doi.org/10.1016/j.compmedimag.2021.102009

[17] Raza, M., Sharif, M., & Amin, J. (2021). Detection of Brain Tumors Utilizing Transfer Learning with RetinaNet. *Computers in Biology and Medicine*, 133, 104362. https://doi.org/10.1016/j.compbiomed.2021.104362

[18] Li, Y., et al. (2022). Deep Learning in Medical Imaging: Three-Dimensional Convolutional Neural Networks and Attention Mechanisms. *Journal of Healthcare Engineering*, 2022, 1–14. https://doi.org/10.1155/2022/2032893

[19] Singh, M., & Mukherjee, A. (2022). CBAM-Enhanced EfficientNet for the Classification of MRI Brain Tumors. *Multimedia Tools and Applications*, 81, 24733–24751. https://doi.org/10.1007/s11042-021-11771-9

[20] Wahid, K., et al. (2022). Segmentation and Classification of Brain Tumors Utilizing a Hybrid CNN-SVM Framework. *Sensors*, 22(4), 1403. https://doi.org/10.3390/s22041403

[21] Nayak, D. R., Dash, R., Majhi, B., & Naik, B. (2021). A Dual-Module Method for Brain Tumor Classification Utilizing Adaptive Filtering and Support Vector Machines. *Journal of King Saud University – Computer and Information Sciences*. https://doi.org/10.1016/j.jksuci.2021.06.010

[22] Nickparvar, M. (2022). *Magnetic Resonance Imaging Dataset for Brain Tumors*. Kaggle. https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset

[23] Nickparvar, M. (2022). *Magnetic Resonance Imaging Dataset for Brain Tumors*. Kaggle. https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset

[24] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely Connected Convolutional Networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4700-4708). https://doi.org/10.1109/CVPR.2017.243

[25] Powers, D. M. W. (2011). Assessment: Transitioning from Precision, Recall, and F-Measure to ROC, Informedness, Markedness, and Correlation. *Journal of Machine Learning Technology*, 2(1), 37–63.

[26] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely Connected Convolutional Networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4700-4708). https://doi.org/10.1109/CVPR.2017.243

[27] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4700–4708). https://doi.org/10.1109/CVPR.2017.243

[28] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778). https://doi.org/10.1109/CVPR.2016.90

[29] Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1251–1258). https://doi.org/10.1109/CVPR.2017.195

[30] Howard, A. G., et al. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861.* https://arxiv.org/abs/1704.04861

[31] Deng, J., et al. (2009). ImageNet: A comprehensive hierarchical database of images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* https://doi.org/10.1109/CVPR.2009.5206848

[32] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4700–4708). https://doi.org/10.1109/CVPR.2017.243

[33] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for picture classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778). https://doi.org/10.1109/CVPR.2016.90

[34] Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1251–1258). https://doi.org/10.1109/CVPR.2017.195

[35] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Optimized convolutional neural networks for mobile visual applications. *arXiv preprint arXiv:1704.04861.* https://arxiv.org/abs/1704.04861

[36] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual elucidations from deep networks by gradient-based localization. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 618–626). https://doi.org/10.1109/ICCV.2017.74

[37] Afshar, P., Mohammadi, A., & Plataniotis, K. N. (2020). Classification of brain tumor types using capsule networks. In *Proceedings of the 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 3283–3286). https://doi.org/10.1109/EMBC44109.2020.9175871

[38] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2021). An image conveys the equivalent of 256 words: Transformers for large-scale image recognition. In *International Conference on Learning Representations*. https://arxiv.org/abs/2010.11929

[39] Wang, L., Yu, J., & Dou, Q. (2022). Swin-UNet: A Unet-inspired Pure Transformer for Medical Image Segmentation. *arXiv preprint arXiv:2105.05537*. https://arxiv.org/abs/2105.05537

[40] Hussain, E., Bird, J. J., & Faria, D. R. (2021). An investigation of CNN transfer learning for picture classification. *Neural Computing and Applications*, 33(7), 2951–2969. https://doi.org/10.1007/s00521-020-05121-3