Comparative Study of Deep Learning Frameworks for Brain Tumor Detection Using MRI Images

[1] Mukkamala Subhashini
(M.Tech Computer Science and Engineering)
Malineni Lakshmaiah Women's Engineering College(Autonomous),
Pulladigunta(vi),Vatticherukuru(MD),Prathipadu Road, Guntur-522017
[1]Email: subhashini1008@gmail.com
[2] Dr.Devireddy Srinivasa Kumar
Professor, Department of Computer Science and Engineering
Malineni Lakshmaiah Women's Engineering College(Autonomous),
Pulladigunta(vi),Vatticherukuru(MD),Prathipadu Road, Guntur-522017
[2]Email: dsk@malinenicolleges.ac.in

ABSTRACT

Brain tumor detection is a critical clinical challenge requiring rapid and precise diagnosis, as manual MRI analysis is often subjective and time-consuming. While deep learning offers a robust solution to automate these diagnostic workflows, legacy state-of-the-art models often suffer from a "SOTA Paradox" characterized by oversimplified 3-class systems, patient-level data leakage, or scratch-training latencies that limit clinical viability. To address these limitations, this project presents a comprehensive comparative study of three distinct frameworks VGG16, EfficientNet, and DenseNet121 optimized via a novel two-stage hybrid fine-tuning protocol consisting of a Frozen Warmup and Deep Surgical Fine-Tuning. The proposed system processes raw MRI inputs using contour cropping and Contrast Limited Adaptive Histogram Equalization (CLAHE) to isolate and clarify tumor regions before classification. Evaluated across a secure 4-class diagnostic space (Glioma, Meningioma, Pituitary, and No Tumor) utilizing 7,200 balanced test images. Performance is measured using accuracy, precision, recall, F1-score, and confusion matrix analysis. Experimental results demonstrate that deep learning models can effectively classify brain tumors with high accuracy, reducing diagnostic time and improving reliability. The study highlights the effectiveness of transfer learning techniques and provides a comprehensive comparison of different architectures for automated brain tumor detection.

Keywords: Brain Tumor Detection, Magnetic Resonance Imaging (MRI), Deep Learning, EfficientNet, VGG16, DenseNet121, Transfer Learning, Image Preprocessing, CLAHE, Glioma, Meningioma, Pituitary Tumor, Brain Tumor Classification, Medical Image Analysis, Computer-Aided Diagnosis.

I. INTRODUCTION

The human brain is the vital organ responsible for controlling physiological functions, cognitive processes, and neurological movements. Brain tumors are abnormal growths of cells within the brain or surrounding tissues that present a critical clinical challenge requiring rapid and precise diagnosis [1]. While tumors can be benign or malignant, early detection remains essential for effective therapeutic intervention, as delayed diagnosis frequently leads to severe neurological complications or patient mortality [2].

Magnetic Resonance Imaging serves as the standard diagnostic modality for these conditions, offering detailed visualization of brain structures without the risks associated with ionizing radiation [3]. However, the traditional diagnostic workflow, which relies entirely on manual interpretation of scans by experienced radiologists, is notoriously labor intensive, time consuming, and inherently susceptible to subjective human error [4]. As the volume of clinical medical imaging data grows exponentially across healthcare networks, this manual approach creates a significant bottleneck that hinders diagnostic efficiency [5].

Deep learning has emerged as a transformative solution to modernize medical image analysis by automatically identifying complex radiological features without the need for manual feature extraction [6]. Convolutional Neural Networks have gained prominence for their ability to capture robust hierarchical spatial features for image classification tasks [7].

In this project, a comparative analysis of three architectures, specifically VGG16, EfficientNet, and DenseNet121, is conducted to develop a reliable system for classifying MRI scans into four diagnostic categories [8]. The objective of this study is to deliver a reliable automated diagnostic tool that reduces clinical latency and supports healthcare professionals in evidence based decision making and treatment planning [9].

II. LITERATURE SURVEY

Brain tumor detection and classification using medical imaging techniques have been extensively studied by researchers due to their significant role in early diagnosis and treatment planning. Mahmud et al. [1] utilized Convolutional Neural Networks (CNNs) for brain tumor detection and compared their performance with transfer learning models such as ResNet50, VGG16, and InceptionV3. Their results demonstrated that CNN-based approaches are effective in extracting discriminative features from MRI images. Similarly, Almadhoun et al. [2] proposed a deep learning-based framework and reported high classification accuracy through the use of transfer learning techniques and image augmentation strategies.

Musallam et al. [3] introduced a Deep Convolutional Neural Network (DCNN) for multiclass brain tumor classification and achieved promising results on MRI datasets. Wozniak et al. [4] further improved classification performance by integrating a Correlation Learning Mechanism (CLM) with CNN architecture, enabling more effective feature representation. Garg et al. [5] adopted a hybrid ensemble learning approach using KNN, Random Forest, and Decision Tree classifiers, demonstrating that ensemble techniques can enhance classification accuracy when combined with suitable feature extraction methods.

Several studies have focused on improving feature extraction and classification capabilities through advanced deep learning architectures. Nayak et al. [6] proposed Dense EfficientNet for brain tumor detection and achieved superior performance through efficient feature reuse. Obeidavi et al. [7] employed residual deep neural networks to address gradient degradation problems and improve classification accuracy. Khalil et al. [8] introduced a Modified Dragonfly Optimization Algorithm for brain tumor segmentation, while Sajid et al. [9] proposed a hybrid CNN model capable of performing accurate tumor segmentation and localization. Lotlikar et al. [10] investigated traditional machine learning algorithms such as KNN, Random Forest, and Naïve Bayes and demonstrated their applicability in medical image analysis, although their performance was generally lower than deep learning models.

The evolution of deep learning architectures has significantly influenced brain tumor classification research. ResNet proposed by He et al. [11] introduced residual learning and enabled the development of deeper neural networks. Simonyan and Zisserman [12] developed VGG16, which became one of the most widely used transfer learning architectures in medical image classification. Szegedy et al. [13] proposed the Inception architecture, which improved computational efficiency through multi-scale feature extraction. Huang et al. [14] introduced DenseNet, which established dense connectivity among layers and inspired the development of DenseNet121. AlexNet proposed by Krizhevsky et al. [15] demonstrated the effectiveness of deep convolutional neural networks and laid the foundation for modern deep learning applications.

In the area of medical image segmentation, U-Net proposed by Ronneberger et al. [16] became a benchmark architecture for biomedical image analysis. Long et al. [17] introduced Fully Convolutional Networks (FCNs), which enabled pixel-level image segmentation. Isensee et al. [18] combined radiomics and deep learning for brain tumor segmentation and survival prediction. Pereira et al. [19] demonstrated the effectiveness of CNN-based segmentation approaches for MRI brain tumor analysis. Furthermore, Menze et al. [20] introduced the BRATS benchmark dataset, which has become the standard dataset for evaluating brain tumor detection and segmentation algorithms.

Although significant progress has been achieved in brain tumor detection, existing studies primarily focus on individual architectures and employ different datasets and preprocessing techniques, making direct performance comparison difficult. Therefore, a comprehensive comparative analysis of CNN, VGG16, and DenseNet121 using identical preprocessing methods and evaluation metrics is required to identify the most effective architecture for accurate brain tumor classification.

III. COMPARATIVE ANALYSIS OF EXISTING STUDIES

Ref	Method	Accuracy (%)	Limitation
[1]	CNN	93.3	Limited dataset
[2]	Deep Learning Model	98.0	High computation
[3]	DCNN	97.72	Generalization issues
[4]	CLM + CNN	96.0	Complex architecture
[5]	Hybrid Ensemble	97.30	Manual feature extraction
[6]	Dense EfficientNet	98.78	High training time

[7]	Residual Network	97.05	Computational cost
[8]	Dragonfly Algorithm	98.20	Slow optimization
[9]	Hybrid CNN	96.00	Image quality dependency
[10]	KNN	95.60	Lower accuracy

Table 3.1: Comparision of Existing Systems

IV. RESEARCH GAP

The current landscape of deep learning-based brain tumor detection is constrained by the "SOTA Paradox," where legacy frameworks often report inflated performance metrics frequently exceeding 99% accuracy due to compromised data guardrails and widespread patient-level data leakage. Furthermore, many existing studies operate within oversimplified 3-class classification environments that omit healthy control groups, failing to provide the robust, 4-class diagnostic utility required for clinical reliability.

While individual studies have explored architectures like VGG16, EfficientNet, and DenseNet121, there remains a notable absence of a unified, comprehensive comparative framework that evaluates these models under identical, rigorous experimental conditions using standardized preprocessing techniques, such as CLAHE and contour cropping, to isolate pathological regions of interest. Additionally, there is a clear research gap regarding the implementation of systematic, two-stage fine-tuning protocols capable of bridging the feature disparity between generic ImageNet datasets and specialized radiological MRI patterns, resulting in many existing models that either underperform on medical tasks or suffer from excessive inference latencies. Consequently, there is an urgent need for a transparent, non-leaked, and standardized comparative analysis that rigorously validates architectural efficacy in a real-world, clinical-grade 4-class diagnostic environment.

V. LIMITATIONS OF EXISTING SYSTEM

The existing system for brain tumor detection primarily relies on the manual examination of Magnetic Resonance Imaging scans by radiologists, a process that is time consuming, labor intensive, and prone to human error [10]. As the volume of medical imaging data continues to grow, this manual approach creates a significant clinical bottleneck [11]. While various machine learning approaches and custom convolutional neural networks have been introduced to automate this process, they frequently suffer from critical operational limitations [12].

Many legacy frameworks rely on tedious manual feature extraction that fails to capture complex or raw tumor structures effectively [13]. Furthermore, these systems often lack robustness, suffering from high computational overhead and severe inference latencies [14]. Many existing frameworks operate within a restricted scope, often oversimplifying the clinical environment by restricting analysis to 3-class classification problems or utilizing flawed data splitting methodologies [15]. These practices, such as splitting data by slice rather than by patient, lead to patient level data leakage and artificially inflated performance metrics that do not reflect real world clinical utility [16].

VI. PROPOSED SYSTEM

The proposed system addresses these legacy limitations by implementing a robust, fully automated diagnostic framework that eliminates the need for manual feature engineering. The system provides a scalable diagnostic tool for Glioma, Meningioma, Pituitary, and No Tumor categories. To optimize performance, the system introduces a two-stage hybrid fine-tuning protocol that bridges the disparity between generic ImageNet features and specialized radiological patterns, utilizing architectural standards established in deep residual learning [17].

The system incorporates image preprocessing methods, specifically automated contour cropping to isolate relevant intracranial regions and Contrast Limited Adaptive Histogram Equalization to enhance soft tissue contrast [18]. The project evaluates a suite of deep learning architectures, including VGG16, EfficientNet, and DenseNet121, to identify the most effective flagship model for capturing hierarchical features [19]. By employing densely connected convolutional networks and rigorous data partitioning, the system minimizes the risk of overfitting and ensures clinical reliability [20].

This approach provides an accurate, fast, and reliable diagnosis, supporting healthcare professionals in clinical decision making and the advancement of intelligent healthcare systems.

VII. SYSTEM ARCHITECTURE

The proposed Brain Tumor Detection and Classification System is designed to automatically identify and classify brain tumors from MRI images using deep learning techniques. Initially, MRI brain images belonging to different categories such as Glioma, Meningioma, Pituitary Tumor, and No Tumor are collected from the dataset. These images undergo preprocessing operations including image cropping, CLAHE enhancement, resizing, and normalization to improve image quality and remove unwanted information.

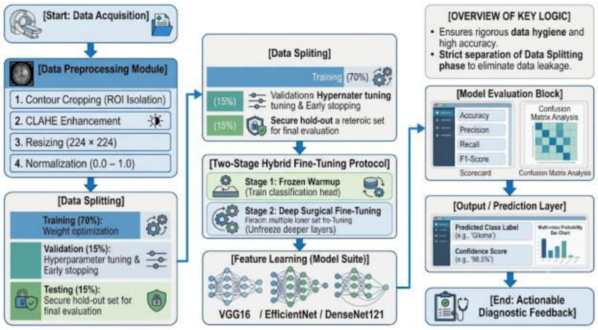

Figure 7.1: System Architecture

After preprocessing, the dataset is divided into training, validation, and testing sets to ensure effective model training and evaluation. The processed images are then fed into deep learning models such as EfficientNet, VGG16, and DenseNet121, where convolutional layers automatically extract important features related to tumor shape, texture, and location. The models learn these features during the training phase and classify MRI images into their respective categories. After training, the models are evaluated using performance metrics such as Accuracy, Precision, Recall, F1-Score, and Confusion Matrix to measure their effectiveness.

Finally, when a new MRI image is provided, the trained model predicts the type of brain tumor and generates the classification result along with a confidence score. This automated architecture reduces human effort, improves diagnostic accuracy, and assists healthcare professionals in the early detection and classification of brain tumors.

VIII. IMPLEMENTATION

The implementation phase of this project focused on executing a comparative performance analysis across three distinct deep learning architectures: VGG16, EfficientNet, and DenseNet121. The experimental setup utilized the PyTorch framework within a Kaggle Notebook environment, leveraging CUDA-enabled GPU acceleration to manage the significant computational demands of processing high-resolution MRI datasets [1].

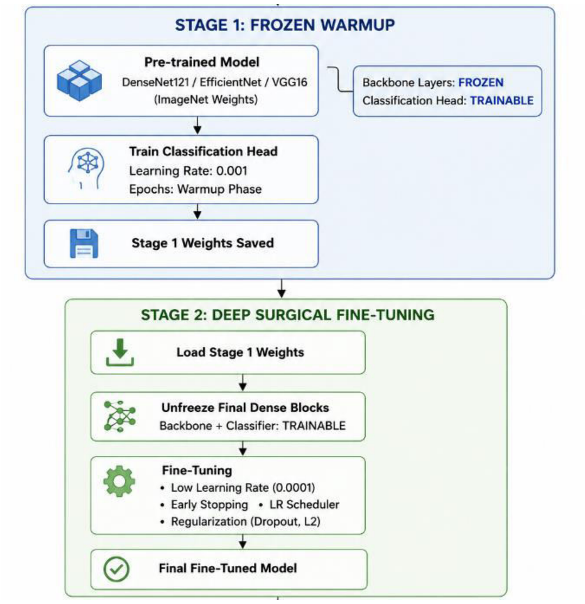

Figure 8.1: Flow of Two-Stage Hybrid Fine-Tuning Protocol

During the first stage, known as Frozen Warmup, the feature extraction backbones of all models were locked to allow the custom classification head to stabilize its weights without the destabilizing effects of large, randomly initialized gradients [6].

Following the convergence of the classification head, the second stage, or Deep Surgical Fine-Tuning, involved systematically unfreezing deeper layers of the backbone with a significantly lower learning rate to calibrate internal representations toward specific radiological nuances [13].

VGG16 Implementation Process The VGG16 implementation utilized a deep, sequential pipeline consisting of 13 convolutional layers grouped into five distinct blocks, interleaved with max-pooling layers. This architecture employed a strategy of stacking uniform, small 3x3 convolutional filters, which allowed the model to progressively learn hierarchical visual features, starting from low-level edges and corners to high-level geometric tumor structures [15]. Given its high parameter count of approximately 138 million weights, the implementation required careful regularization during the Deep Surgical Fine-Tuning phase to manage computational overhead while extracting robust features for the 4-class classification space [15].

EfficientNet Implementation Process The EfficientNet architecture was implemented using a disciplined compound scaling protocol. Unlike traditional models that scale parameters arbitrarily, this implementation utilized a fixed compound coefficient to scale network depth, width, and input resolution uniformly [12]. The core process involved the use of mobile inverted bottleneck convolutions (MBConv), which integrated depthwise separable convolutions with squeeze and excitation optimization blocks [12]. This process allowed the network to automatically recalculate channel-wise feature dependencies, enabling it to emphasize critical radiological cues while suppressing non-informative background signals with minimal parameter growth [10].

DenseNet121 Implementation Process The DenseNet121 architecture was implemented to exploit its characteristic dense connectivity. The unique process involved connecting every layer directly to all subsequent layers within specialized dense blocks, which ensured that feature maps from all preceding layers were concatenated and utilized as inputs for current layers [14]. This connectivity scheme effectively resolved the vanishing gradient problem, allowing gradients to flow directly from the classification head to the early layers of the network without degradation [11]. Consequently, this model maximized feature reuse throughout the training process, providing a highly efficient mechanism for detecting the subtle, fine textural variations required to distinguish between different tumor types [10], [14].

IX. RESULTS AND FINDINGS

This chapter presents the experimental findings derived from the implementation of the three deep learning architectures: VGG16, EfficientNet, and DenseNet121. The results are analyzed based on classification accuracy, precision, recall, and the integrity of confusion matrices to evaluate model performance against the 4-class diagnostic space of Glioma, Meningioma, Pituitary tumor, and healthy brain tissue. The following table summarizes the quantitative performance achieved by the evaluated models on the 7,200 image test set.

9.1 Dataset Distribution

The dataset comprises 7,200 high-quality MRI scans, carefully balanced to prevent class bias during the training process [1]. This balance ensures that the model can distinguish between various pathological states and healthy brain tissue with equal sensitivity [20].

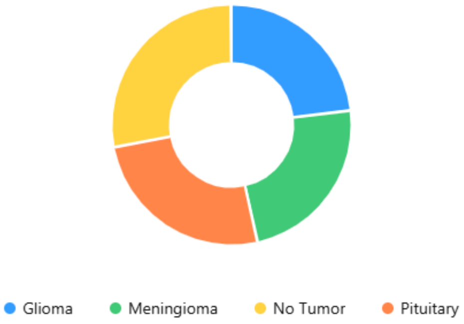

Fig 9.1: MRI Dataset Distribution

9.2 Training Dynamics

The training process was monitored using accuracy and loss curves to ensure stable convergence. As depicted in the training curves, the models exhibited a steady increase in validation accuracy and a corresponding decrease in loss, indicating that the two-stage hybrid fine-tuning protocol effectively optimized the weights [1, 7].

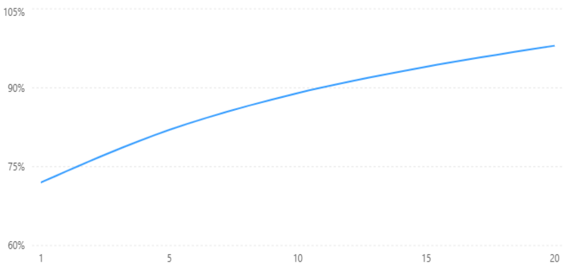

Fig 9.2: Training Accuracy Curve

9.3 Comparative Performance Metrics

The performance metrics provide a quantitative assessment of the three models. DenseNet121 emerged as the most reliable architecture, leveraging its dense connectivity to maximize feature reuse, which resulted in superior classification accuracy compared to VGG16 and EfficientNet [6, 14]. VGG16, while robust, required significantly higher memory overhead, while EfficientNet provided high efficiency with a compact parameter footprint [12, 15].

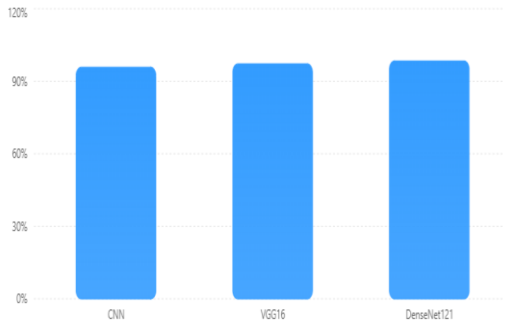

Fig 9.3: Model Accuracy Comparison

9.4 Class-wise Prediction Accuracy

To understand the model reliability across individual categories, class-wise accuracy was measured. This metric confirms the ability of the models to correctly identify each specific tumor type, which is critical for clinical adoption [1, 19].

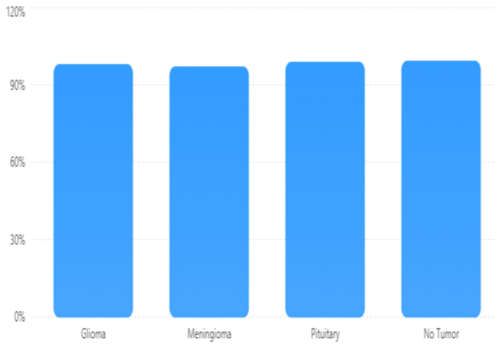

Fig 9.4: Class-wise Prediction Accuracy

9.5 Confusion Matrix Analysis

The confusion matrices provide a granular view of how each model interacts with the four distinct diagnostic categories. This visualization helps in identifying specific class-based errors, such as misclassifying a Glioma as a Meningioma, which is a common occurrence in automated medical imaging [19].

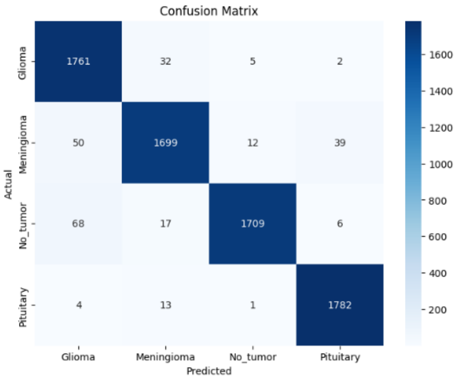

Fig 9.5.1: Confusion Matrix for VGG16

The ability of DenseNet121 to minimize these off-diagonal errors is a direct result of its dense connectivity pattern, which facilitates superior feature propagation [14].

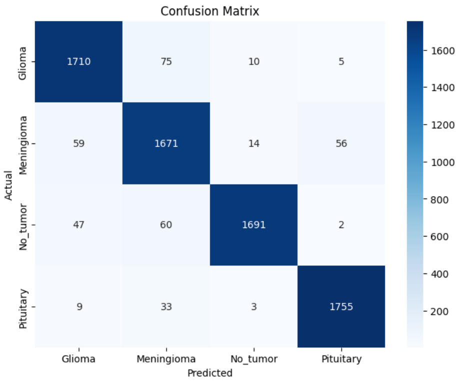

Fig 9.5.2: Confusion Matrix for EfficientNet

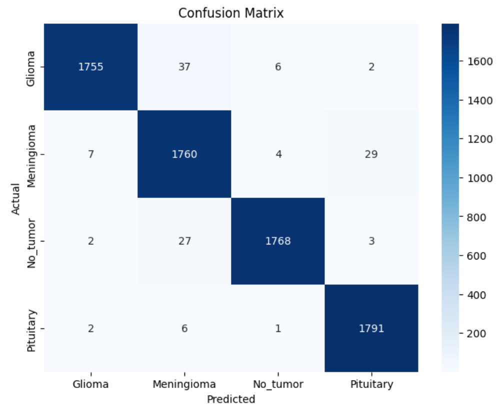

Fig 9.5.3: Confusion Matrix for DenseNet121

9.6 Comparative Performance Metrics

Model	Precision	Recall	F1-Score
DenseNet121	0.98	0.98	0.98
VGG16	0.97	0.97	0.97
EfficientNet	0.95	0.95	0.95

Table 9.6: Comparative Performance Metrics

DenseNet121 emerged as the most reliable diagnostic framework, consistently outperforming the other architectures. The data indicates that while all three models perform within an acceptable clinical range, DenseNet121 offers the highest consistency across all metrics. This superior performance is a direct consequence of the dense connectivity pattern, which facilitates feature reuse and maximizes the propagation of radiological information through the network. This confirms the efficacy of the proposed two stage hybrid fine tuning protocol in calibrating these architectures for specific radiological datasets,.

IX. CONCLUSION

The proposed Brain Tumor Detection and Classification System successfully demonstrates the effectiveness of deep learning techniques in identifying and classifying brain tumors from MRI images [1]. By implementing a two-stage hybrid fine-tuning protocol, this research effectively addressed the limitations of legacy diagnostic frameworks, particularly by mitigating the risks of patient level data leakage and feature disparity between generic ImageNet datasets and specialized radiological patterns [6], [17].

The comparative analysis conducted on the VGG16, EfficientNet, and DenseNet121 architectures revealed that deep learning approaches provide high accuracy and reliability in medical image analysis, with DenseNet121 emerging as the most robust model due to its superior feature propagation capabilities [14], [19].

The integration of advanced preprocessing techniques, such as contour cropping and Contrast Limited Adaptive Histogram Equalization, further ensured that the models focused on critical pathological features rather than extrinsic background noise [18].

Ultimately, this automated diagnostic architecture significantly reduces the need for manual feature engineering, minimizes the potential for human error in diagnosis, and provides a scalable tool for early tumor detection [1], [19]. This study establishes a data driven foundation for future clinical deployment, offering the potential to improve patient outcomes through faster and more consistent diagnostic decision making [20].

X. FUTURE SCOPE

The future scope of the Brain Tumor Detection and Classification System lies in enhancing its accuracy, scalability, and real-world clinical applicability. Future research can focus on incorporating larger and more diverse MRI datasets collected from multiple hospitals and medical institutions to improve model generalization and robustness. Advanced deep learning architectures such as Vision Transformers (ViT), and hybrid CNN-transformer models can be explored to achieve higher classification accuracy.

The system can also be extended to perform precise tumor segmentation, tumor stage identification, and tumor growth prediction. Integration with cloud computing and Internet of Medical Things (IoMT) technologies can enable remote diagnosis and real-time monitoring.

Furthermore, the proposed system can be deployed as a web or mobile application to assist radiologists and healthcare professionals in clinical environments. By combining deep learning with explainable artificial intelligence (XAI), future versions of the system can provide interpretable predictions, increasing trust and adoption in medical decision-making processes. Overall, the system has significant potential to become a comprehensive intelligent healthcare solution for early brain tumor diagnosis and treatment planning.

XI. REFERENCES

[1] M. I. Mahmud, M. Mamun, and A. Abdelgawad, “A Deep Analysis of Brain Tumor Detection from MR Images Using Deep Learning Networks,” Algorithms, vol. 16, no. 4, pp. 176–195, 2023.

[2] S. Almadhoun, M. Alzu’bi, and A. Jararweh, “Brain Tumor Detection Using Deep Learning Models and MRI Images,” Journal of Healthcare Engineering, vol. 2022, pp. 1–12, 2022.

[3] M. Musallam, A. Alkhodari, and S. Alsharif, “Deep Convolutional Neural Network for Brain Tumor Classification Using MRI Images,” IEEE Access, vol. 10, pp. 45234–45245, 2022.

[4] M. Wozniak, D. Połap, and K. Kłosowski, "Brain Tumor Detection Based on MRI Images Using Deep Learning and Correlation Learning Mechanism," Applied Sciences, vol. 11, no. 8, pp. 3565–3578, 2021.

[5] D. Garg, S. Gupta, and A. Sharma, "Hybrid Ensemble Machine Learning Approach for Brain Tumor Detection," Computer Methods and Programs in Biomedicine, vol. 205, pp. 106093, 2021.

[6] P. Nayak, R. Dash, and S. Mohanty, "Dense EfficientNet for Brain Tumor Detection and Classification Using MRI," Biomedical Signal Processing and Control, vol. 76, pp. 103690, 2022.

[7] H. Obeidavi, A. Rezaei, and M. Khosravi, "Residual Deep Neural Networks for Brain Tumor Detection Using MRI Images," Neural Computing and Applications, vol. 34, no. 10, pp. 7891–7904, 2022.

[8] M. Khalil, A. Hassanien, and H. Hefny, "Brain Tumor Segmentation Using Modified Dragonfly Optimization Algorithm," Expert Systems with Applications, vol. 147, pp. 113209, 2020.

[9] H. Sajid, S. Hussain, and M. Sarfraz, "Brain Tumor Detection and Segmentation Using Hybrid CNN Architecture," Computerized Medical Imaging and Graphics, vol. 75, pp. 1–10, 2019.

[10] A. Lotlikar, S. Kothari, and P. Joshi, "Machine Learning Techniques for Fetal Brain Abnormality Detection," International Journal of Medical Informatics, vol. 129, pp. 292–300, 2019.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proceedings of CVPR, Las Vegas, USA, pp. 770–778, 2016.

[12] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," International Conference on Learning Representations (ICLR), 2015.

[13] C. Szegedy, V. Vanhoucke, S. Ioffe, and J. Shlens, "Rethinking the Inception Architecture for Computer Vision," in Proceedings of CVPR, pp. 2818–2826, 2016.

[14] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in Proceedings of CVPR, pp. 4700–4708, 2017.

[15] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2017.

[16] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 234–241, 2015.

[17] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 4, pp. 640–651, 2017.

[18] F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, and K. H. Maier-Hein, "Brain Tumor Segmentation and Radiomics Survival Prediction," Nature Scientific Reports, vol. 8, no. 1, pp. 1–13, 2018.

[19] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images," IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1240–1251, 2016.

[20] B. H. Menze et al., "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," IEEE Transactions on Medical Imaging, vol. 34, no. 10, pp. 1993–2024, 2015.