2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI) | 979-8-3503-6052-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/IATMSI60426.2024.10502596

# Computer-Aided Diagnostic System for Brain Tumor Classification using Explainable AI

S.T. Padmapriya
*Department of Applied Mathematics and Computational Science*
stpca@tce.edu

M.S. Gayathri Devi
*Thiagarajar College of Engineering, Madurai*
TamilNadu, India
Msg3dv@gmail.com

***Abstract*—The field of computer-aided diagnosis (CAD) of brain tumors has been transformed by developments in medical imaging and artificial intelligence. The accuracy and interpretability of brain tumor classification are improved in this research using Explainable AI (XAI) techniques. A timely and correct diagnosis of brain tumors is essential for the best possible care and treatment of the patient. In contrast, traditional machine learning models often lack transparency and interpretability, making it difficult for clinicians to rely on their judgment. This study uses a Grad-Cam algorithm to create an understandable and interpretable CAD system for classifying brain tumors. Our approach not only achieves high classification accuracy, but also provides physicians with insights into the decision-making process, improving their understanding and confidence in the system's recommendations. We evaluate our model using a large and diverse dataset and compare it to modern deep learning models and traditional CAD systems. The results show that our CAD system with XAI extensions not only achieves improved accuracy but also provides useful insights into the decision-making process. Using this method, doctors may be able to diagnose patients more accurately.**

***Index Terms*—Explainable AI (XAI), Computer Aided Diagnostic System (CAD), Brain Tumor Classification, VGG-16.**

## I. Introduction

Early detection of brain tumors is essential since it directly correlates with improved treatment outcomes. Numerous neurological symptoms can be brought on by brain tumors, and early diagnosis enables doctors to develop efficient treatment regimens. Early intervention can slow the growth of a tumor, reduce symptoms, and enhance the likelihood that a patient will receive an effective treatment, potentially saving lives. The burden of cancer in India is greatly increased by brain tumors. Effective brain tumor identification and therapy are crucial parts of comprehensive cancer care as the nation's cancer incidence rate rises. Early detection can aid in the effective and efficient management of this burden and ensuring that resources are used correctly to fulfill the needs of patients with brain tumors. Accurate diagnosis of brain tumors influences patients' quality of life as well as their physical condition. Brain tumors can produce debilitating symptoms such as cognitive impairment, movement dysfunction, and sensory issues. These symptoms can be relieved and the general health and functioning of brain tumor patients can be enhanced with early detection and treatment of these tumors. Individual treatment regimens are necessary for brain tumors of various forms and sites. Healthcare professionals can create treatment strategies that are specifically suited to the unique characteristics of the tumor with early and accurate detection. Early discovery guarantees that patients receive the most effective and appropriate care, whether that involves chemotherapy, radiation therapy, surgical resection, or a combination of these modalities.

It may be expensive and necessitate surgery, diagnostic procedures, and continuing medical attention to treat a brain tumor. Early identification reduces overall treatment expenses by preventing costly surgeries that could be necessary later on. This financial assistance enables patients and their families to manage and lessen the cost of long-term therapy. Early detection enables patients to participate in research studies and clinical trials, which are essential to expanding our knowledge of brain tumors and creating more efficient treatments. By identifying cases early on, researchers can look into novel therapies and diagnostics, potentially advancing the discipline and offering hope for improved outcomes in the future.

The classification of brain tumors has been profoundly and fundamentally changed by Explainable AI (XAI), which also offers various advantages in the fields of medical imaging and diagnostics. One of the primary issues with AI in healthcare, particularly for life-threatening illnesses like brain tumors, is the lack of confidence and understanding among healthcare professionals. XAI, which offers visible and understandable insights into how AI algorithms arrive at a diagnosis, fills this gap. Since AI systems are more transparent, doctors are more likely to adopt them. Healthcare professionals can use AI technologies more effectively thanks to XAI solutions. Doctors can make more informed judgments about patient treatment by having a deeper understanding of the rationale behind AI-generated diagnoses. Increased diagnostic accuracy is the outcome of the synergy between human expertise and AI capabilities created by this collaboration.

XAI systems reduce diagnostic uncertainty by providing a concise and understandable justification for their findings. Increased diagnostic confidence in brain tumor cases can result in earlier interventions, more accurate treatment planning, and improved patient outcomes. XAI can assist in finding inaccuracies or biases in the information or algorithms used to categorize brain tumors. Healthcare professionals can spot mistakes made by the AI model and correct them, preserving the fairness and accuracy of the AI system. XAI systems can adapt and learn by interacting with medical personnel. The AI model's ability to categorize brain tumors becomes more accurate and efficient over time because to this iterative learning process. In the healthcare industry, ethical principles including transparency, justice, and accountability are common. By offering an audit trail for decision-making processes, XAI addresses these problems and makes it simpler to abide by moral standards and legal requirements. Brain tumor classification is a challenging task, and misinterpretation can be fatal. XAI systems can assist decrease the likelihood of misdiagnosis by allowing doctors to counter-verify and confirm the results produced by AI by providing thorough explanations of their conclusions.

Related works are discussed in Section II. The resources and measures used in this study are described in Section III. The methodology of the article is discussed in Section IV. The conclusions and discussions are explained in Section V. Section VI talks about the conclusion and upcoming works.

## II. RELATED WORKS

The following works relate to the classification of brain tumors. The authors of [1] investigate the use of deep learning techniques to classify brain tumors and offer doctors promising perspectives. It shows how deep learning can be used to increase diagnostic accuracy and efficiency. The benefits come from achieving high classification accuracy, which is crucial for proper diagnosis and treatment planning. However, this has the disadvantage of requiring extensive computing resources and a large data set for training, which can be difficult in resource-limited healthcare environments. Waghmare et al. conducted a study on deep learning to classify brain tumors, highlighting the confluence of technology in healthcare. The advantage is that it enables real-time monitoring and remote diagnosis, which has the potential to revolutionize healthcare. However, a major disadvantage is that the integration of IoT devices leads to security and privacy issues in healthcare systems. This requires strict security and privacy procedures [2]. The study in [3] uses multimodal data and rigorous feature selection to improve the accuracy of brain tumor diagnosis. Benefits include increased accuracy and complete diagnosis. Feature selection and feature engineering still require expert knowledge, which can take time.

The authors show how pre-trained models can be adapted in medical image analysis, minimizing the need for large tagged data. Benefits include reduced labeled data requirements and the potential for improved performance. However, the selection of pre-trained models and their suitability for specific tasks can have an impact on performance [4]. The authors of [5] present a decision support system that supports doctors in diagnostic processes. Benefits include improved diagnostic accuracy and better patient outcomes. However, this requires integration into existing healthcare systems, which can be difficult. The authors of [6] studied brain tumor categorization using hybrid techniques that used multiple machine learning algorithms. The benefits come from improved categorization results achieved by combining approaches. However, model selection and hyperparameter optimization can be difficult. The authors of [7] use a CNN-LSTM technique to identify brain tumors in MRI scans, considering sequential data. The benefits come from a more comprehensive study of brain tumor detection. However, complex sequential models such as LSTM can be computationally intensive. The authors of [8] focus on advanced deep learning approaches to improve the accuracy of brain tumor detection and classification. Benefits include improved accuracy in diagnosing brain tumors. However, training may require larger datasets and complicated models may be computationally intensive. The authors of [9] describe a unique brain tumor classification strategy that combines deep feature fusion with well-known machine learning classifiers. The benefits come from providing a clear perspective for classification. However, the interpretability of the model and the selection of the classifier can be difficult.

The authors of [10] use fine-tuned models and ensemble approaches to increase the robustness and reliability of brain tumor categorization. The advantages arise from increased classification robustness. Still, additional computing resources as well as careful model combination tactics may be required. The authors of [11] address the classification of brain tumors using deep learning algorithms, with a focus on the possibility of better diagnostic accuracy. Benefits include excellent categorization accuracy and possible health benefits. Deep learning models may require a huge data set for training, which may be a limitation for some healthcare organizations. The authors of [12] offer a deep learning system to categorize brain tumors using MRI scans, offering a promising opportunity for accurate diagnosis. Deep learning models are capable of detecting complicated patterns in MRI scans, potentially increasing classification accuracy. Model training can require a large amount of labeled data as well as computational resources. A hybrid deep learning-based brain tumor classification system with multiple methods for improved diagnostic accuracy is described by the authors in [13]. Hybrid models are able to achieve more reliable classification results by taking advantage of many methods. It can be challenging to select a model, tune the hyperparameters, and integrate the components.

Using Deep Neural Network (DNN) and Convolutional Neural Network (CNN) models, the authors of [14] identify and classify brain tumors. Due to the expertise of CNN and DNN models in image processing, tumors can be identified precisely. Large data sets may be required for model training and computationally intensive complex neural network topologies. The authors of [15] investigate image processing for MRI-based categorization of brain tumors using deep learning techniques. The categorization process can be simplified through deep learning, which can automatically extract relevant information from MRI scans. The performance of the deep learning model can be significantly influenced by the amount and quality of the training data. Using a combination of deep learning and machine learning techniques, [16] focuses on the early detection of brain tumors in MRI images. Early diagnosis of brain tumors can lead to better treatment outcomes and higher patient survival rates. Data collection can be difficult and computationally intensive for hybrid approaches. The authors of [17] classify brain tumors in magnetic resonance imaging (MRI) using deep learning and wavelet transform. When wavelet transform and deep learning are combined, it is possible to thoroughly examine MRI data to provide accurate categorization. Wavelet transform and other preparation methods could make the classification pipeline more complex.

The authors of [18] describe an intelligently regulated deep residual learning system for classifying brain tumors based on MRI images. Deep residual learning frameworks are well suited to image classification problems and have the potential for high accuracy. These models can be computationally intensive and may require large data sets for training. The authors of [19] describe a hybrid deep learning model for brain tumor classification that integrates many strategies for increased diagnostic accuracy. Hybrid models can take advantage of the properties of different methods and potentially lead to more robust classification results. Model selection, hyperparameter tuning, and component integration can be difficult. The authors of [20] focus on efficient automatic categorization of brain tumors using an optimized hybrid deep neural network. Optimized deep neural networks can enable efficient and accurate categorization. Model optimization may require a thorough understanding of deep learning techniques and architectures.

## III. MATERIALS AND METRICS

Two datasets were used in our experiments. The Kaggle brain tumor dataset was used for validation and testing after training the model on the BraTS 2021 dataset. The Brain Tumor Segmentation Challenge (BraTS) 2021 dataset is an important tool for research in neuro-oncology and medical imaging. It includes 3D MRI scans of the brain created using many imaging modalities, including T1-weighted, T1-weighted post contrast, T2-weighted, and FLAIR images. This information is used to build and test advanced machine learning techniques and models to increase the precision of brain tumor segmentation, classification, and complete characterization. BraTS 2021 continues to play an important role in pushing the boundaries of brain tumor detection and therapy using state-of-the-art computational methods, ultimately benefiting patients and neuro-oncology researchers, even though the actual number of volumes is 2000.

The “Brain Tumor” and “No Brain Tumor” classes in the Kaggle dataset indicate the presence or absence of brain tumors. It contains 155 files in the brain tumor category and 98 files in the non-brain tumor category. The collection is 16 MB in size, with the images mostly taken using MRI (magnetic resonance imaging) and capturing brain structures and disorders. Photos vary in width and height and are saved in various file formats, including JPG. Although the actual number of photos in the dataset for each class is unknown, this dataset is an excellent resource for training and testing models in the field of brain tumor detection and classification. Metrics such as accuracy, precision, recall and F1 score were derived to evaluate the model.

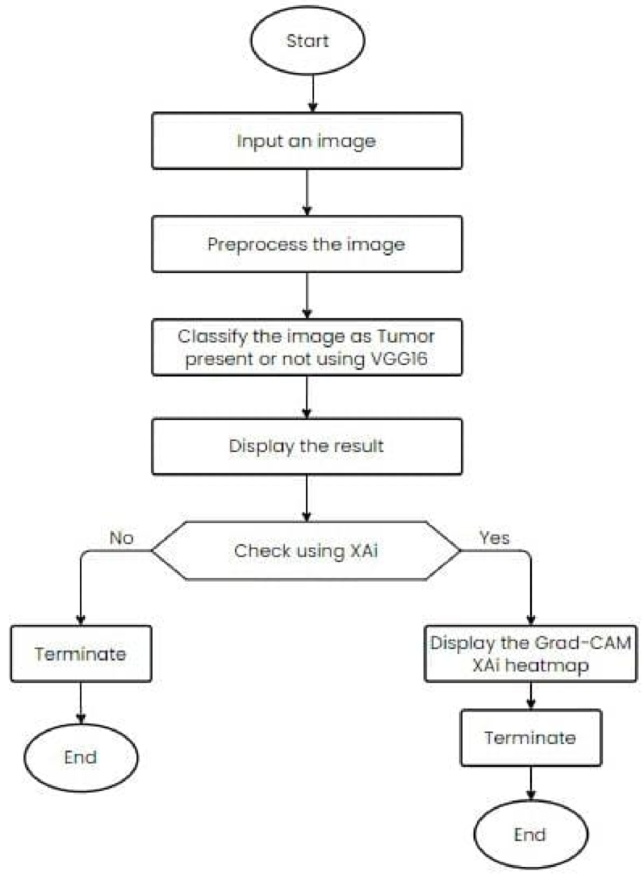


Fig. 1. Workflow of the Proposed Work

## IV. METHODOLOGY

The methodology of the proposed work is shown in Figure 1. In the proposed work, the VGG-16 model is used. The VGG-16 architecture is shown in Figure 2. An image with a fixed size of $224 \times 224$ pixels can be input into VGG16. A series of convolutional layers are applied to the input image. These layers use small squares of the input data to perform convolution operations on the input and learn properties such as edges, corners, and textures. There are a total of 13 convolution layers in VGG16. Each convolutional layer is followed by the SWISH activation function. SWISH imparts nonlinearity to the model, allowing it to recognize complicated data patterns. After a few convolution layers, the max-pooling layers are inserted. By using max-pooling, the spatial dimensions of the input volume can be reduced while still preserving the most important data. It helps reduce model computation and control the complexity of overfitting. One or more fully connected layers are applied after flattening and traversing the output of the final convolution and pooling layers. These layers acquire high-level properties and communicate them in a way that is conducive to classification. Three fully connected layers form VGG16. The number of neurons in the final fully connected layer is equal to the number of classes in the data set. Using a softmax activation function, the raw values from the previous level are converted into probabilities. The class with the highest probability is the projected class for the input image.

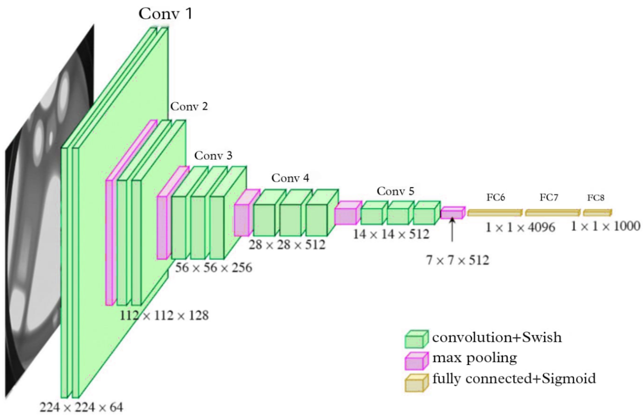


Fig. 2. VGG-16 Architecture for Brain Tumor Classification

We used Grad-Cam to show doctors the exact location of the tumor and how they arrived at their prognosis. Grad-CAM (Gradient-weighted Class Activation Mapping) is a method for viewing and understanding which components of an image are critical for a neural network to generate a specific prediction. Grad-CAM works in six steps. After a target image is passed through the neural network, the predicted class value is calculated. In other words, the network predicts which class the input image belongs to. The gradient of the predicted score with respect to the last convolutional layer of the network is calculated. This gradient reflects how much the output value would change if the activations of the last convolutional layer were changed. The gradients are then globally averaged to determine the significance weights of the neurons. This is an important phase because it captures the meaning of each feature map in the final convolutional layer.

The average pooling step guarantees that the important weights match the overall importance of the feature map. In the final convolutional layer, each feature map is multiplied by its corresponding significance weight. This phase reveals the feature map regions that had the greatest impact on the final prediction. Grad-CAM identifies which areas of the feature maps had the greatest influence on the prediction by weighting the feature maps. After the weighted combination, ReLU activation is used. This step ensures that only positive influences are taken into account. Negative influences (feature maps that can reduce predictive value) are set to zero and areas that contribute to the prediction are highlighted. The Grad-CAM heat map is created by summing the ReLU-enabled weighted feature maps. This heat map shows the areas of the input image where the neural network focused its attention to make the prediction. This heat map is often up scaled to fit the proportions of the source image before being displayed on the input image.

## V. RESULTS AND DISCUSSIONS

The dataset is divided into 70% for training, 15% for testing, and 15% for validation. If the validation accuracy exceeds 99.5%, training stops to avoid overfitting the model. With the default Image Net weights, the pre-trained model VGG16 is used. Contained in the hidden layers of the dense layers is the swish activation function with a fully convolutional activation layer that is intended to be sigmoid for binary classification. The model was trained for 500 epochs using the "binary cross-entropy" loss, the Adam optimizer, and the "accuracy" metrics. This resulted in a training accuracy of 99.4 percent. The model was saved as a 21.1 MB .h5 file. Grad Cam was trained using this model and a visualization "jet" color map that highlights the pixels that help classify and detect brain tumors in the input image. The trained model is deployed for free on the Streamlet cloud platform. Using the provided MRI scans of the brain, the website takes an image and determines whether a tumor is present or not. If a tumor is present and explained using Grad Cam, a button is pressed and the image is visualized using Grad Cam XAI.

VGG16's deep layers quickly extract detailed patterns from medical images, enabling the identification of small features such as tumors or lesions, which is crucial for accurate medical diagnoses. VGG16 can use pre-trained weights from datasets such as Image Net. This enables transfer learning, where the model is tuned to medical images, saving time and processing resources while benefiting from the general image properties gained. VGG16's transparent layered architecture enables the display of learned features. This interpretability is critical in medical applications because it helps doctors understand which visual elements the model is focusing on and aids in diagnosis and decision making. The model was trained for 500 epochs. The accuracy and loss graph are shown in Figure 3 and Figure 4, respectively. The computer-aided diagnostic system is shown in Figure 5. When the image to be predicted is loaded, the classification results with the Grad-Cam feature map are displayed.

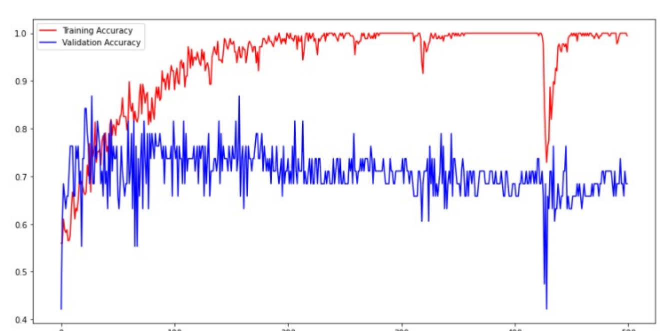


Fig. 3. Accuracy Graph of the Proposed Model

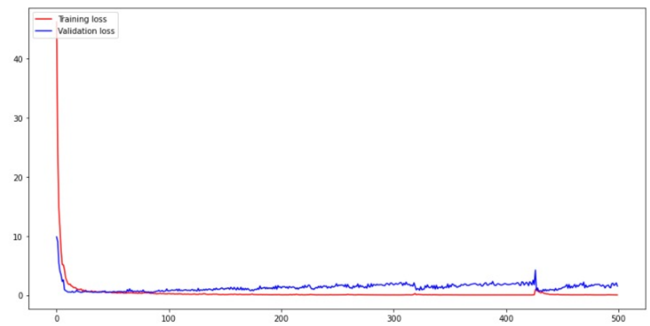


Fig. 4. Loss Graph of the Proposed Model

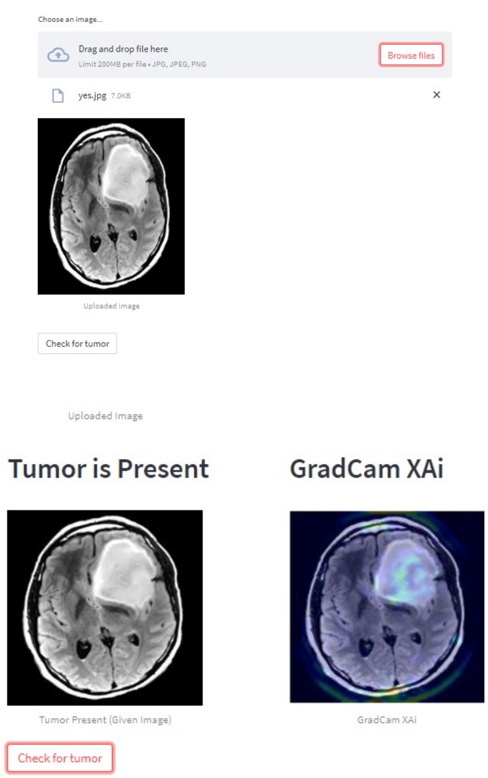


Fig. 5. Computer Aided Diagnostic System for Brain Tumor Detection using XAI

## VI. Conclusion and Future Works

The use of Explainable AI (XAI) methods in this study represents a significant step forward in addressing the field's critical difficulties regarding accuracy and interpretability. Although timely and accurate identification of brain tumors is an essential part of good patient care, the historical opacity of standard machine learning models has created a trust barrier for physicians. We have successfully developed a CAD system that not only achieves remarkable classification accuracy, but also provides clear insight into the decision-making process, increasing healthcare professionals' confidence in the system's recommendations. Our research highlights the improved performance and critical insights our XAI-enhanced CAD system delivers through rigorous validation on a broad dataset and comparisons to traditional CAD systems and state-of-the-art deep learning models. Ultimately, this technology has the potential to empower physicians by improving diagnostic precision, reducing subjectivity, and providing deep insights into the complexities of medical data, transforming the landscape of brain tumor diagnosis and ultimately patient outcomes.

## References


[1] A. Saleh, R. Sukaik, and S. S. Abu-Naser, "Brain tumor classification using deep learning," in *2020 International Conference on Assistive and Rehabilitation Technologies (iCareTech)*. IEEE, 2020, pp. 131–136.

[2] V. K. Waghmare and M. H. Kolekar, "Brain tumor classification using deep learning," *Internet of things for healthcare technologies*, pp. 155–175, 2021.

[3] M. A. Khan, I. Ashraf, M. Alhaisoni, R. Damaševičius, R. Scherer, A. Rehman, and S. A. C. Bukhari, "Multimodal brain tumor classification using deep learning and robust feature selection: A machine learning application for radiologists," *Diagnostics*, vol. 10, no. 8, p. 565, 2020.

[4] A. Rehman, S. Naz, M. I. Razzak, F. Akram, and M. Imran, "A deep learning-based framework for automatic brain tumors classification using transfer learning," *Circuits, Systems, and Signal Processing*, vol. 39, pp. 757–775, 2020.

[5] M. I. Sharif, M. A. Khan, M. Alhussein, K. Aurangzeb, and M. Raza, "A decision support system for multimodal brain tumor classification using deep learning," *Complex & Intelligent Systems*, pp. 1–14, 2021.

[6] S. Rinesh, K. Maheswari, B. Arthi, P. Sherubha, A. Vijay, S. Sridhar, T. Rajendran, Y. A. Waji *et al.*, "Investigations on brain tumor classification using hybrid machine learning algorithms," *Journal of Healthcare Engineering*, vol. 2022, 2022.

[7] R. Vankdothu, M. A. Hameed, and H. Fatima, "A brain tumor identification and classification using deep learning based on CNN-LSTM method," *Computers and Electrical Engineering*, vol. 101, p. 107960, 2022.

[8] T. Sadad, A. Rehman, A. Munir, T. Saba, U. Tariq, N. Ayesha, and R. Abbasi, "Brain tumor detection and multi-classification using advanced deep learning techniques," *Microscopy Research and Technique*, vol. 84, no. 6, pp. 1296–1308, 2021.

[9] H. Kibriya, R. Amin, A. H. Alshehri, M. Masood, S. S. Alshamrani, A. Alshehri *et al.*, "A novel and effective brain tumor classification model using deep feature fusion and famous machine learning classifiers," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.

[10] N. Noreen, S. Palaniappan, A. Qayyum, I. Ahmad, and M. O. Alassafi, "Brain Tumor Classification Based on Fine-Tuned Models and the Ensemble Method," *Computers, Materials & Continua*, vol. 67, no. 3, 2021.

[11] K. S. Kumar, A. Bansal, and N. P. Singh, "Brain Tumor Classification Using Deep Learning Techniques," in *International Conference on Machine Learning, Image Processing, Network Security and Data Sciences*. Cham: Springer Nature Switzerland, 2022, pp. 68–81.

[12] M. Aamir, Z. Rahman, Z. A. Dayo, W. A. Abro, M. I. Uddin, I. Khan, A. S. Imran, Z. Ali, M. Ishfaq, Y. Guan *et al.*, "A deep learning approach for brain tumor classification using MRI images," *Computers and Electrical Engineering*, vol. 101, p. 108105, 2022.

[13] A. Raza, H. Ayub, J. A. Khan, I. Ahmad, A. S. Salama, Y. I. Daradkeh, D. Javeed, A. Ur Rehman, and H. Hamam, "A hybrid deep learning-based approach for brain tumor classification," *Electronics*, vol. 11, no. 7, p. 1146, 2022.

[14] C. L. Choudhury, C. Mahanty, R. Kumar, and B. K. Mishra, "Brain tumor detection and classification using convolutional neural network and deep neural network," in *2020 International Conference on Computer Science, Engineering and Applications (ICCSEA)*. IEEE, 2020, pp. 1–4.

[15] K. N. Qodri, I. Soesanti, and H. A. Nugroho, "Image analysis for mri-based brain tumor classification using deep learning," *IJITEE (International Journal of Information Technology and Electrical Engineering)*, vol. 5, no. 1, pp. 21–28, 2021.

[16] E. M. Senan, M. E. Jadhav, T. H. Rassem, A. S. Aljaloud, B. A. Mohammed, Z. G. Al-Mekhlafi *et al.*, "Early diagnosis of brain tumour MRI images using hybrid techniques between deep and machine learning," *Computational and Mathematical Methods in Medicine*, vol. 2022, 2022.

[17] A. M. Sarhan *et al.*, “Brain tumor classification in magnetic resonance images using deep learning and wavelet transform,” *Journal of Biomedical Science and Engineering*, vol. 13, no. 06, p. 102, 2020.
[18] H. Mehnatkesh, S. M. J. Jalali, A. Khosravi, and S. Nahavandi, “An intelligent driven deep residual learning framework for brain tumor classification using MRI images,” *Expert Systems with Applications*, vol. 213, p. 119087, 2023.
[19] M. Rasool, N. A. Ismail, W. Boulila, A. Ammar, H. Samma, W. M. Yafooz, and A.-H. M. Emara, “A hybrid deep learning model for brain tumour classification,” *Entropy*, vol. 24, no. 6, p. 799, 2022.
[20] S. Shanthi, S. Saradha, J. Smitha, N. Prasath, and H. Anandakumar, “An efficient automatic brain tumor classification using optimized hybrid deep neural network,” *International Journal of Intelligent Networks*, vol. 3, pp. 188–196, 2022.