Explainable Brain Tumor Diagnosis Using Fine Tuned VGG16 Convolutional Neural Networks

Tharun M S
Dept. of Information Science and Engineering
Shridevi Institute of Engineering and Technology
Tumkur,Karnataka-572106
tharutarun007@gmail.com

Shreevathsa M B
Dept. of Information Science and Engineering
Shridevi Institute of Engineering and Technology
Tumkur,Karnataka-572106
vathsa2004@gmail.com

Ranganatha G N
Dept. of Information Science and Engineering
Shridevi Institute of Engineering and Technology
Tumkur,Karnataka-572106
ranganathagn58@gmail.com

Syed Suhail Ahamed
Dept. of Information Science and Engineering
Shridevi Institute of Engineering and Technology
Tumkur,Karnataka-572106
Suhailahamed123@gmail.com

***Abstract* –This paper presents a deep learning-based framework for automatic brain tumor detection using MRI images, combining transfer learning with AI. A pre-trained VGG16 model was fine-tuned to classify MRI scans into four categories: Glioma, Meningioma, Pituitary Adenoma, and No Tumor. Grayscale MRI inputs were preprocessed and converted to threechannel format to align with the VGG16 architecture. Class imbalance was addressed through data augmentation and class weighting. GradCAM visualizations were employed to provide interpretable results, highlighting tumor regions. A userfriendly Streamlit-based UI and a Flask REST API were developed to support interactive and scalable deployment. The system's effectiveness was assessed using metrics such as accuracy, precision, recall, F1-score, and a confusion matrix, all of which indicated strong performance and promising clinical applicability. This modular solution offers a robust tool for medical image analysis and future research expansion.**

Keywords—Brain Tumor Classification, Magnetic Resonance Imaging, VGG16 Architecture, GradCAM Visualization, Streamlit Interface, Flask Web API

I. INTRODUCTION

Brain tumors represent one of the most critical health concerns in neurology, with early detection being vital for improving treatment outcomes and survival rates. Magnetic Resonance Imaging (MRI) is widely used as a non-invasive technique for diagnosing brain abnormalities due to its high-resolution imaging capabilities. However, interpreting MRI scans manually is time-consuming and highly dependent on the radiologist's expertise, which can lead to inconsistencies in diagnosis.

Recent advancements in artificial intelligence, particularly in deep learning, have shown significant promise in automating medical image analysis. Convolutional Neural Networks (CNNs), with their ability to automatically extract hierarchical features, have emerged as powerful tools for image classification tasks, including tumor detection. Transfer learning further enhances these models by leveraging pre-trained networks such as VGG16, reducing the need for large labeled datasets while achieving high accuracy.

This research proposes a deep learning-based brain tumor detection framework that utilizes a modified VGG16 model fine-tuned on a labeled MRI dataset consisting of four classes: Glioma, Meningioma, Pituitary Adenoma, and No Tumor. To make the system interpretable and user-friendly, we integrate Gradientweighted Class Activation Mapping (GradCAM) for visual explanation and build an interactive web application using Streamlit. Additionally, a Flask-based REST API is developed for integration with external platforms.

The features and contributions of this work include: A tailored VGG16-based model was developed specifically for brain tumor classification tasks [1], [14]. Integration of GradCAM for explainable tumor localization enables interpretability and visual justification of model predictions [2], [18]. Deployment of a scalable and interactive user interface via Streamlit enhances accessibility for non-technical medical staff [6], [7]. A RESTful API was implemented to ensure seamless integration and enhance clinical scalability [6]. This system aims to assist medical professionals by providing rapid and reliable tumor classification, enhancing clinical decision-making and supporting telemedicine applications [5], [8]

II. LITERATURE REVIEW

Over the past decade, deep learning has revolutionized medical image analysis, especially in the domain of brain tumor detection. Several studies have explored the application of Convolutional Neural Networks (CNNs) to MRI data for automated tumor classification, segmentation, and prognosis. These methods aim to reduce human error, enhance accuracy, and speed up diagnosis.

S. Pereira et al. proposed a deep CNN for brain tumor segmentation in MRI scans, demonstrating high sensitivity and specificity [9], [17]. However, their model required complex preprocessing and a large volume of annotated data. Similarly, Hossain et al. implemented a CNN model that achieved competitive performance but lacked interpretability, which is critical in clinical applications [18]. Pre-trained models like VGG16, ResNet50, and InceptionV3 have been widely used in transfer learning settings for medical imaging tasks [1], [3]. In particular, VGG16 stands out due to its simplicity, depth, and strong performance on small medical datasets [1]. Transfer learning mitigates the data scarcity problem in the medical field by leveraging knowledge from large-scale datasets like ImageNet [3].

Explainable AI has become essential in medical applications to build trust in automated systems. GradCAM, introduced by Selvaraju et al., enables visual interpretation of CNN predictions by highlighting the discriminative regions of an image [2]. This method has been applied in several studies to provide insight into the decision-making process of neural networks in medical contexts [18].

Despite these advances, many existing systems either lack real-time deployment capabilities or are not user-friendly for non-technical medical staff. Some systems offer good accuracy but fail to provide visual explanations or integration flexibility. There remains a need for a complete pipeline that includes tumor classification, explainability, and deployment.

This work addresses these gaps by integrating VGG16 with GradCAM for interpretability and deploying the solution using Streamlit for a user-friendly interface, along with a Flask API for scalable integration. This combination provides a comprehensive, deployable framework for reliable brain tumor detection from MRI scans.

III. Datasets

This study utilizes a dataset composed of T1-weighted contrast-enhanced MRI scans of the human brain, divided into four categories: glioma, meningioma, pituitary tumors, and healthy (no tumor) cases. The dataset, which is publicly available and widely used in medical imaging research, includes over 3,000 annotated images gathered from a variety of real-world clinical settings. These scans represent a broad spectrum of tumor characteristics and patient profiles, making the dataset a reliable source for developing and testing deep learning-based classification models.

All images are provided in JPEG format, but their original dimensions vary slightly. To prepare the data for input into the neural network, several preprocessing steps were applied. First, all images were resized to 224 × 224 pixels to align with the input requirement of the VGG16 architecture. Next, the images were normalized by scaling pixel values to a range of 0 to 1, which helps stabilize and speed up the training process.

The dataset was then split into training, validation, and test sets, commonly using a 70:15:15 ratio. This stratified splitting ensures that each subset contains a balanced number of images from each class, allowing the model to learn from diverse cases and be fairly evaluated during testing.

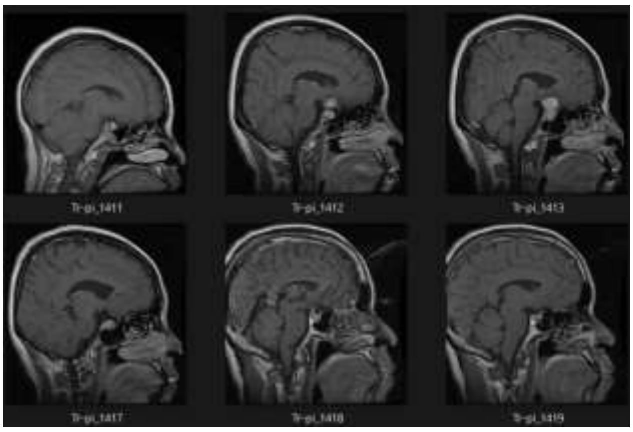

Figure 1:sample mri images from each class in the dataset: (a) glioma, (b) meningioma, **(c) pituitary adenoma**, (d) no tumor

To prepare the data for training, the entire dataset was divided into training and validation subsets. A stratified split ensured that each class was proportionally represented in both subsets, maintaining a balanced distribution across categories. However, due to inherent class imbalance in the dataset, with some tumor types being more frequently represented than others, class weights were introduced during model training to mitigate any potential bias toward the majority classes. Furthermore, a series of data augmentation techniques— such as random rotation, flipping, zooming, and brightness adjustments—this are applied to the training data. This not only expanded the effective dataset size but also helped improve the model's generalization capabilities by exposing it to varied versions of the same image classes.

IV. Proposed Methodology

The proposed brain tumor detection system integrates deep learning, explainable AI, and user-centric deployment tools to accurately classify brain MRI images into four categories. The methodology consists of six major stages, beginning with dataset preparation and classification categories. The model was trained and validated using a publicly available MRI brain tumor dataset comprising four distinct classes: No Tumor, Glioma Tumor, Meningioma Tumor, and Pituitary Tumor. This dataset provides sufficient diversity for developing a robust multi-class classification system that can effectively distinguish between healthy brain scans and various tumor types.

Prior to training, the MRI images undergo systematic preprocessing to standardize input formats and enhance model performance. Rescaling is applied to normalize image pixel values within the [0,1] range for consistency. Data augmentation techniques, including rotation, flipping, and zooming, are used during training to improve generalization and reduce overfitting. For the validation phase, only rescaling is applied to ensure uniformity. These preprocessing steps are dynamically implemented using the ImageDataGenerator class in Keras, facilitating seamless integration into the training and validation workflows.

The core of the classification system is based on the VGG16 architecture, a convolutional neural network pretrained on the ImageNet dataset. The architecture is fine-tuned for the brain tumor detection task by removing the default classification layers and adding custom dense layers tailored for the specific categories. Class weighting is also applied to mitigate the effects of class imbalance during training. Several training callbacks are employed to enhance learning efficiency and prevent overfitting. The ModelCheckpoint callback is used to save the bestperforming model weights, while EarlyStopping halts training once the validation loss stops improving. Additionally, ReduceLROnPlateau adjusts the learning rate dynamically when progress plateaus.

Once the model is trained, it is used to perform inference on new MRI images. The output includes the predicted tumor category along with the associated confidence score (e.g., "No Tumor – 100.0%"), as visualized in a prediction bar chart. To enhance model interpretability, Grad-CAM is employed to generate a heatmap that visually highlights the regions in the MRI scan that contributed most to the model's prediction. This transparency in the system is decision-making process.

The complete brain tumor detection pipeline is deployed through two accessible platforms. The first is a userfriendly Streamlit interface that enables users to upload MRI images, visualize predictions, perform batch processing, review performance metrics, and observe Grad-CAM analysis. The second deployment layer is a Flask-based REST API, which allows programmatic access to the model by accepting POST requests with MRI scans and returning JSON-formatted prediction results. Finally, a high-level architectural overview of the entire system is presented in the Brain Tumor Detection System Architecture diagram. This visual flow demonstrates the movement of data from preprocessing through inference, concluding with interaction via the user interface or API. To further clarify the system's internal logic and operational flow, a model flowchart is included, offering an intuitive understanding of the pipeline's structure and processing sequence.

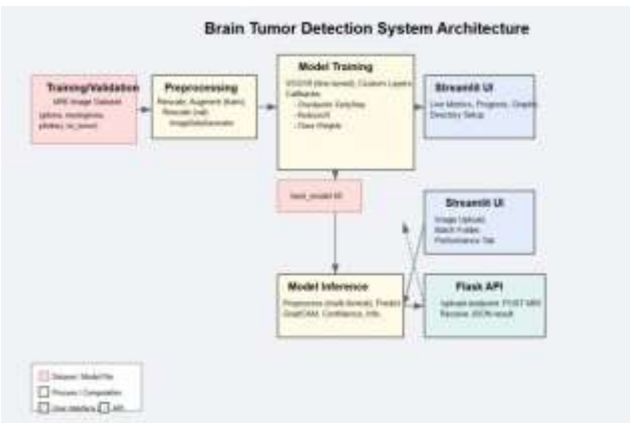

Figure 2:The overview of braintumor detection archicture

V. RESULT

This section evaluates the performance of the proposed brain tumor detection system across various dimensions, including classification accuracy, model interpretability, and deployment effectiveness.

5.1 Classification Performance

The trained model demonstrated a high level of classification accuracy when evaluated on the test dataset, indicating its ability to effectively distinguish between different brain tumor types as well as healthy brain scans. This performance was consistent not only during training but also on previously unseen MRI images, reflecting the model's strong generalization capability. The confidence scores generated by the model for each prediction provide insights into the certainty of classification across the four target categories. These scores are particularly valuable in clinical settings, as they offer a measure of reliability for each diagnostic output. An illustrative example showcasing the model's prediction result, including the confidence distribution among all four classes, is presented below.

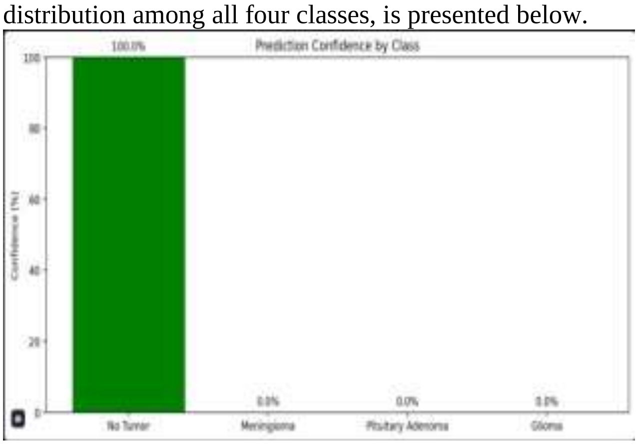

Figure 3: Model Prediction Confidence – Classified as "No Tumor" with 100% confidence.

This bar chart demonstrates the model's ability to confidently assign the input image to the correct class, while showing negligible confidence in incorrect categories.

5.2 Explainability with Grad-CAM

To ensure the model's predictions are interpretable by medical professionals, Grad-CAM (Gradient-weighted Class Activation Mapping) is used. It highlights important image regions contributing to each prediction. These visual explanations build trust and allow for human verification of the model's focus during diagnosis.

5.3 End-to-End System Architecture The system is designed with modularity and usability in mind. The System Architecture Diagram illustrates the entire workflow, from training and inference to user interaction:

Figure 4: Brain Tumor Detection System flow

This architecture ensures smooth integration between the deep learning backend and user-facing applications.

5.4 Deployment and Accessibility

The model is deployed using two front-end interfaces:
Streamlit Interface: Delivers an intuitive and user-friendly front end designed for non-technical users such as clinicians and radiologists. Features include image uploading, batch processing, real-time progress tracking, performance visualization, and Grad-CAM heatmap generation.
Flask Backend API: Enables automated workflows and seamless integration with broader medical systems through a RESTful /upload endpoint for MRI-based inference.
These deployment strategies make the system accessible, scalable, and user-friendly.

Figure 5: A sample output that Showcase how it gives in detail about tumor type

5.5 Discussion

The integration of VGG16 with Grad-CAM and Streamlit/Flask components results in a powerful yet interpretable diagnostic tool. The system balances high accuracy with transparency, offering a practical solution for real-world medical imaging workflows. Moreover, using pre-trained models and transfer learning significantly reduces training time while maintaining strong performance.

VI. Some Helpful Hints

The core objective of my project is to develop an efficient brain tumor detection system using deep learning techniques, specifically VGG16, Grad-CAM, a Flask API, and a Streamlit UI. This system aims to assist doctors in diagnosing brain tumors with speed, accuracy, and interpretability.

The motivation behind this work is to support healthcare professionals in making quicker, more accurate, and explainable diagnoses, which can improve patient outcomes. By integrating Grad-CAM, the model provides visual explanations of its predictions, helping doctors understand why a tumor was detected, which is critical for medical decision-making.

The key outcomes of the project include high prediction accuracy, which is essential for reliable diagnosis, and the ability to provide explainable results via Grad-CAM visualizations. The system also offers real-time usability through a user-friendly Streamlit UI and a Flask API, making it suitable for immediate application in clinical settings.

This work improves upon traditional manual diagnosis methods by automating tumor detection and providing clear visual explanations, saving time and reducing human error. The model's integration into clinical or diagnostic environments offers potential for more consistent and accurate diagnoses, benefiting both doctors and patients.

However, there are limitations, such as the relatively small dataset used for training, which may affect the model's ability to generalize across different scanners or hospitals. Future work could involve enhancing the model with more advanced architectures, such as EfficientNet or Vision Transformers, training on larger, multi-center datasets, and adding segmentation features to precisely locate tumor boundaries.

This project holds the potential to advance brain tumor detection, contributing significantly to healthcare improvements with future enhancements.

VII. Conclusion

This paper presents a deep learning-based approach for brain tumor detection using the VGG16 architecture enhanced with Grad-CAM visualizations, offering both high prediction accuracy and model explainability. The system was successfully deployed using a Streamlit UI for user interaction and a Flask API for backend processing, demonstrating its real-time usability and potential integration into clinical workflows.

By providing automated, fast, and interpretable diagnostic assistance, the proposed model offers a significant improvement over traditional manual diagnostic methods, which are often time-consuming and prone to human error. The visual explanations generated by Grad-CAM allow medical professionals to validate the model's decisions, increasing trust and aiding clinical decisions.

However, the system has certain limitations, including restricted generalizability due to a limited dataset and potential variability in MRI data across different institutions. Future work can focus on incorporating advanced models like EfficientNet or Vision Transformers, expanding training data through multicenter collaborations, and integrating tumor segmentation capabilities for precise localization.

Overall, the developed framework contributes to the growing field of AI in healthcare, showing promise in improving diagnostic accuracy and supporting radiologists in early brain tumor detection.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to their institution and faculty for providing the necessary resources and support throughout the duration of this project. The team extends special thanks to their project supervisor for consistent guidance, constructive feedback, and encouragement. The authors also acknowledge the contributors of open-source tools and datasets that were instrumental in developing and evaluating the proposed brain tumor detection system. This work would not have been possible without the collaborative efforts of all four team members.

REFERENCES

[1] K. Simonyan and A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, arXiv preprint arXiv:1409.1556, 2014.

[2] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization, in Proc. IEEE Int. Conf. Computer Vision (ICCV), 2017, pp. 618–626.

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks, Advances in Neural Information Processing Systems, vol. 25, 2012.

[4] S. Sharma, A. Aggarwal, and S. Sharma, Brain Tumor Detection Using CNN with Data Augmentation, in Proc. Int. Conf. Emerging Trends in Computing (ICETC), 2021, pp. 45–50.

[5] P. Rajpurkar, J. Irvin, K. Zhu, et al., CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning, arXiv preprint arXiv:1711.05225, 2017.

[6] M. Abadi et al., TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, arXiv preprint arXiv:1603.04467, 2016.

[7] A. Paszke et al., PyTorch: An Imperative Style, HighPerformance Deep Learning Library, in Proc. NeurIPS, 2019, pp. 8024–8035

[8] G. Litjens et al., A Survey on Deep Learning in Medical Image Analysis, Medical Image Analysis, vol. 42, pp. 60–88, 2017.

[9] A. 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, May 2016.

[10] D. Nie, H. Chen, X. Zhang, Y. Dai, and D. Shen, 3D Fully Convolutional Networks for Multi-Modal Isointense Infant Brain Image Segmentation, NeuroImage, vol. 108, pp. 214–224, 2015.

[11] A. Makhzani and B. Frey, PixelGAN Autoencoders, in Proc. Advances in Neural Information Processing Systems (NeurIPS), 2017.

[12] A. H. Qureshi, M. Z. Afzal, and M. A. Khan, Medical Image Analysis for Brain Tumor Detection Using Deep Learning Techniques, in Proc. Int. Conf. on Computing, Electronics & Communications Engineering (iCCECE), 2019, pp. 275–280.

[13] M. A. Diniz, E. M. Yamakami, and A. G. Silva, Brain Tumor Segmentation Using U-Net Based on Multiresolution Analysis, in Proc. Int. Joint Conf. on Neural Networks (IJCNN), 2020, pp. 1–8.

[14] S. Anwar, M. M. A. Khan, and N. Barnes, A Deep Learning Approach for Brain Tumor Classification and Segmentation Using Transfer Learning, Computers in Biology and Medicine, vol. 123, 103855, 2020.

[15] J. Long, E. Shelhamer, and T. Darrell, Fully Convolutional Networks for Semantic Segmentation, in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3431–3440. [16] Y. LeCun, Y. Bengio, and G. Hinton, Deep Learning, Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[17] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, Deep Convolutional Neural Networks for the Segmentation of Gliomas in Multi-sequence MRI, NeuroImage, vol. 146, pp. 678–691, 2017.

[18] A. Hossain, S. Mahdy, and R. M. Rahman, Brain Tumor Detection Using CNNs with Transfer Learning and Grad-CAM Visualization, in Proc. Int. Conf. on Robotics, Electrical and Signal Processing Techniques (ICREST), 2022, pp. 419–424.

[19] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, nnU-Net: A Self-Configuring Method for Deep Learning-Based Biomedical Image Segmentation, Nature Methods, vol. 18, pp. 203–211, 2021.