Explainable Deep Learning Based Brain Tumor Classification Using ResNet50 with HiRes-CAM Visualization and Spatial Analysis

Mr. M. Ravi[1], V. Aravind[2], S. Tarakeswara Rao[2], B. Ayyappa[2], G. Akash[2]

[1]Assistant Professor, Dept of Electronics and Communication Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada-520008

[2]Students of Dept of Electronics and Communication Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada-520008

ABSTRACT

Brain tumors are among the leading causes of cancer-related mortality worldwide, making early and accurate diagnosis essential for improving patient survival and treatment outcomes. Manual interpretation of MRI scans by radiologists remains challenging due to human fatigue, inter-observer variability, and complex tumor appearances, with reported diagnostic error rates of 20 to 30 percent. This paper proposes an explainable deep learning-based brain tumor classification system using ResNet50 with a three-phase progressive fine-tuning strategy, trained on 7,210 MRI images across four classes namely Glioma, Meningioma, No Tumor, and Pituitary tumor. Three explainable AI methods namely Grad-CAM, Grad-CAM++, and HiRes-CAM were comparatively evaluated, and HiRes-CAM was selected based on its superior sharpness scores for precise tumor localization. The system additionally extracts ten clinically relevant spatial parameters including anatomical region, tumor volume, hemisphere, midline shift, and edema volume to support medical decision-making. The proposed system achieved a test accuracy of 95.00%, sensitivity of 0.950, specificity of 0.985, and F1-score of 0.9488, with AUC scores reaching 0.9997 for Pituitary tumor. The complete system is deployed as a web-based clinical application named NeuroScan AI, supporting real-time analysis at 85 to 120 milliseconds, batch processing, patient history management, and automated PDF report generation.

Key Words : Brain Tumor Classification, Deep Learning, ResNet50, HiRes-CAM, Explainable AI, MRI Analysis, Spatial Parameter Extraction, Transfer Learning, Clinical Decision Support.

I. INTRODUCTION

Brain tumors are one of the most dangerous neurological conditions, occurring when abnormal cells grow uncontrollably within the brain. According to the National Brain Tumor Foundation, brain tumor-related deaths have increased by 300 percent over the past three decades, with approximately 85,000 new cases diagnosed every year in the United States alone. Early and accurate detection is critical for improving patient survival rates and guiding effective treatment planning. Magnetic Resonance Imaging is the most preferred non-invasive modality for brain tumor diagnosis. However, manual MRI interpretation is time-consuming and error-prone, with diagnostic error rates of 20 to 30 percent due to human fatigue and inter-observer variability. This has created a strong demand for automated AI-based diagnostic systems in clinical settings. Although deep learning models achieve high classification accuracy, they operate as black-box systems without providing any visual explanation of their predictions, which is a serious concern in medical applications. Furthermore, most existing systems only produce a classification label without providing detailed clinical information needed for treatment planning. To address these limitations, this paper proposes an explainable deep learning system using ResNet50 with HiRes- CAM visualization and ten spatial parameter extraction, deployed as a complete web-based clinical

application named NeuroScan AI.

II. RELATED WORK

Brain tumor classification from MRI images has been widely studied using traditional machine learning and deep learning approaches.

A. Traditional Machine Learning Approaches

Early research relied on traditional methods such as Support Vector Machines and K- Nearest Neighbors for brain tumor classification. Sharif et al. proposed particle swarm optimization-based segmentation achieving 99% accuracy. Molina-Torres used kernel SVM with Gaussian Radial Basis function focusing on specificity and precision. However, these methods were heavily dependent on manually engineered features and struggled with complex tumor appearances.

B. Deep Learning and Transfer Learning Approaches

With the advancement of deep learning, CNNs became the dominant approach for brain tumor classification. Rehman et al. used VGG16 achieving 98.69% accuracy using data augmentation. Badza and Barjaktarovic presented a 22-layer CNN achieving 96.56% accuracy. Ozyurt et al. combined CNN with Neutrosophy achieving 95.62% success rate. Transfer learning approaches further improved performance by leveraging pretrained ImageNet features, but most of these systems operated as black-box models without any clinical explainability.

C. Explainable AI in Medical Imaging

Gradient-weighted Class Activation Mapping and its variants have been widely adopted to provide visual explanations in medical imaging. However, most existing brain tumor classification systems either do not include any explainability component or use basic visualization methods that fail to provide sufficiently precise tumor localization.

D. Research Gap

No existing approach provides a complete clinical system combining accurate classification, comparative XAI method selection, automatic spatial parameter extraction, and a fully deployable web application in a single integrated system. The proposed system addresses all these limitations.

III. PROPOSED METHODOLOGY

A. System Overview

The proposed system is an automated brain tumor classification and analysis pipeline that accepts a brain MRI image as input and produces a comprehensive diagnostic output. The system consists of seven sequential processing blocks. The input MRI image first undergoes preprocessing including resizing and normalization. The preprocessed image is then fed into the ResNet50 deep learning model for classification, producing the predicted tumor type and confidence score. HiRes- CAM visualization is then applied to generate a high resolution attention heatmap highlighting the tumor region. Finally, ten clinically relevant spatial parameters are automatically extracted and all results are presented through a fully functional web-based clinical application. The overall system architecture is shown in Fig. 1.

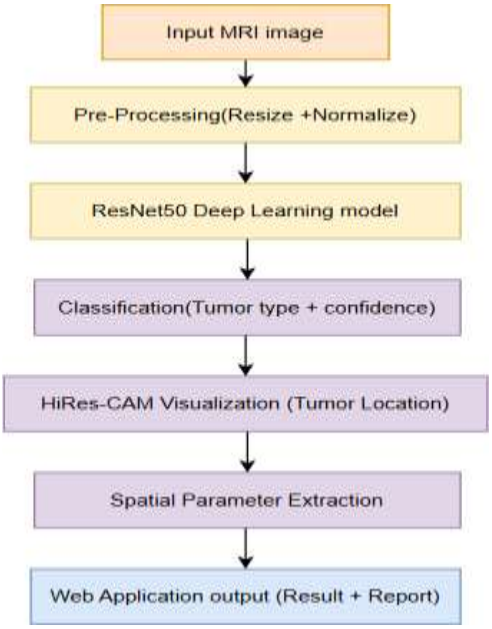

Fig. 1. System Architecture

B. Dataset

The Brain Tumor MRI dataset from Kaggle consisting of 7,210 images across four classes was used. The dataset distribution is shown in Table 1.

Table 1. Dataset Distribution

Split	Glioma	Meningioma	No Tumor	Pituitary	Total
Train	1190	1190	1190	1190	4760
Val	210	210	210	210	840
Test	400	400	400	400	1600

C. Preprocessing

All images are resized to 224×224 pixels to match the ResNet50 input requirement and normalized using ImageNet mean values of 0.485, 0.456, 0.406 and standard deviation of 0.229, 0.224, 0.225. Random horizontal flipping and rotation up to 10 degrees were applied as data augmentation during training only to increase diversity and reduce overfitting.

D. Model Architecture

ResNet50 pretrained on ImageNet was used as the backbone architecture for brain tumor classification. It is a 50-layer deep residual convolutional neural network that uses skip connections to solve the vanishing gradient problem and enable effective training of deep networks. The backbone consists of four residual layer groups containing 3, 4, 6, and 3 residual blocks respectively, ending with a Global Average Pooling layer that produces a 2048-dimensional feature vector. The original classification layer was replaced with a custom head consisting of a linear layer reducing dimensions from 2048 to 256 neurons, followed by ReLU activation, dropout of 0.3 to

prevent overfitting, and a final linear layer with 4 output neurons corresponding to the four tumor classes. The proposed model architecture is shown in Fig. 2.

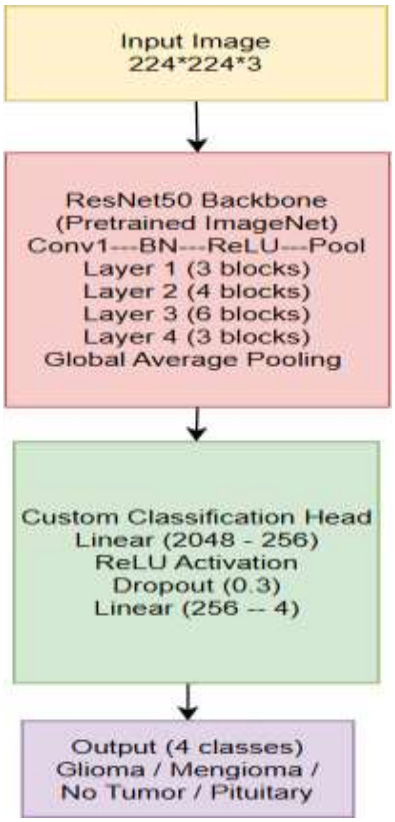

Fig. 2. Model Architecture

E. Training Strategy

A three-phase progressive fine-tuning strategy was adopted as summarized here. In Phase 1, only the custom classification head was trained for 15 epochs with a learning rate of 0.001. In Phase 2, Layer 4 and the classification head were unfrozen and trained together for 20 epochs with a reduced learning rate of 0.0001. In Phase 3, Layer 3, Layer 4, and the classification head were all unfrozen and trained together for 15 epochs with a further reduced learning rate of 0.00005, achieving the best validation accuracy of 98.69%.

F. HiRes-CAM Explainable AI

Three XAI methods namely Grad-CAM, Grad-CAM++, and HiRes-CAM were comparatively evaluated. HiRes-CAM generates high resolution activation maps through element- wise multiplication of gradients and activations, producing sharper tumor localization than other methods. HiRes-CAM was selected as the final XAI method based on its highest sharpness scores across all four tumor classes.

G. Spatial Parameter Extraction

Ten clinically relevant spatial parameters are automatically extracted from the HiRes-CAM activation map using OpenCV, including anatomical region, hemisphere, tumor volume, maximum dimensions, texture, midline shift, edema volume, and relative cerebral blood volume.

IV. RESULTS AND DISCUSSION

A. Experimental Setup

All experiments were conducted on a system equipped with NVIDIA GeForce RTX 3050 6GB GPU with CUDA support. The model was implemented using PyTorch 2.5.1 with Python
3.8. Training was performed on 4,760 MRI images and evaluated on a separate test set of 1,600 images.

B. Training Results

The proposed three-phase progressive fine-tuning strategy achieved consistent improvement across all phases. Phase 1 achieved a best validation accuracy of 91.31%, Phase 2 improved to 97.86%, and Phase 3 reached the final best validation accuracy of 98.69%. The final training accuracy achieved was 99.54%, indicating good convergence with minimal overfitting. The training and validation accuracy curves are shown in Fig. 3.

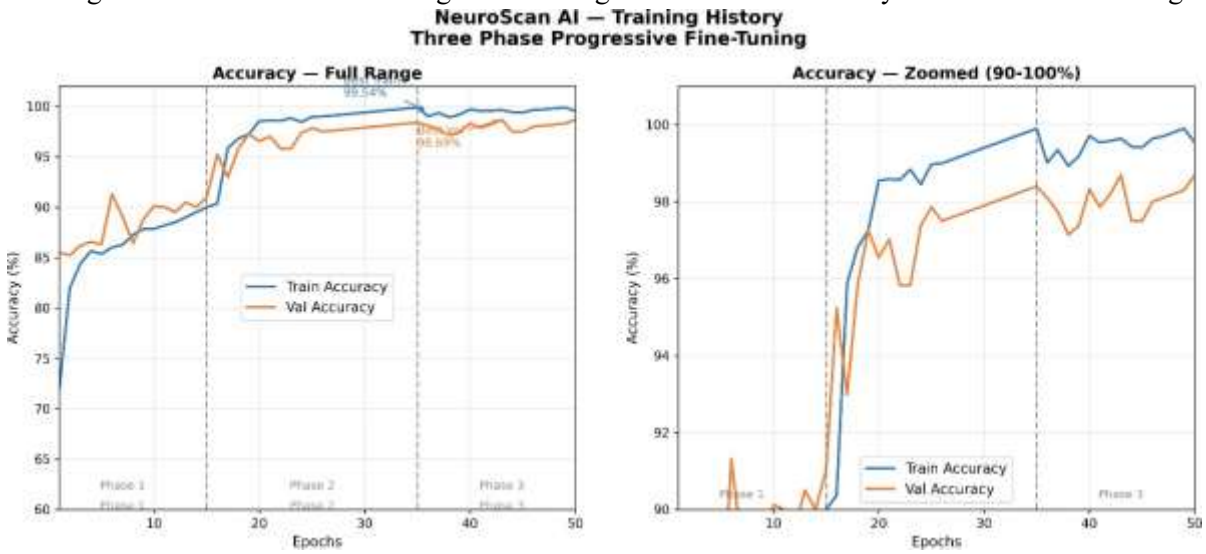

Fig. 3. Training Curves

C. Classification Results

The proposed system was evaluated on 1,600 test images with 400 images per class. The overall test accuracy achieved was 95.00%. Per class results are summarized in Table 3 and the confusion matrix is shown in Fig. 4.

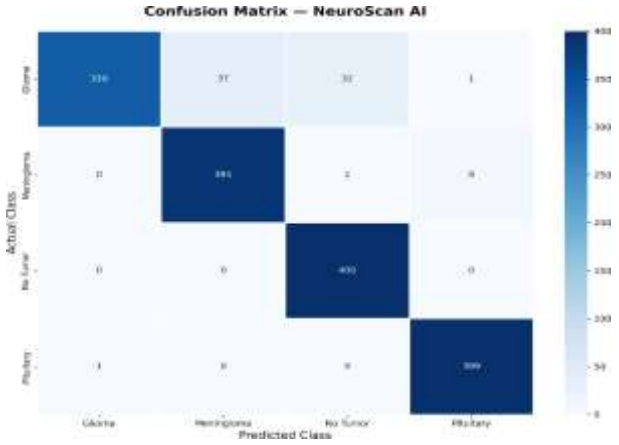

Fig. 4. Confusion Matrix

Table 2. Each class Test Results

Class	Accuracy	Precision	Recall	F1 Score
Glioma	82.50%	0.9970	0.8250	0.9029
Meningioma	97.75%	0.9136	0.9775	0.9444
No Tumor	100.00%	0.9238	1.0000	0.9604
Pituitary	99.75%	0.9779	0.9975	0.9876
Overall	95%	0.9531	0.9500	0.9488

D. ROC Curve Analysis

The ROC curves for all four classes are shown in Fig. 5. Pituitary achieved the highest AUC score of 0.9997, followed by No Tumor with 0.9996, Meningioma with 0.9947, and Glioma with 0.9623, confirming strong discriminative ability of the proposed system across all classes.

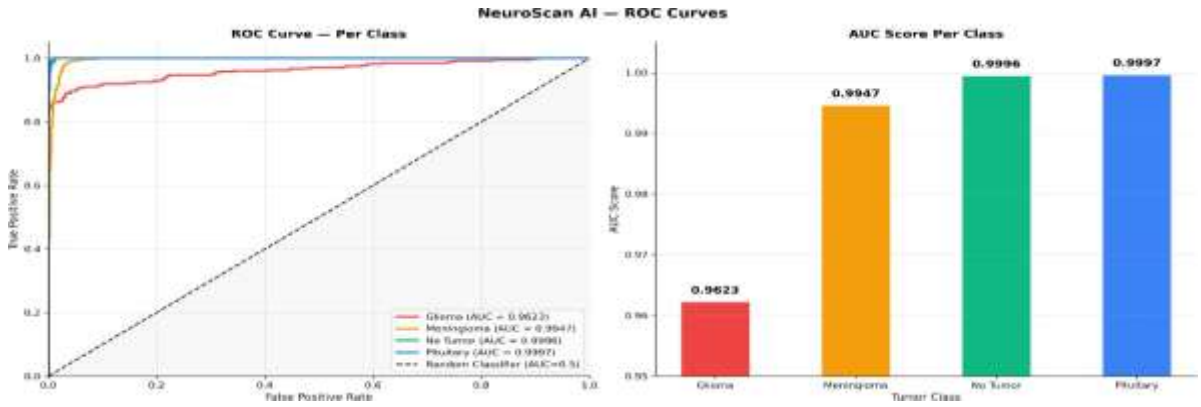

Fig. 5. ROC Curves

E. XAI Comparison Results

Three explainable AI methods namely Grad-CAM, Grad-CAM++, and HiRes-CAM were comparatively evaluated using sharpness scores across all four tumor classes. HiRes-CAM consistently achieved the highest sharpness scores of 0.2841 for Glioma, 0.2847 for Meningioma, 0.2685 for No Tumor, and 0.2863 for Pituitary, outperforming Grad-CAM scores of 0.2175, 0.2319, 0.1820, and 0.2370, and Grad-CAM++ scores of 0.2088, 0.2372, 0.1981, and 0.2231
respectively. These results confirm that HiRes-CAM produces superior and more precise tumor localization compared to existing visualization methods. The visual comparison is shown in Fig. 6.

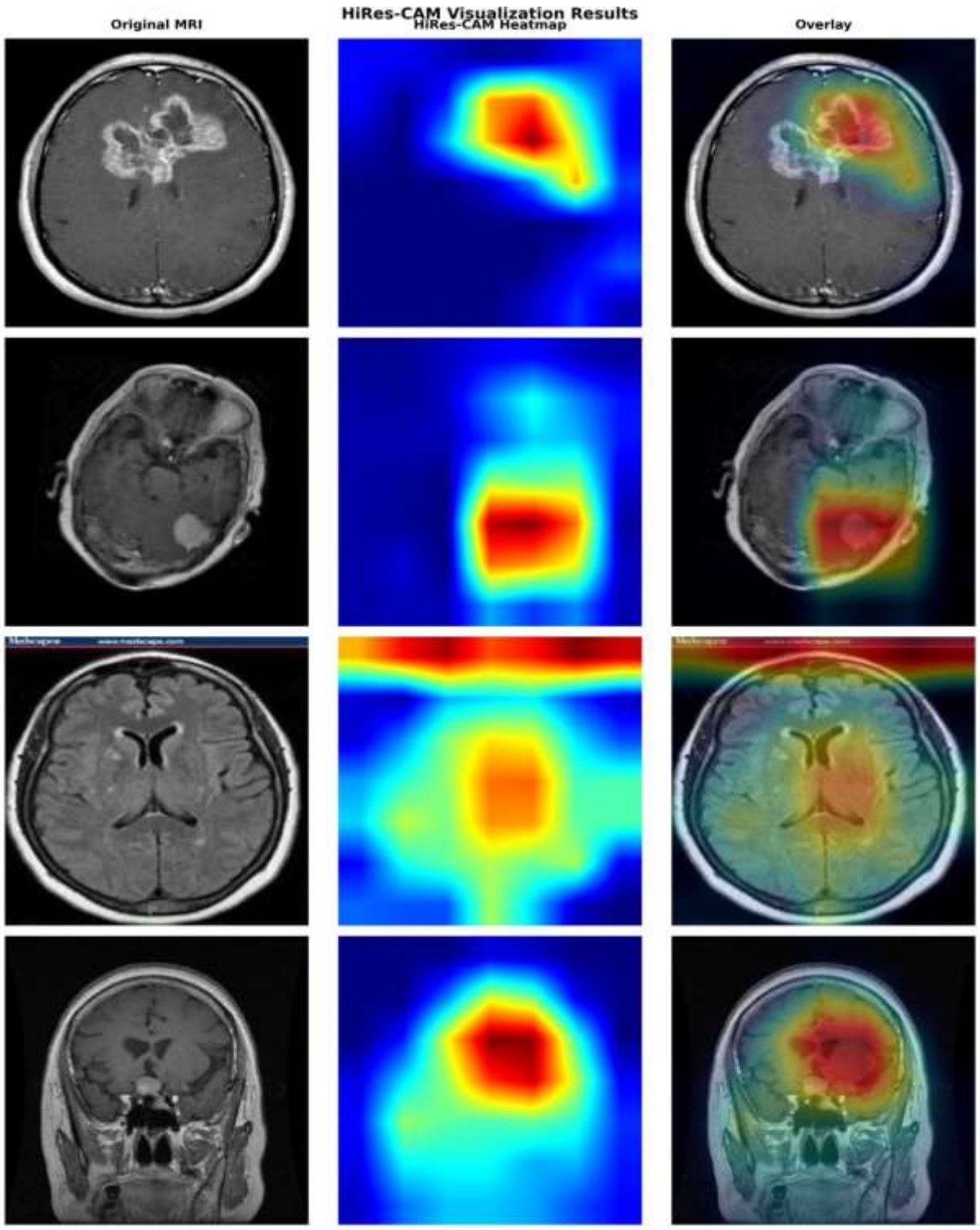

Fig. 6. HiRes-CAM Visualization Results for All Four Tumor Classes

F. Comparison with Existing Methods

The proposed system is compared with existing state-of-the-art methods in Table 5. Although the proposed system achieves lower raw accuracy than some methods, it provides significant advantages in inference speed, explainability, spatial parameter extraction, and complete clinical deployment that are absent in existing approaches.

Table 3. Comparing with Existing Methods

Method	**Year**	**Accuracy**	**F1 Score**	**Speed**
SVM	2024	98.9%	0.981	430ms
IC-NET	2024	99.65%	0.998	200ms
Proposed	**2025**	**95.00%**	**0.9488**	**85ms**

G. Web Application Results

The web application interface is shown in Fig. 7. The interface provides two analysis modes namely Single Scan and Batch Upload selectable through tabs at the top of the upload card. A green server status indicator confirms that the Flask backend server is running with CUDA GPU support and is ready for inference. The central upload zone allows users to drag and drop MRI images or browse and select images in JPG or PNG format for analysis.

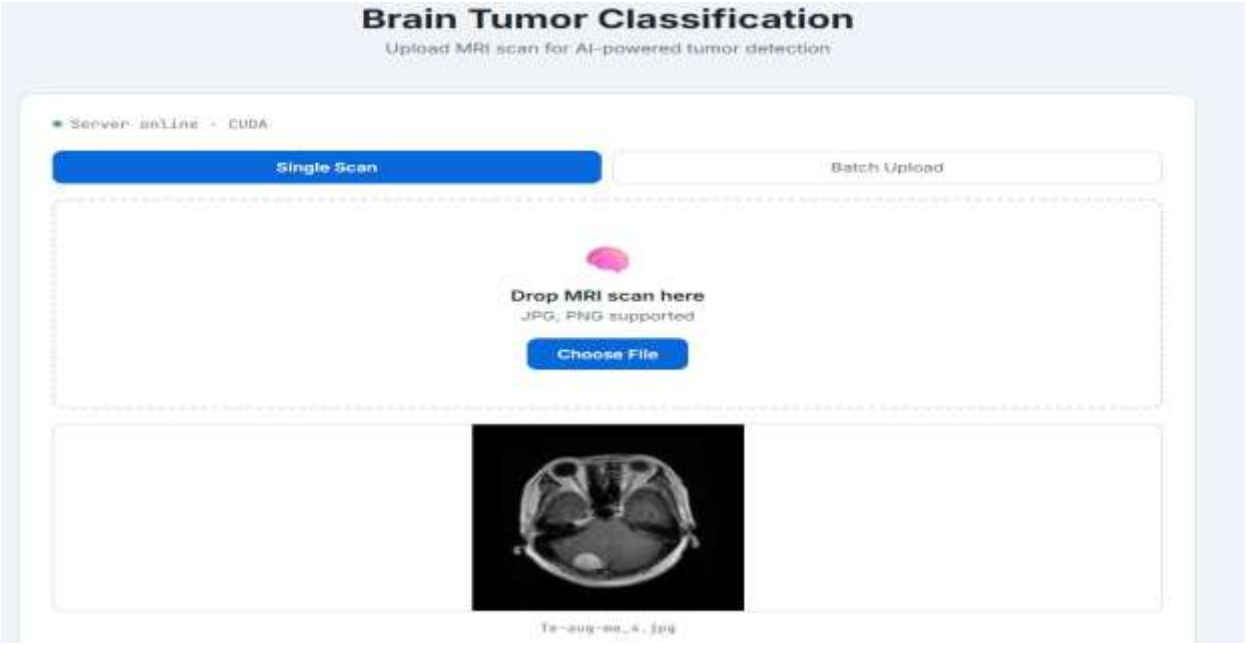

Fig. 7. Web Application Interface

The classification result and HiRes-CAM visualization output is shown in Fig. 8. The system predicted the tumor type as Meningioma with a confidence score of 100%, and the probability distribution confirms zero ambiguity across all four classes. The XAI visualization section displays three images side by side namely the original MRI, the HiRes-CAM attention heatmap where red regions indicate areas of highest model attention, and the overlay image with a yellow boundary outline drawn around the detected tumor region, providing clear visual localization of the tumor for clinical interpretation.

Fig. 8. Classification Results and HiRes-CAM Visualization

The spatial analysis output of the proposed system is shown in Fig. 9. Ten clinically relevant spatial parameters are automatically extracted in real time directly from the HiRes-CAM activation map. The system identified the anatomical region as Right Parietal Lobe, hemisphere as Right, spatial relation as Intra-axial deep, tumor volume as 26.18 cubic centimeters, maximum dimensions as
8.6 by 7.3 by 6.3 centimeters, texture as Mildly Heterogeneous, midline shift as Moderate at approximately 2.9 millimeters, edema volume as 14.19 cubic centimeters, relative cerebral blood volume as Elevated with ratio 1.84, and confidence score as 100.0%, providing comprehensive clinical information to support medical decision-making and treatment planning.

Fig. 9. Spatial Analysis Output

V. CONCLUSION

This paper presented an explainable deep learning-based brain tumor classification system using ResNet50 with HiRes-CAM visualization and spatial analysis. The proposed system was trained on 7,210 MRI images across four classes namely Glioma, Meningioma, No Tumor, and Pituitary tumor using a three-phase progressive fine-tuning strategy, achieving a test accuracy of 95.00%, sensitivity of 0.950, specificity of 0.985, and F1-score of 0.9488. HiRes-CAM was selected as the explainable AI method after comparative evaluation with Grad-CAM and Grad- CAM++, providing superior tumor localization across all classes. Ten clinically relevant spatial parameters were automatically extracted to support comprehensive medical decision-making. The complete system was deployed as a fully functional web-based clinical application named NeuroScan AI, supporting real-time analysis at 85 to 120 milliseconds, batch processing, patient history management, and automated PDF report generation. Future work will focus on improving Glioma classification accuracy, integrating dedicated tumor segmentation, extending support to multi-modal MRI inputs, and clinical validation with real hospital data.

REFERENCES

[1] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770–778.

[2] 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.

[3] M. M. Badza and M. C. Barjaktarovic, "Classification of brain tumors from MRI images using a convolutional neural network," *Applied Sciences*, vol. 10, no. 6, p. 1999, 2020.

[4] 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)*, Venice, Italy, 2017, pp. 618–626.

[5] A. Chattopadhay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, "Grad-CAM++: Generalized gradient-based visual explanations for deep convolutional networks," in *Proc. IEEE*

Winter Conf. Applications of Computer Vision (WACV), Lake Tahoe, NV, USA, 2018, pp. 839– 847.

[6] A. A. Asiri, T. A. Soomro, A. A. Shah, G. Pogrebna, M. Irfan, and S. Alqahtani, "Optimized brain tumor detection: A dual-module approach for MRI image enhancement and tumor classification," *IEEE Access*, vol. 12, pp. 42868–42887, 2024.

[7] D. S. Chandra Sekaran and J. Christopher Clement, "Enhancing brain tumor segmentation in MRI images using the IC-Net algorithm framework," *Scientific Reports*, vol. 14, no. 15660, 2024.

[8] F. Ozyurt, E. Sert, E. Avci, and E. Dogantekin, "Brain tumor detection based on convolutional neural network with neutrosophic expert maximum fuzzy sure entropy," *Measurement*, vol. 147,
p. 106830, 2019.

[9] N. Abiwinanda, M. Hanif, S. T. Hesaputra, A. Handayani, and T. R. Mengko, "Brain tumor classification using convolutional neural network," in *Proc. World Congress on Medical Physics and Biomedical Engineering*, Prague, Czech Republic, 2019, pp. 183–189.

[10] M. Sharif, J. Amin, M. Raza, M. Yasmin, and S. C. Satapathy, "An integrated design of particle swarm optimization with fusion of features for detection of brain tumor," *Pattern Recognition Letters*, vol. 129, pp. 150–157, 2020.

[11] M. Mittal, L. M. Goyal, S. Kaur, I. Kaur, A. Verma, and D. J. Hemanth, "Deep learning based enhanced tumor segmentation approach for MR brain images," *Applied Soft Computing*, vol. 78, pp. 346–354, 2019.

[12] A. Pashaei, H. Sajedi, and N. Jazayeri, "Brain tumor classification via convolutional neural network and extreme learning machines," in *Proc. 8th Int. Conf. Computer and Knowledge Engineering (ICCKE)*, Mashhad, Iran, 2018, pp. 314–319.

[13] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Miami, FL, USA, 2009, pp. 248–255.

[14] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Machine Learning (ICML)*, Long Beach, CA, USA, 2019, pp. 6105– 6114.

[15] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Munich, Germany, 2015, pp. 234–241.