



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: <https://doi.org/10.22214/ijraset.2026.79816>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Multi-Class Brain Tumor Classification from MRI Using ResNet50V2 Transfer Learning with Two-Stage Augmentation, Grad-CAM Explainability, and Cross-Dataset Validation

Shalma Muji M¹, Nithin T², Lavanya D³, Abhiram VS⁴, S Sreeja⁵, A Kousalya⁶

^{1,2,3,4}B.Tech Student, ⁵M.E Student, ⁶Associate Professor Department of Artificial Intelligence and Data Science United Institute of Technology, Coimbatore, Tamil Nadu, India

Abstract: Brain tumor classification from magnetic resonance imaging (MRI) remains a challenging clinical task due to inter-class visual similarity and significant variation across imaging protocols and acquisition equipment. Accurate and timely classification of glioma, meningioma, pituitary tumor, and normal brain tissue is critical for treatment planning and patient outcomes. This work proposes a two-phase transfer learning framework based on ResNet50V2 pretrained on ImageNet for four-class brain MRI classification. A two-stage augmentation pipeline combining offline degradation augmentation (Gaussian noise (SNR=18.39dB), Gaussian blur (SSIM \geq 0.93), and resolution degradation (SSIM \geq 0.85))—with online geometric augmentation was employed to improve cross-scanner robustness. Class imbalance was addressed through balanced class weighting. Grad-CAM visualizations were generated to provide anatomically interpretable explanations of model decisions. Evaluated on a held-out test set of 310 images from the Daneshmand dataset, the proposed model achieved 89.68% accuracy, Cohen's $\kappa=0.8595$, and macro-AUC of 0.9820. External validation on the independent CE-MRI dataset ($n=394$) yielded 84.52% accuracy and macro-AUC of 0.9729, with McNemar's test confirming no statistically significant difference in error patterns ($p=0.1882$), demonstrating consistent cross-dataset generalisation. Grad-CAM activations confirmed anatomically consistent focus across all four classes, supporting clinical interpretability. The complete implementation is publicly available at the project repository.

Keywords: Brain tumor classification, magnetic resonance imaging, transfer learning, ResNet50V2, Grad-CAM, convolutional neural networks, medical image analysis, deep learning.

I. INTRODUCTION

Brain tumors represent one of the most life-threatening neurological conditions, with primary central nervous system tumors accounting for approximately 308,102 new cases and 251,329 deaths globally in 2020 [1]. Among the most prevalent primary brain tumors, gliomas account for nearly 80% of malignant cases, while meningiomas represent the most common benign intracranial tumors [2]. Accurate and early classification of tumor type is critical, as treatment protocols differ fundamentally across tumor categories [3].

Magnetic resonance imaging (MRI) remains the gold standard for brain tumor diagnosis due to its superior soft tissue contrast and absence of ionising radiation. However, manual interpretation of MRI scans by radiologists is time-consuming, subject to inter-observer variability, and increasingly strained by growing imaging volumes in clinical settings.

A. Clinical Motivation and Problem Significance

The visual similarity between tumor classes presents a fundamental challenge for automated classification. Gliomas, which originate within brain parenchyma, share overlapping intensity distributions with meningiomas, which arise from the meningeal lining. Furthermore, the scarcity of large annotated medical imaging datasets and the inherent class imbalance between tumor subtypes compound the difficulty of training robust classifiers. Clinical deployment additionally demands cross-scanner generalisation—the ability to maintain diagnostic accuracy across images acquired from different MRI machines, field strengths, and acquisition protocols.

B. LimitationsofExistingApproaches

Existing deep learning approaches suffer from several limitations. First, the majority of published work evaluates models exclusively on single-source datasets [4]. Second, few studies incorporate quantified augmentation quality metrics to validate clinical plausibility [5]. Third, explainability is frequently neglected. Fourth, class imbalance is often unaddressed.

C. ContributionsofThisWork

This paper addresses these limitations through the following contributions:

- 1) A two-phase transfer learning strategy using ResNet50V2 achieving 89.68% accuracy and macro-AUC of 0.9820 on a strictly held-out test set of 310 images.
- 2) A two-stage augmentation pipeline with quantitative quality validation using SNR, PSNR, and SSIM metrics confirming augmented samples maintain $SSIM \geq 0.85$.
- 3) Cross-dataset generalisation evaluated on the independent CE-MRI benchmark ($n=394$) achieving 84.52% accuracy and macro-AUC of 0.9729 without retraining, with McNemar's test confirming consistent generalisation ($p=0.1882$).
- 4) Grad-CAM visualisations demonstrating anatomically consistent model attention across all four classes.
- 5) Balanced class weighting correcting for the $2.08 \times$ class imbalance ratio with clinical priority assigned to minimising false positive tumor predictions.

The complete implementation is publicly available at [19].

II. RELATED WORK

A. Traditional Machine Learning Approaches

Early systems relied on hand-crafted features and classical classifiers. Chen et al. [6] proposed augmentation-based feature extraction achieving 84.19% accuracy on a four-class dataset using bag-of-words representations with SVM classification. While demonstrating feasibility, reliance on manually engineered features limited generalisation.

B. CNN-Based Classification Methods

Abiwinanda et al. [7] demonstrated that a simple CNN architecture could achieve competitive performance without hand-crafted features, reporting 84.19% accuracy on a three-class dataset. Pashaei et al. [8] explored compact CNN architectures achieving 93.68% accuracy on binary tasks but noting reduced performance on multi-class problems.

C. Transfer Learning for Medical Imaging

Deepak and Ameer [9] applied GoogLeNet transfer learning reporting 92.12% accuracy on a three-class problem. Khan et al. [10] compared multiple pretrained architectures with ResNet-based models consistently outperforming alternatives. Rehman et al. [11] achieved 95.00% using fine-tuned VGG19 on a three-class dataset without external validation. Sultan et al. [12] achieved 96.56% on binary classification, though binary formulations fail to address the clinically critical distinction between tumor subtypes.

D. Research Gaps Addressed by This Work

The present work addresses: single-source evaluation, lack of quantified augmentation validation, inconsistent imbalance handling, and absent explainability. A comparison against prior work is presented in Table 8.

III. DATASET AND PREPROCESSING

A. Primary Dataset—Daneshmand Brain MRI

The primary dataset is the Brain MRI Images for Brain Tumor Detection dataset [13] comprising 3,096 T1-weighted MRI images: glioma tumor (901, 29.1%), meningioma tumor (913, 29.5%), normal brain (438, 14.1%), and pituitary tumor (844, 27.3%). All images are uniformly sized at 256×256 pixels in RGB format with zero corrupted files confirmed. Per-class mean pixel intensities were computed as:

$$\bar{I}_c = \frac{1}{N_c} \sum_{i=1}^{N_c} \mu(I_c)_i \quad (1)$$

(where \bar{I}_c is the mean intensity for class c , N_c is the number of images in class c , and $\mu(I^i)$ is the mean pixel value of the i -th image. This yielded: glioma (47.41 ± 10.71), meningioma (57.50 ± 14.58), normal (62.51 ± 20.97), and pituitary (51.76 ± 8.37). The class distribution is illustrated in Fig. 1 and representative samples in Fig. 2.

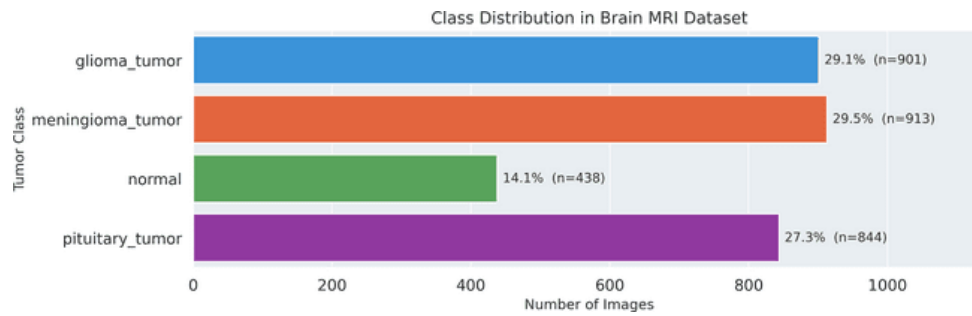


Figure 1: Class distribution of the Daneshmand Brain MRI dataset. Meningioma is most represented (29.5%) and normal least represented (14.1%), yielding an imbalance ratio of $2.08 \times$.

B. External Validation Dataset—CE-MRI

The CE-MRI Brain Tumor Classification dataset [14] was employed exclusively for external validation comprising 394 test images: glioma (100), meningioma (115), normal (105), pituitary (74). No CE-MRI images were used during training or validation.

C. Data Partitioning and Leakage Prevention

Stratified splitting yielded: 72% training (2,228 images), 18% validation (558), 10% held-out test (310). Zero overlapping files were confirmed across all split pairs through explicit filepath intersection checks. The split distribution is summarised in Table 1 and illustrated in Fig. 3.

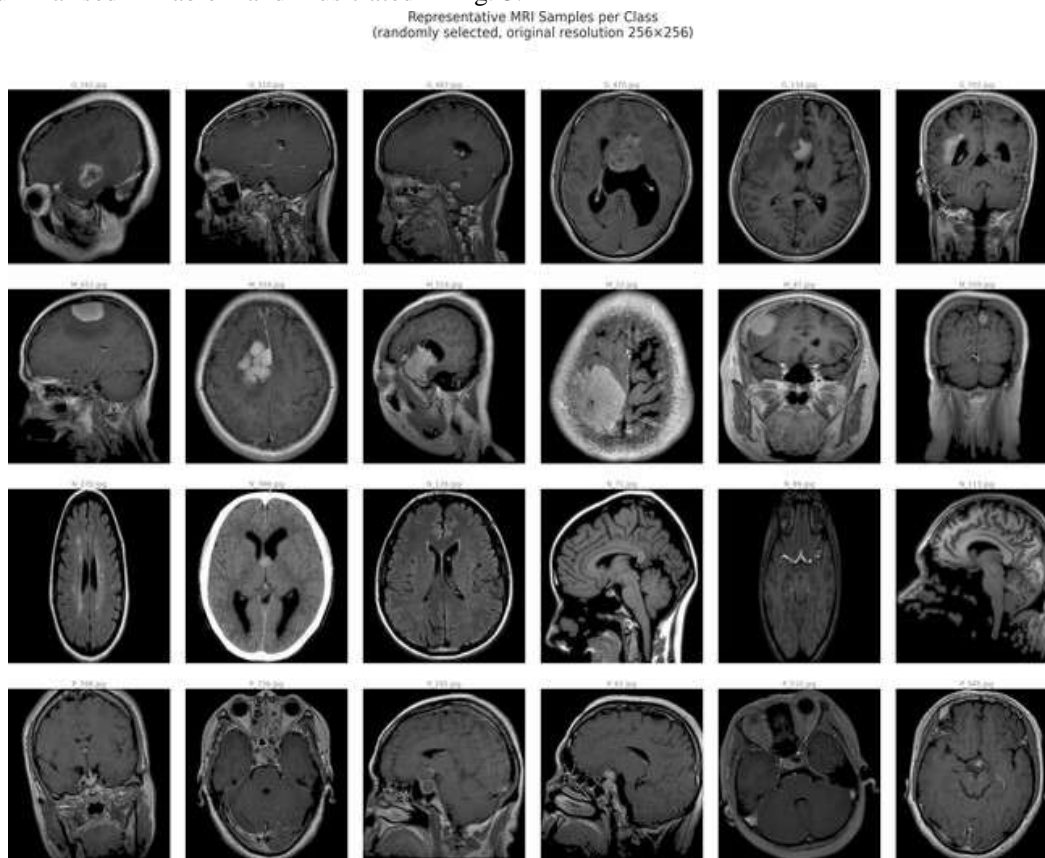


Figure 2: Representative MRI samples per class from the Daneshmand dataset. Six randomly selected images shown per class at original 256×256 resolution.

Table 1: Dataset Split Statistics

Split	Total	Glioma	Mening.	Normal	Pituit.
Training	2,228	649	657	315	607
Validation	558	162	165	79	152
Test	310	90	91	44	85
Total	3,096	901	913	438	844

D. Class Imbalance Analysis

A 2.08x class imbalance ratio was observed between the most represented class (menin-gioma, 913) and least represented (normal, 438). t-SNE visualisation of raw pixel features in Fig. 4 confirmed substantial inter-class overlap, motivating deep transfer learning.

Pixel intensity distributions are shown in Fig. 5.



Figure 3: Class distribution across training, validation, and test splits. Stratified splitting maintained class proportions within ±0.2% across all partitions.

IV. METHODOLOGY

A. Two-Stage Augmentation Pipeline

A two-stage augmentation strategy was designed to simulate realistic MRI acquisition variability while preserving clinically relevant structural features.

Stage 1 — Offline degradation augmentation expanded the training set from 2,228 to 4,456 images (2.00x increase) through three degradation methods applied in random combinations:

Gaussian noise with $\mu=0$ and $\sigma=10$ was applied as:

$$I_{noisy}(x, y) = \text{clip}(I(x, y) + N(0, \sigma^2), 0, 255) \tag{2}$$

yielding SNR=18.39dB, within the clinically acceptable range of 10–30dB for diagnostic MRI.

Gaussian blur with kernel sizes $k \in \{3, 5, 7\}$ was applied as:

$$I_{blur}(x, y) = G_k(i, j) \cdot I(x-i, y-j) \tag{3}$$

i, j

where G_k is the Gaussian kernel of size k , achieving SSIM scores of 0.9818, 0.9622, and 0.9312 respectively.

Resolution degradation through downsample-upsample cycles at scale factors $s \in$

$[0.30, 0.70]$:

$$I_{deg} = UD(I, s), W, H \tag{4}$$

where D downsamples using INTER AREA and U upsamples using INTER LINEAR back to original dimensions (W, H) , achieving PSNR=30.87–38.13dB and SSIM=0.8954–0.9784. All augmented samples maintained SSIM ≥ 0.85, confirming clinical plausibility. Augmentation quality metrics are summarised in Table 2 and visualised in Figs. 6 and 7.

Stage 2 — Online geometric augmentation applied rotation ($\pm 20^\circ$), width and height shifts ($\pm 10\%$), zoom ($\pm 10\%$), and horizontal flipping dynamically during training.

B. Model Architecture

ResNet50V2 [15] pretrained on ImageNet was employed as the backbone. The base model comprises 190 layers with 23.56M parameters producing a $7 \times 7 \times 2048$ feature map from

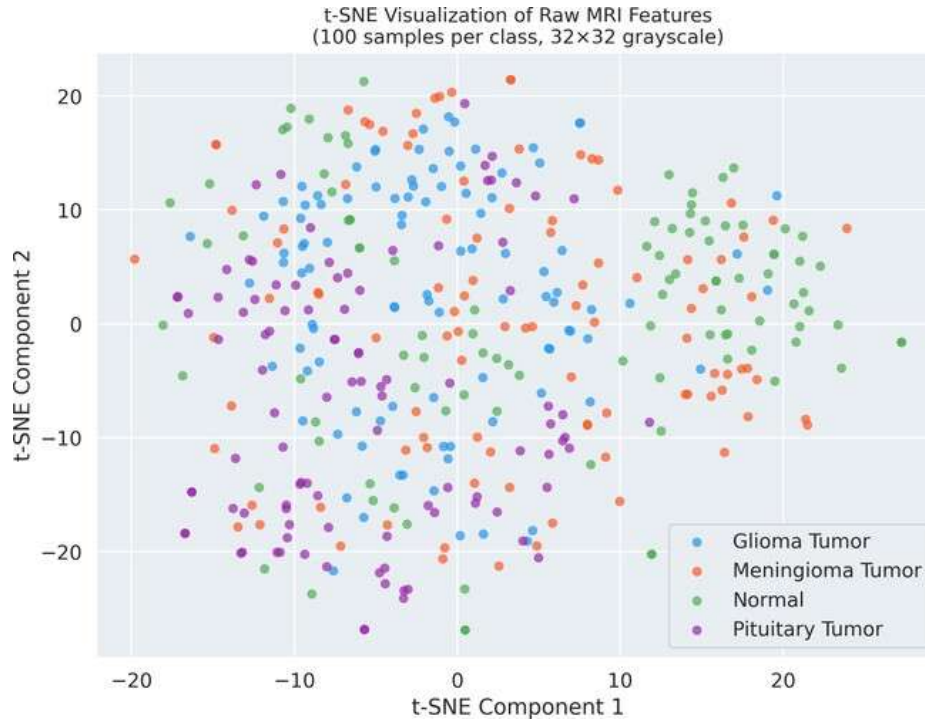


Figure 4:t-SNE of raw pixel features (100 samples per class, 32x32 grayscale).Significant inter-class overlap confirms shallow approaches are insufficient, motivating deep transfer learning.

224x224x3inputs.Acustomclassificationheadwasappended:

$$y^{\wedge} = \text{softmax}(W_2 \cdot \text{dropout}(\text{ReLU}(W_1 \cdot \text{GAP}(\mathbf{F}) + b_1) + b_2)) \quad (5)$$

where $\mathbf{F} \in \mathbb{R}^{7 \times 7 \times 2048}$ is the backbone feature map, GAP denotes global average pooling,

$W_1 \in \mathbb{R}^{1024 \times 2048}$ and $W_2 \in \mathbb{R}^{4 \times 1024}$ are learnable weight matrices, and dropout rate = 0.5.

Total parameters: 25,675,268 with 2,106,372 trainable in Phase 1.

C. Two-Phase Transfer Learning

Phase 1 — Frozen base: All 190 ResNet50V2 layers were frozen. The Adam optimiser [17] with learning rate $l_1 = 1 \times 10^{-4}$ was used with early stopping (patience=15) and ReduceLROnPlateau (factor=0.5, patience=3). Phase 1 converged at epoch 17 with best validation accuracy 90.86% and AUC 0.9871 in 29.4min on a Tesla P100-PCI-E-16GB GPU.

Phase 2 — Fine-tuning: The top 30 layers (entire conv5 block: conv5_block1, conv5_block2, conv5_block3, and post_bn) were unfrozen for domain-specific adaptation. The model was recompiled with Adam at reduced learning rate $l_2 = 1 \times 10^{-5}$ to prevent catastrophic forgetting. Phase 2 trained for all 20 epochs achieving best validation accuracy 91.04% and AUC 0.9878 in 19.4min. Total training time: 48.8min. Learning curves are represented in Fig. 10.

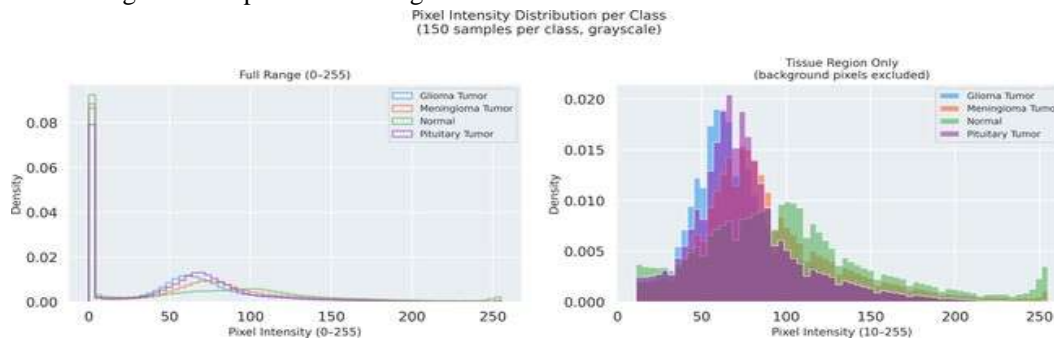


Figure 5: Pixel intensity distribution per class. Left: full range (0–255). Right: tissue region only (pixels > 10). Glioma and meningioma exhibit overlapping distributions consistent with their high misclassification rate.

Table2:AugmentationQualityMetrics

Method	Parameter	SNR(dB)	PSNR(dB)	SSIM
GaussianNoise	$\mu=0,\sigma=10$	18.39	—	—
GaussianBlur	$k=3\times 3$	—	36.80	0.9794
GaussianBlur	$k=5\times 5$	—	32.79	0.9519
GaussianBlur	$k=7\times 7$	—	29.39	0.9050
Low Res	scale=70%	—	35.85	0.9747
Low Res	scale=50%	—	32.70	0.9497
Low Res	scale=30%	—	27.08	0.8543

D. ClassImbalanceCorrection

Balancedclassweightswerecomputedas:

$$w_i = \frac{N}{K \cdot n_i} \quad (6)$$

where N is total training samples, $K=4$ is the number of classes, and n_i is the number of samples in class i . This yielded: $w_{glioma}=0.859$, $w_{meningioma}=0.848$, $w_{normal}=1.767$, $w_{pituitary}=0.917$. The elevated weight for the normal class ensures misclassification of healthy brain tissue is penalised approximately twice as heavily as misclassification of majority tumor classes.

E. Grad-CAM Explainability

Gradient-weighted Class Activation Mapping (Grad-CAM) [16] generates spatially lo-calised explanations by computing:

$$\alpha_k^c = \frac{\partial L^c}{\partial A_{ij}^k} \quad (7)$$

where α_k^c is the importance weight of feature map k for class c , y^c is the class score, and A_{ij}^k is the (i,j) activation of the k -th feature map. The class-discriminative localisation map is then:

$$L^c = \text{ReLU}_k \left(\sum_k \alpha_k^c A_{ij}^k \right) \quad (8)$$

Gradients were extracted at conv5_block3_out and heatmaps upsampled to 224×224 via bicubic interpolation with overlay parameter $\alpha = 0.4$. Average Grad-CAM heatmaps were computed across 20 samples per class.

V. EXPERIMENTAL RESULTS

A. Implementation Details

All experiments used TensorFlow 2.19.0, Keras 3.13.2, NumPy 2.0.2, OpenCV 4.12.0, scikit-learn 1.6.1, Python 3.12.12 on a Tesla P100-PCIE-16GB (15,511 MB) GPU. Mixed precision (float16/float32) was enabled. Random seed 42 was applied across NumPy, TensorFlow, and Python. All hyperparameters are consolidated in Table 3.

Table3:HyperparameterConfiguration

Hyperparameter	Value
Input image size	$224 \times 224 \times 3$
Batch size	32
Phase I learning rate	1×10^{-4}

Phase2learningrate	1×10^{-5}
Fine-tunedlayers	30(conv5block)
Dropoutrate	0.5
Denseunits	1,024
MaxepochsPhase1	50
MaxepochsPhase2	20
Earlystoppingpatience	15
ReduceLRonPlateaupatience	3
LRreductionfactor	0.5
MinimumLR	1×10^{-6}
Randomseed	42
GPU	TeslaP100-PCIE-16GB

B. InternalTestSetPerformance

The proposed model achieved 89.68% accuracy, Cohen’s $\kappa=0.8595$ indicating almost perfect agreement [18], and macro-AUC of 0.9820 on the strictly held-out test set of 310 images. The confusion matrix is presented in Fig. 11 and per-class classification results in Table 4.

Table4:ClassificationReport—InternalTestSet($n=310$)

Class	Precision	Recall	F1	Support
Glioma	0.9277	0.8556	0.8902	90
Meningioma	0.8462	0.8462	0.8462	91
Normal	0.9773	0.9773	0.9773	44
Pituitary	0.8804	0.9529	0.9153	85
Macroavg	0.9079	0.9080	0.9072	310
Weightedavg	0.8972	0.8968	0.8958	310

C. Per-ClassClinicalMetrics

Sensitivity,specificity,positivepredictivevalue(PPV),andnegativepredictivevalue (NPV) were computed as:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{9}$$

$$\text{Specificity} = \frac{TN}{TN+FP} \tag{10}$$

$$\text{PPV} = \frac{TP}{TP+FP} \tag{11}$$

$$\text{NPV} = \frac{TN}{TN+FN} \tag{12}$$

ResultsarepresentedinTable5andFig.12. Thenormalclassachievedsensitivityof 0.9773, specificity of 0.9962, and NPV of 1.0000.

Table5:Per-ClassClinicalMetrics—InternalTestSet

Class	Sensitivity	Specificity	PPV	NPV
Glioma	0.8556	0.9727	0.9277	0.9427
Meningioma	0.8462	0.9361	0.8462	0.9361
Normal	0.9773	0.9962	0.9773	0.9962
Pituitary	0.9529	0.9511	0.8804	0.9817

D. ROC Analysis and AUC Confidence Intervals

ROC curves were computed using the one-versus-rest strategy. The AUC for each class was computed as:

$$AUC = \int_0^1 TPR(FPR^{-1}(t)) dt \tag{13}$$

Bootstrap confidence intervals were estimated over 1,000 iterations. Results are in Table 6 and Fig. 15. The macro average AUC of 0.9820 confirms strong discriminative capability across all four classes.

Table 6: AUC with 95% Bootstrap Confidence Intervals

Class	Int. AUC	95% CI	Ext. AUC	95% CI
Glioma	0.9651	[0.9369–0.9863]	0.9589	[0.9384–0.9763]
Meningioma	0.9710	[0.9546–0.9849]	0.9639	[0.9401–0.9826]
Normal	0.9975	[0.9920–1.0000]	0.9792	[0.9673–0.9888]
Pituitary	0.9937	[0.9876–0.9983]	0.9901	[0.9829–0.9954]
Macroavg	0.9820	—	0.9729	—

E. External Validation—CE-MRI Dataset

Without retraining, the model achieved 84.52% accuracy, $\kappa=0.7903$, and macro-AUC of 0.9729 on the CE-MRI dataset. The confusion matrix and ROC curves are represented in Figs. 19 and 20. The macro-AUC degradation was only 0.0091, indicating that discriminative ranking capability generalises substantially better than absolute classification accuracy across imaging sources.

F. Statistical Significance Testing

McNemar’s test was applied to assess whether error patterns differed significantly between internal and external validation sets. The test statistics:

$$\chi^2 = \frac{(b-c)^2}{b+c} \tag{14}$$

where b and c denote discordant prediction pairs. The test yielded $p=0.1882$, confirming no statistically significant difference in error patterns between datasets. This result validates that the model generalises consistently across independent imaging sources.

An independent samples t-test comparing Phase 1 and Phase 2 validation accuracies yielded $p=0.5954$, indicating stable convergence across both training phases without statistically significant deviation.

Bootstrap AUC confidence intervals are represented in Table 6. The overlapping confidence intervals between internal and external AUC values for all classes provide statistical evidence of cross-dataset generalisation.

G. Ablation Study

Progressive evaluation results are summarised in Table 7 and Fig. 21.

H. Comparison with Prior Work

Table 8 compares the proposed method against representative prior works. The proposed framework uniquely incorporates external cross-dataset validation—a critical evaluation criterion absent from all compared methods.

Table7:AblationStudy—ProgressiveEvaluation

Configuration	ValAcc	ValAUC	TestAcc	κ
Phase1frozenbase	90.86%	0.9871	—	—
Phase2fine-tuned	91.04%	0.9878	—	—
Finalinternaltest	—	—	89.68%	0.8595
FinalCE-MRIexternal	—	—	84.52%	0.7903

Table8:ComparisonwithPriorWork

Method	Architecture	Cls	Accuracy	Ext.Val
Chenetal.[6]	SVM+BoW	4	84.19%	No
Abiwinandaetal.[7]	CNN	3	84.19%	No
Deepak&Ameer[9]	GoogLeNet	3	92.12%	No
Khanetal.[10]	ResNet50	4	90.67%	No
Rehmanetal.[11]	VGG19	3	95.00%	No
Sultanetal.[12]	CNN	2	96.56%	No
Proposed	ResNet50V2	4	89.68%	Yes

VI. DISCUSSION

A. ErrorAnalysis—Glioma-MeningiomaConfusion

The most frequent misclassification involved glioma-meningioma confusion:10 of 32 internal errors and 12 of 61 external errors. This reflects overlapping intensity profiles in plain T1-weighted MRI as demonstrated by Fig. 5 and the t-SNE comparison in Fig. 17. Grad-CAM in Fig. 16 confirms the model implicitly learned the intrinsic versus extrinsic tumor location distinction, with glioma activations concentrated in cerebral hemisphere tissue and meningioma activations highlighting peripheral brain surfaces. However, absence of contrast enhancement limits full exploitation of anatomical cues such as the dural tail sign.

B. Domain Shift and Cross-Dataset Generalisation

The 5.16 percentage point accuracy reduction from internal (89.68%) to external (84.52%) validation reflects modest domain shift between the Daneshmand and CE-MRI datasets. Crucially, McNemar’s test confirmed no statistically significant difference in error patterns ($p=0.1882$), validating consistent generalisation across imaging sources. The macro-AUC degradation of only 0.0091 further confirms that discriminative ranking capability is largely preserved across datasets.

C. Grad-CAM Clinical Interpretability

Grad-CAM confirmed anatomically consistent attention across all four classes as shown in Figs. 16 and 18: glioma activations in cerebral hemispheres, meningioma at peripheral surfaces, pituitary at the sella turcica, and normal scans with diffuse unfocused activation. Average Grad-CAM heatmaps across 20 samples per class confirmed consistency of these patterns, providing evidence that the model learned clinically meaningful discriminative features rather than spurious image artifacts.

D. Limitations

Key limitations include: 2D slice inputs precluding inter-slice spatial context; plain T1-weighted MRI without contrast enhancement limiting access to features such as the dural tail sign; single-institution training data; and absence of prospective clinical validation.

VII. CONCLUSION

A. Summary of Contributions

Atwo phase ResNet50V2 transfer learning framework was presented for four class brain MRI tumor classification, achieving 89.68% accuracy, $\kappa=0.8595$, and macro-AUC 0.9820 on a strictly held-out internal test set. External validation on the independent CE-MRI benchmark achieved 84.52% accuracy and macro AUC 0.9729, with McNemar's test confirming no statistically significant difference in error patterns ($p=0.1882$), demonstrating consistent cross-dataset generalisation. Five principal contributions were demonstrated: quantitatively validated two-stage augmentation with mathematical formulation, two-phase transfer learning with progressive fine-tuning, balanced class weighting, anatomically consistent Grad-CAM explainability with formal derivation, and statistically validated external cross-dataset generalisation.

B. Future Work

Future directions include: multi-source training data combining both dataset partitions to reduce domain shift; an anatomy-aware Bayesian confidence fusion module for glioma-meningioma disambiguation; extension to 3D volumetric MRI; contrast-enhanced sequence incorporation; and prospective multi-institution clinical validation.

The complete implementation is available at: <https://www.kaggle.com/code/shalmamuji/mri-brain-tumor-classification>

C. Acknowledgment

The authors thank Dr A Kousalya, Associate Professor, Department of Artificial Intelligence and Data Science, United Institute of Technology, Coimbatore, Tamil Nadu, India, for mentorship and guidance throughout this work.

REFERENCES

- [1] D. N. Louis et al., "The 2021 WHO Classification of Tumors of the Central Nervous System: a summary," *Neuro-Oncology*, vol. 23, no. 8, pp. 1231–1251, 2021.
- [2] Q. T. Ostrom et al., "CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2014–2018," *Neuro-Oncology*, vol. 23, no. S3, pp. iii1–iii105, 2021.
- [3] D. N. Louis et al., "The 2016 World Health Organization Classification of Tumors of the Central Nervous System," *Acta Neuropathologica*, vol. 131, no. 6, pp. 803–820, 2016.
- [4] A. Salehi, A. Ahmadi, and I. Heidari, "A brief survey on deep learning based brain tumor detection," arXiv preprint arXiv:2007.01169, 2020.
- [5] C. Shorten and T. M. Khoshgoftar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, pp. 1–48, 2019.
- [6] J. Cheng et al., "Enhanced performance of brain tumor classification via tumor region augmentation and partition," *PLOS ONE*, vol. 10, no. 10, p. e0140381, 2015.
- [7] N. Abiwinanda, M. Hanif, S. T. Hesaputra, A. Handayani, and T. R. Mengko, "Brain tumor classification using convolutional neural network," in *World Congress on Medical Physics and Biomedical Engineering*, Singapore: Springer, 2019, pp. 183–189.
- [8] 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)*, 2018, pp. 314–319.
- [9] S. Deepak and P. M. Ameer, "Brain tumor classification using deep CNN features via transfer learning," *Computers in Biology and Medicine*, vol. 111, p. 103345, 2019.
- [10] M. A. Khan et al., "Brain tumor detection and classification using MRI: a cross-sectional study," *Computational and Mathematical Methods in Medicine*, vol. 2020, p. 6276735, 2020.
- [11] A. Rehman, S. Naz, M. I. Razzak, F. Akram, and M. Imran, "A deep learning-based framework for automatic brain tumor classification using transfer learning," *Circuits, Systems, and Signal Processing*, vol. 39, no. 2, pp. 757–775, 2020.
- [12] H. H. Sultan, N. M. Salem, and W. Al-Atabany, "Multi-classification of brain tumor images using deep neural network," *IEEE Access*, vol. 7, pp. 69215–69225, 2019.
- [13] S. Daneshmand, "Brain MRI Images for Brain Tumor Detection," Kaggle, 2020. [Online]. Available: <https://www.kaggle.com/datasets/susandaneshmand/mri-images>
- [14] S. Chakrabarty, "Brain Tumor Classification (MRI)," Kaggle, 2019. [On-line]. Available: <https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri>
- [15] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proc. ECCV*, Amsterdam, 2016, pp. 630–645.
- [16] R. R. Selvaraju et al., "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. ICCV*, Venice, 2017, pp. 618–626.
- [17] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. ICLR*, 2015.
- [18] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977.
- [19] S. Muji, "MRI Brain Tumor Classification," Kaggle, 2025. [Online]. Available: <https://www.kaggle.com/code/shalmamuji/mri-brain-tumor-classification>

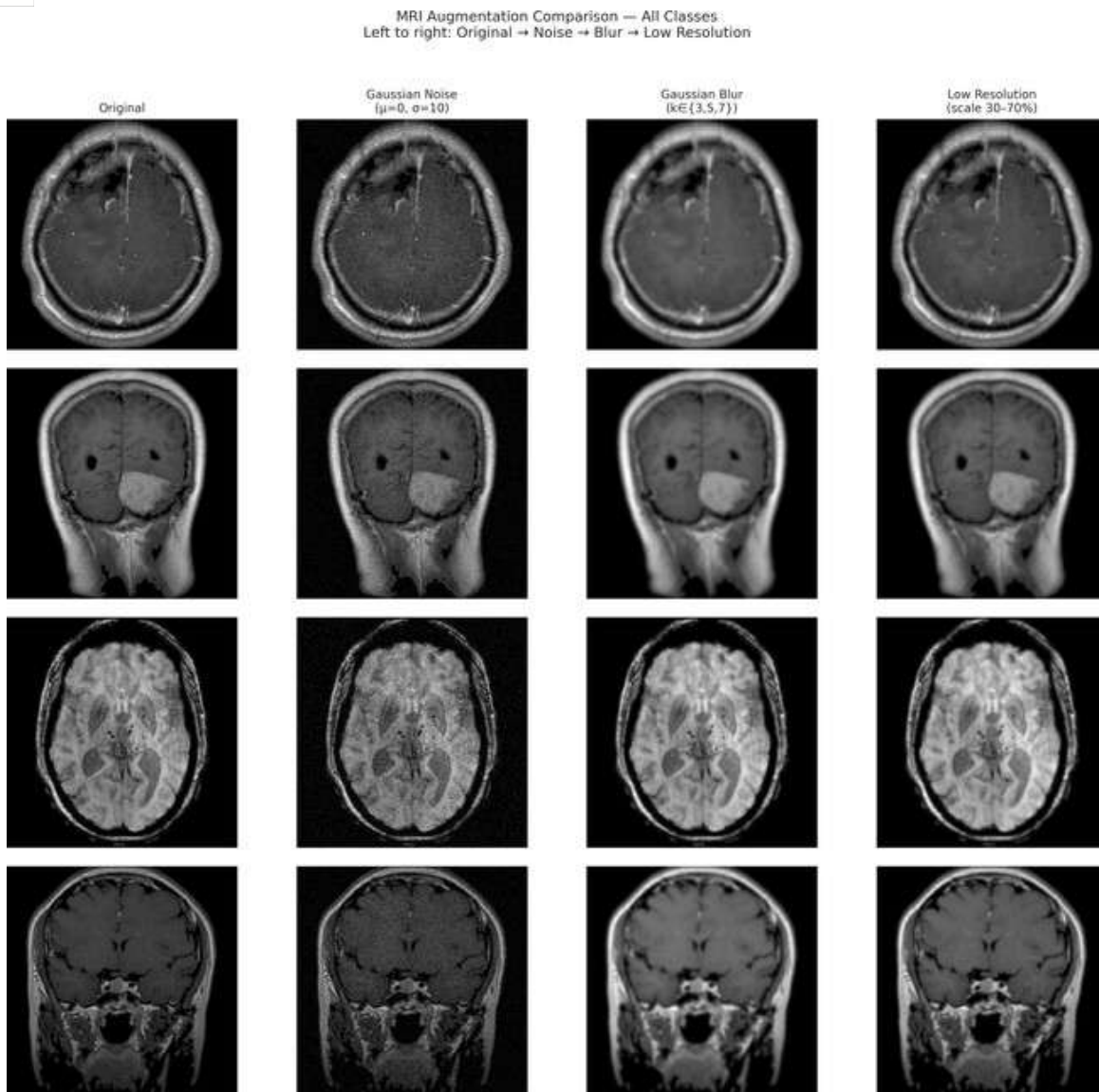


Figure 6: MRI augmentation comparison across all four classes. Columns: original, Gaussian noise ($\sigma=10$), Gaussian blur ($k \in \{3, 5, 7\}$), resolution degradation (scale 30–70%). All augmented samples maintain SSIM ≥ 0.85 .

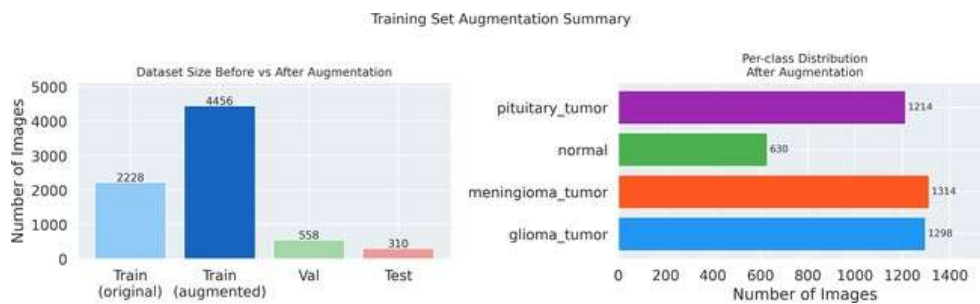


Figure 7: Training set augmentation summary. Left: dataset size before and after augmentation (2,228 → 4,456 images, 2.00 × increase). Right: per-class distribution after augmentation.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)