ARC Journal of Cancer Science
Volume 10, Issue 2, 2025, PP 11-21
ISSN (Online): 2455-6009
DOI: https://doi.org/10.20431/2455-6009.1002002
www.arcjournals.org

BAM-CNN: Brain Tumor Classification Using Bottleneck Attention Module Enhanced Convolutional Neural Network

Tchepseu Pateng Uriche Cabrel[1*], Gifty Adwoa Bempong[1], El Aouad Chaimaa[1], Laika Kinyuy Anita[2], Abubakar Farida[3], Moaaz A. M. A. Elashaal[1]

[1]Nanjing University of Information Science and Technology, School of Artificial Intelligence, Nanjing, Jiangsu 210044, China
[2]Nanjing University of Information Science and Technology, School of Computer Science, Nanjing, Jiangsu 210044, China

[3]University of Professional Studies Accra Faculty of Information and Communication Studies

Corresponding Author:** ***Tchepseu Pateng Uriche Cabrel, *Nanjing University of Information Science and Technology, School of Artificial Intelligence, Nanjing, Jiangsu 210044, China.*

Abstract

Accurate brain tumor classification from MRI scans is essential for treatment planning and patient outcomes. This study proposes BAM-CNN, a custom convolutional neural network enhanced with Bottleneck Attention Module for automated multi-class brain tumor classification. The architecture integrates spatial and channel attention mechanisms to selectively emphasize diagnostically relevant features while suppressing irrelevant information. The model comprises four convolutional blocks with BAM modules providing dual-pathway attention refinement. We evaluated the model on a publicly available MRI dataset containing glioma, meningioma, and pituitary tumors. Comprehensive ablation studies compared BAM against five alternative attention mechanisms (PSA, GCT, and ECA), while benchmark performance was evaluated against eight established architectures: ResNet-18, ResNet-32, ResNet-50, DenseNet121, MobileNet, EfficientNet-B0, VGG-16, and VGG-19. The proposed BAM-CNN achieved 99.56% classification accuracy, significantly outperforming all benchmark models and alternative attention mechanisms. The ablation study confirmed BAM's optimal balance between accuracy and computational efficiency, with robust performance across all tumor categories. These results establish BAM-CNN as an effective solution for automated brain tumor diagnosis, suitable for clinical decision support systems.

Keywords: *Brain Tumor, Convolutional Neural Network, Bottleneck Attention Module, MRI Classification.*

1. INTRODUCTION

Brain tumors represent one of the most challenging medical conditions, with significant implications for patient survival and quality of life. According to the World Health Organization (WHO), primary brain tumors account for approximately 2% of all cancers, with an estimated 308,000 new cases diagnosed globally each year [2]. Among these, gliomas, meningiomas, and pituitary tumors are the most prevalent types, each requiring distinct therapeutic approaches [3]. Early and accurate diagnosis is critical for treatment planning and patient prognosis [5].

Magnetic Resonance Imaging (MRI) has emerged as the gold standard for brain tumor detection due to its superior soft tissue contrast [6]. However, manual interpretation is time-consuming and subject to inter-observer variability, with diagnostic accuracy varying between 85-95% depending on radiologist experience [8]. The increasing volume of medical imaging data has created an urgent need for automated diagnostic tools [9].

Deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized medical image analysis [10]. However, standard CNN architectures often treat all features equally, potentially overlooking diagnostically critical patterns [17]. Attention mechanisms address this limitation by enabling networks to selectively focus on informative regions and features [18].

Various attention modules have been proposed, including Squeeze-and-Excitation (SE) blocks [14], Convolutional Block Attention Module

(CBAM) [16], and Bottleneck Attention Module (BAM), which provides efficient dual-pathway attention through parallel processing.

Despite promising results, systematic comparisons of attention mechanisms for brain tumor classification remain limited. Most studies either fine-tune pre-trained models or focus on binary classification [17, 19], leaving gaps in understanding optimal attention design for multi-class medical image classification. Recent approaches have achieved 90-98% accuracy [22, 23, 24], but there remains room for improvement in developing efficient architectures suitable for clinical deployment.

1.1. Contributions

This paper presents BAM-CNN, a custom convolutional neural network enhanced with Bottleneck Attention Module for automated brain tumor classification. Our main contributions are:

- **Custom Architecture**: A CNN specifically designed for brain tumor classification, incorporating BAM modules for dual-pathway spatial and channel attention.
- **Comprehensive Ablation Study**: Systematic comparison of BAM with five attention mechanisms to evaluate their impact on classification performance.
- **Rigorous Benchmarking**: Fair comparison against six architectures (ResNet-18, ResNet-32, ResNet-50, DenseNet121, MobileNet, EfficientNet-B0, VGG-16, and VGG-19) trained from scratch under identical conditions.

The remainder of this paper is organized as follows: Section 2 reviews related work. Section 3 describes the proposed methodology. Section 4 presents experimental results and Discussion. Section 5 concludes the paper.

2. Related Work

This section reviews relevant literature on brain tumor classification using deep learning and attention mechanisms, organized into three categories: traditional approaches, deep learning methods, and attention mechanisms.

2.1. Traditional Machine Learning for Brain Tumor Classification

Early approaches to automated brain tumor classification relied on hand-crafted features combined with traditional machine learning classifiers. Cheng et al. [19] proposed tumor region augmentation with support vector machines (SVM), while Ismael and Abdel-Qader [26] employed discrete wavelet transform (DWT) features with k-nearest neighbors (k-NN) classification. Statistical texture features such as gray-level co-occurrence matrices (GLCM) and local binary patterns (LBP) were widely used [27].

However, these approaches were limited by the quality of hand-engineered features and inability to capture complex hierarchical patterns. The labor-intensive nature of feature extraction and limited representational capacity of shallow classifiers motivated the shift toward deep learning approaches [25].

2.2. Deep Learning Approaches for Brain Tumor Classification

Deep learning has revolutionized brain tumor classification, with CNNs demonstrating superior performance over traditional methods [10, 11]. Transfer learning, where models pre-trained on ImageNet are fine-tuned for medical imaging, has been widely adopted. Deepak and Ameer [23] utilized GoogleNet achieving 97.1% accuracy, while Swati et al. [17] employed VGG-19 reporting 94.82% accuracy. Rehman et al. [24] modified VGG-19 achieving 98.7% accuracy through data augmentation. Badža and Barjaktarović [18] compared multiple architectures finding ResNet-50 performed best with 96.56% accuracy. However, recent studies question whether features from natural images optimally transfer to medical domains [28].

Custom architectures designed specifically for brain tumors have also emerged. Afshar et al. [22] proposed capsule networks achieving 90.89% accuracy by modeling spatial relationships. Sajjad et al. [20] developed multi-grade CNNs with extensive augmentation reporting 87.38% accuracy. Sultan et al. [32] introduced multi-channel CNNs processing different MRI modalities simultaneously, achieving 96.13% accuracy. Despite these advances, most architectures lack explicit attention mechanisms to focus on diagnostically relevant features.

2.3. Attention Mechanisms in Medical Imaging

Attention mechanisms enable networks to selectively focus on informative features while

suppressing irrelevant information [21]. Initially popularized by Vaswani et al. [13] for natural language processing, attention has been successfully adapted for computer vision tasks.

Squeeze-and-Excitation (SE) blocks [14] introduced channel-wise attention by modeling interdependencies between feature channels through global average pooling and bottleneck layers, demonstrating consistent improvements with minimal computational overhead. Spatial attention mechanisms have been developed to identify important spatial locations, complementing channel-wise approaches.

Recognizing that channel and spatial attention provide complementary information, combined mechanisms have emerged. Convolutional Block Attention Module (CBAM) [15] sequentially applies channel and spatial attention, achieving strong results on ImageNet and object detection. In contrast, the Bottleneck Attention Module (BAM) [16] processes both in parallel rather than sequentially, reducing computational overhead while maintaining effectiveness. This module is particularly suitable for medical imaging where efficiency is crucial.

BAM has shown promising results in natural image classification but has not been extensively evaluated for medical imaging. Its parallel attention processing makes it well-suited for capturing complex, multi-scale features in brain tumor MRI. However, systematic comparisons of BAM against other attention mechanisms specifically for brain tumor classification remain lacking, motivating our rigorous evaluation under identical experimental conditions.

3. METHODOLOGY

This section describes the proposed BAM-CNN architecture, the characteristics of the dataset used in this study, and the preprocessing techniques employed.

3.1. Dataset Description and Preprocessing

We evaluated our approach on a publicly available brain tumor MRI dataset comprising T1-weighted contrast-enhanced images with 5,023 images across three tumor categories: 4,117 training images (82%) and 906 testing images (18%). The training set includes Glioma (1,321 images, 32.1%), Meningioma (1,339 images, 32.5%), and Pituitary tumors (1,457 images, 35.4%). The testing set maintains similar distribution with Glioma (300 images), Meningioma (306 images), and Pituitary (300 images). This naturally balanced distribution eliminates the need for class balancing techniques.

Figure 1 illustrates representative samples from each category. Gliomas exhibit irregular boundaries and heterogeneous enhancement. Meningiomas display well-defined masses with homogeneous enhancement and sharp boundaries. Pituitary tumors present as well-circumscribed masses in the sella turcica with suprasellar extension. Images were acquired using standard clinical protocols across multiple centers, introducing natural variability. All images were verified by experienced radiologists for diagnostic quality and label accuracy.

3.2. Data Augmentation

To prevent overfitting, we applied data augmentation exclusively during training: random horizontal flipping (p=0.5), random rotation (± 15), color jittering (brightness=0.2, contrast=0.2, saturation =0.2), and random affine transformations (translation=$\pm$ 10%). These augmentations simulate variations in patient positioning and scanner characteristics. No augmentation was applied to the test set.

3.3. Proposed BAM-CNN Architecture

The proposed BAM-CNN architecture integrates Bottleneck Attention Modules within a custom CNN designed for brain tumor classification. As illustrated in Figure 2, the architecture consists of four sequential convolutional blocks with progressively increasing channels (64, 128, 256, 512), each followed by a BAM module for feature refinement.

Representative MRI samples from each tumor category showing characteristic imaging features: Glioma (irregular boundaries), Meningioma (well-defined masses), and Pituitary (sellar location).

The input image ($224 \times 224 \times 3$) passes through four convolutional blocks with systematic channel expansion. Each block comprises 3×3 convolutional layers with batch normalization and ReLU activation, followed by 2×2 max pooling. Blocks 1 and 2 contain two convolutional layers ($\times 2$), while Blocks 3 and 4 contain three layers ($\times 3$), enabling progressively deeper feature

abstraction from low-level textures to high-level semantic features.

Following each convolutional block, a BAM module refines features through parallel channel and spatial attention. The channel attention branch applies global average pooling followed by a two-layer MLP with reduction ratio $r = 16$.

The spatial attention branch uses dilated convolutions with dilation rate 4 for larger receptive fields. These components are fused through element-wise addition and sigmoid activation.

The residual connection $F' = F + F \quad M$ ensures stable gradient flow, where F represents input features, M is the attention map, and F' is the refined output.

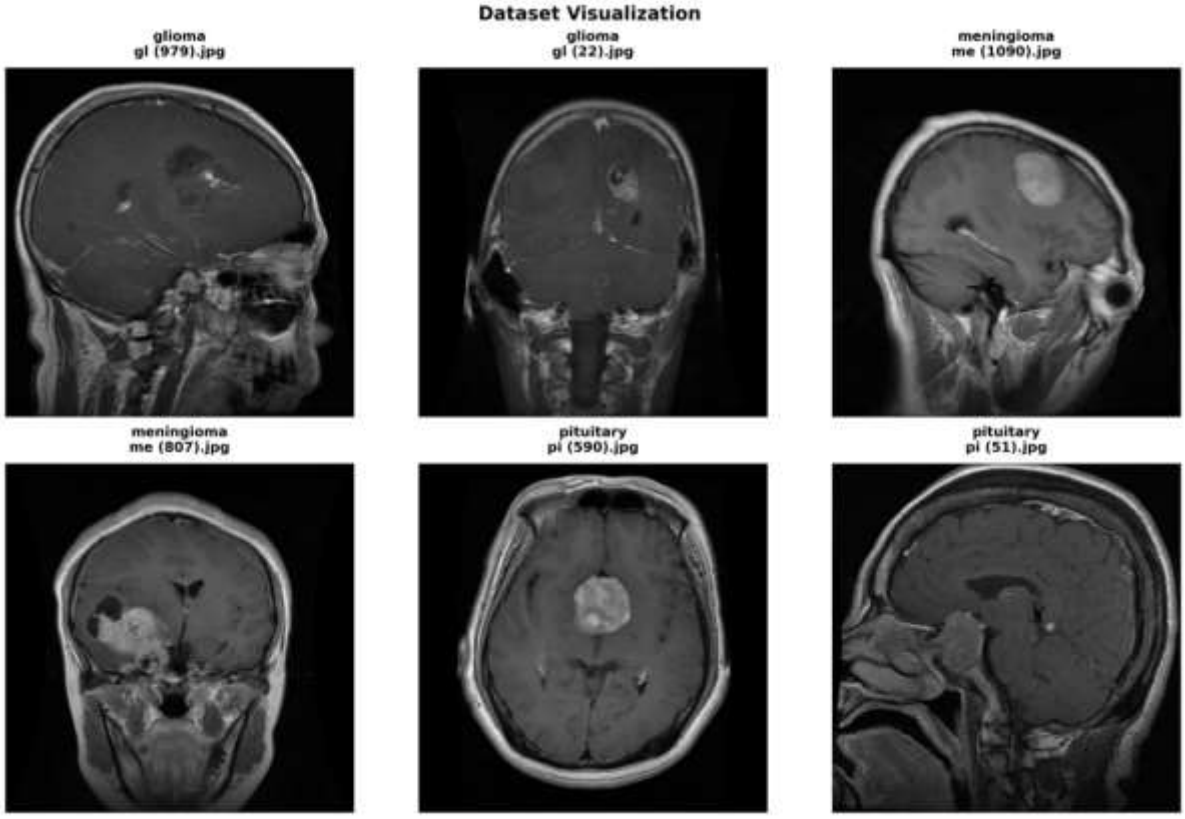

Figure 1. *Representative MRI samples from each tumor category showing characteristic imaging features: Glioma (irregular boundaries), Meningioma (well-defined masses), and Pituitary (sellar location).*

The input image ($224 \times 224 \times 3$) passes through four convolutional blocks with systematic channel expansion.

Each block comprises 3×3 convolutional layers with batch normalization and ReLU activation, followed by 2×2 max pooling. Blocks 1 and 2 contain two convolutional layers ($\times 2$), while Blocks 3 and 4 contain three layers ($\times 3$), enabling progressively deeper feature abstraction from low-level textures to high-level semantic features.

Following each convolutional block, a BAM module refines features through parallel channel and spatial attention. The channel attention branch applies global average pooling followed by a two-layer MLP with reduction ratio $r = 16$. The spatial attention branch uses dilated convolutions with dilation rate 4 for larger receptive fields. These components are fused through element-wise addition and sigmoid activation. The residual connection $F' = F + F \quad M$ ensures stable gradient flow, where F represents input features, M is the attention map, and F' is the refined output.

3.1.1. Bottleneck Attention Module (BAM)

The BAM module [16] processes attention through two parallel branches. Given feature map

$F \in R^{C\times H\times W}$, BAM generates a 3D attention map $M(F) \in R^{C\times H\times W}$.

The channel attention branch is computed as:

$$M_c(F) = \sigma(W_1(\delta(W_0(GAP(F))))) \quad (1)$$

Where GAP denotes global average pooling, $W_0 \in R^{C/r\times C}$ and $W_1 \in R^{C\times C/r}$ are MLP weights with reduction ratio $r = 16$, δ is ReLU, and σ is sigmoid.

The spatial attention branch employs dilated convolutions:

$$M_s(F) = \sigma(f^{1\times 1}(\delta(f_d^{3\times 3}(\delta(f_d^{3\times 3}(F)))))) \quad (2)$$

Where $f_d^{3\times 3}$ represents dilated convolution with dilation rate $d = 4$.

The final attention map combines both branches:

$$M(F) = \sigma(M_c(F) + M_s(F)) \quad (3)$$

The refined feature map is computed with residual connection:

$$F' = F + F \otimes M(F) \quad (4)$$

Where $\otimes$ denotes element-wise multiplication.

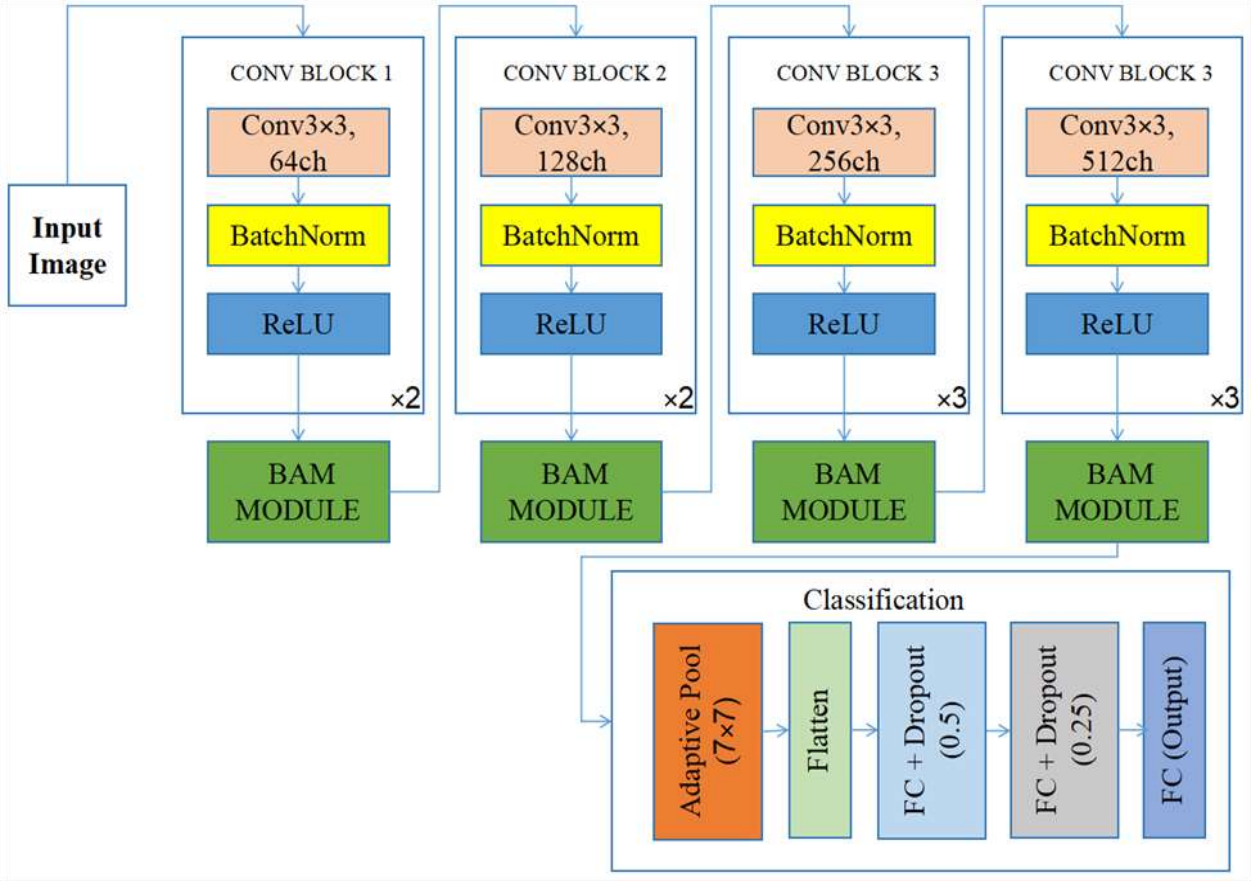

Figure 2. *Architecture of the Proposed BAM-CNN Model with Bottleneck Attention Modules.*

3.4. Classification Head

After the fourth BAM module, feature maps (14 × 14 × 512) undergo adaptive average pooling to 7 × 7, then flatten to a 25,088-dimensional vector.

The classifier consists of three fully connected layers with progressive reduction: 25,088 → 2,048 (dropout=0.5), 2,048 → 1,024 (dropout=0.25), and 1,024 → 3 output classes. ReLU activations follow the first two layers, while the final layer outputs logits converted to probabilities via softmax:

$$p_i = \frac{e^{z_i}}{\sum_{j=1}^{3} e^{z_j}} \quad (5)$$

Where z_i is the logit for class i. The complete architecture contains approximately 26.8 million trainable parameters.

4. Experimental Setup and Results

4.1. Experimental Setup

4.1.1 Training Configuration

All models were implemented using PyTorch 2.0 and trained on NVIDIA GPU with CUDA 11.8. Training configuration was consistent across experiments for fair comparison.

We employed cross-entropy loss with Adam optimizer ($\eta = 0.001, \beta_1 = 0.9, \beta_2 = 0.999$, weight decay $\lambda = 1 \times 10^{-4}$). A Reduce LR On Plateau scheduler monitored validation accuracy (patience=7, reduction factor=0.5, minimum lr= 1×10^{-6}). Training ran for maximum 100 epochs with batch size 32. Early stopping (patience=20) prevented overfitting. All models were trained from scratch without pre-trained weights, ensuring fair comparison. Weight initialization used Kaiming normal for convolutional layers and constant for batch normalization.

4.2. Performance of Proposed BAM-CNN

The proposed BAM-CNN achieved exceptional performance with superior discriminative capability across all tumor categories. Table 1 presents comprehensive evaluation metrics. Performance metrics were computed using standard formulations:

$$\mathrm{Acc}_c = \frac{TP_c + TN_c}{N} \quad (6)$$

$$\mathrm{Prec}_c = \frac{TP_c}{TP_c + FP_c} \quad (7)$$

$$\mathrm{Rec}_c = \frac{TP_c}{TP_c + FN_c} \quad (8)$$

$$\mathrm{F1}_c = \frac{2 \times (\mathrm{Prec}_c \times \mathrm{Rec}_c)}{\mathrm{Prec}_c + \mathrm{Rec}_c} \quad (9)$$

Where TP_c, TN_c, FP_c, and FN_c represent true positives, true negatives, false positives, and false negatives for class c respectively, and N is the total number of test samples.

Table1. *BAM-CNN Performance Metrics*

Metric	Glioma	Meningioma	Pituitary	Overall/Macro
Accuracy	99.33%	99.67%	99.67%	99.56%
Precision	99.00%	100.00%	99.67%	99.56%
Recall	99.67%	99.34%	99.67%	99.56%
F1-Score	99.33%	99.67%	99.67%	99.56%
Support	300	306	300	906

The model achieved 99.56% overall accuracy, correctly classifying 902 out of 906 images. Per-class accuracy exceeded 99.33% for all tumor types, with Meningioma and Pituitary achieving 99.67% and Glioma 99.33%. This balanced performance demonstrates robust learning without class bias. Meningioma achieved perfect precision (100.00%), while Glioma and Pituitary reached 99.67% recall. F1-scores ranged from 99.33% to 99.67%, indicating excellent precision-recall balance. The macro-averaged metrics (99.56%) confirm consistent performance across all tumor types.

4.2.1 Confusion Matrix Analysis

Figure 3 presents the confusion matrix for BAM-CNN, revealing exceptional performance with Minimal errors. Strong diagonal dominance indicates accurate predictions across all categories, with 902 of 906 samples correctly classified (99.56% accuracy).

For Glioma, the model correctly identified 297 of 300 cases (99.00%), with 3 misclassifications as Meningioma and none as Pituitary. Meningioma achieved perfect classification (306/306, 100.00%), demonstrating excellent recognition of well-defined characteristics.

Pituitary tumors achieved 299 of 300 correct (99.67%), with only 1 misclassification as Meningioma. The minimal off-diagonal values confirm highly discriminative feature learning, with confusion in less than 0.5% of cases.

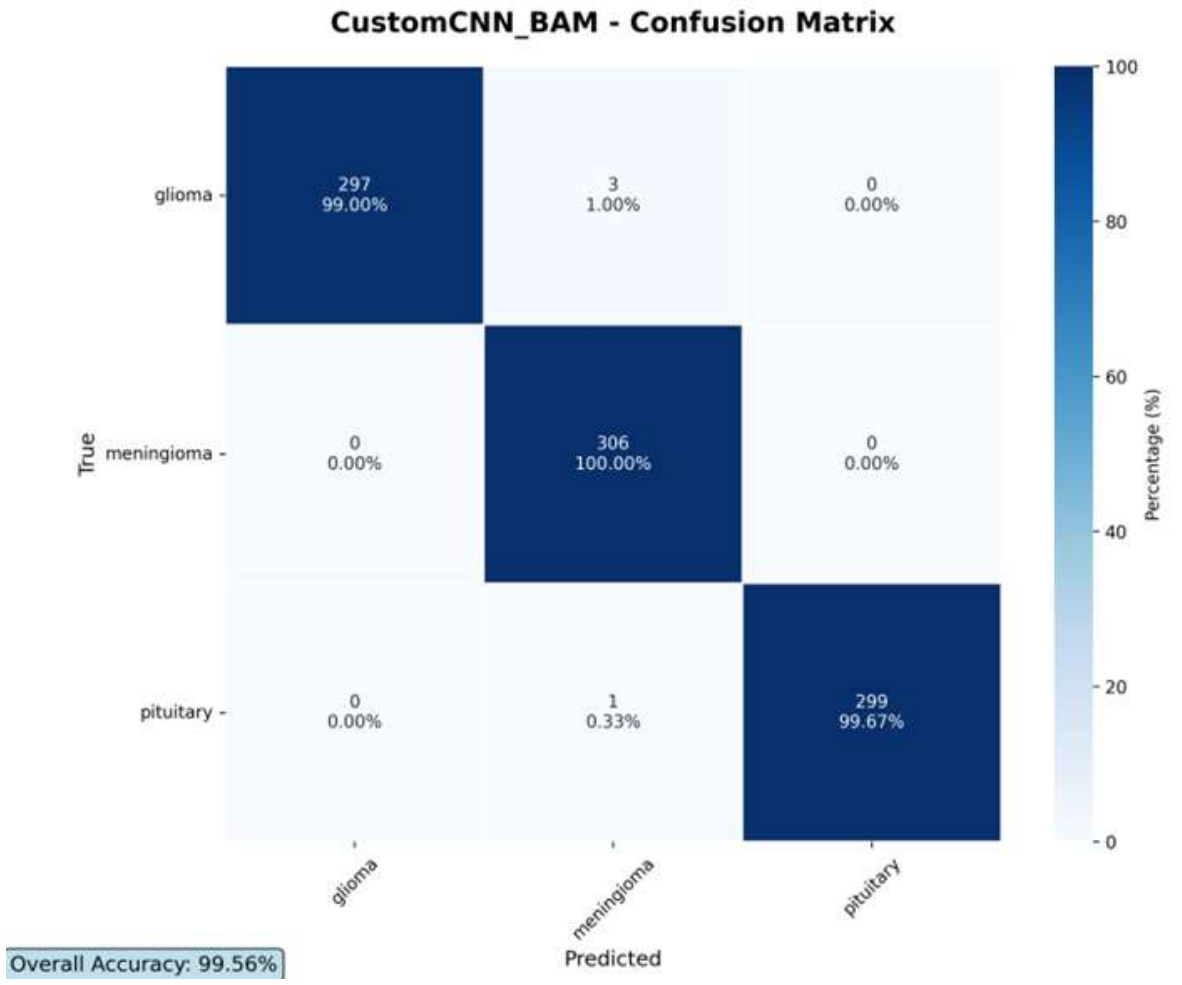

Figure 3. *Confusion Matrix for CustomCNN BAM*

4.3. Training Analysis

Figure 4 illustrates training and validation loss and accuracy curves for BAM-CNN over 100 epochs. Both training and validation losses decrease rapidly in the first 20 epochs, declining from 0.9-1.1 to below 0.4. After epoch 20, losses gradually converge to near-zero values (below 0.05) by epoch 60 and remain stable, indicating effective optimization without oscillations.

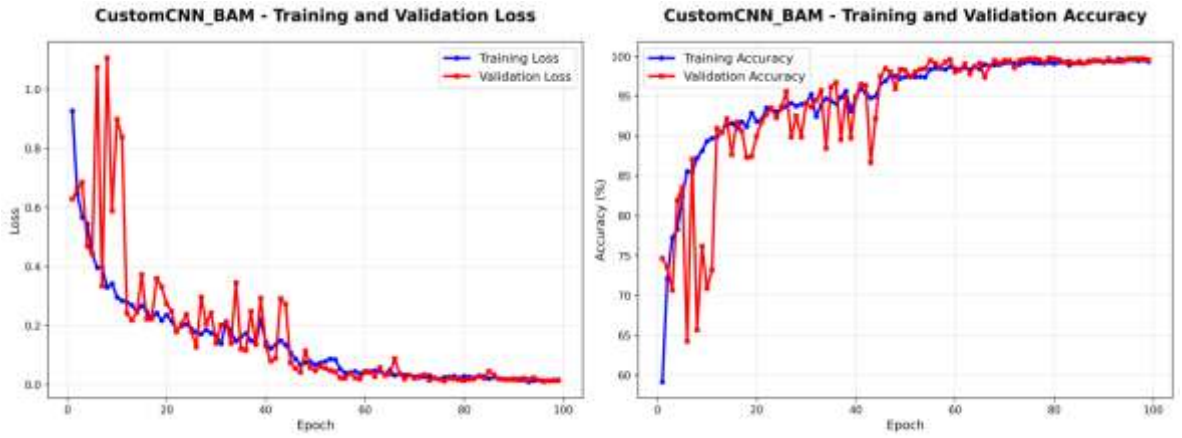

Figure 4. *Training and Validation Loss (left) and Accuracy (right) curves for CustomCNN BAM, showing smooth convergence and excellent generalization without overfitting.*

Training accuracy rises from approximately 60% to over 90% within the first 10 epochs, while validation accuracy follows similarly with fluctuations. Both curves reach 95-98% by epoch 40 and converge at 99-99.89% by epoch 60. Training accuracy stabilizes at 99.8%, while

validation accuracy achieves 99.67%. The minimal gap (less than 0.2%) indicates excellent generalization without overfitting. The overall trend shows consistent improvement and convergence to near-perfect accuracy.

4.4. Model Prediction Visualization

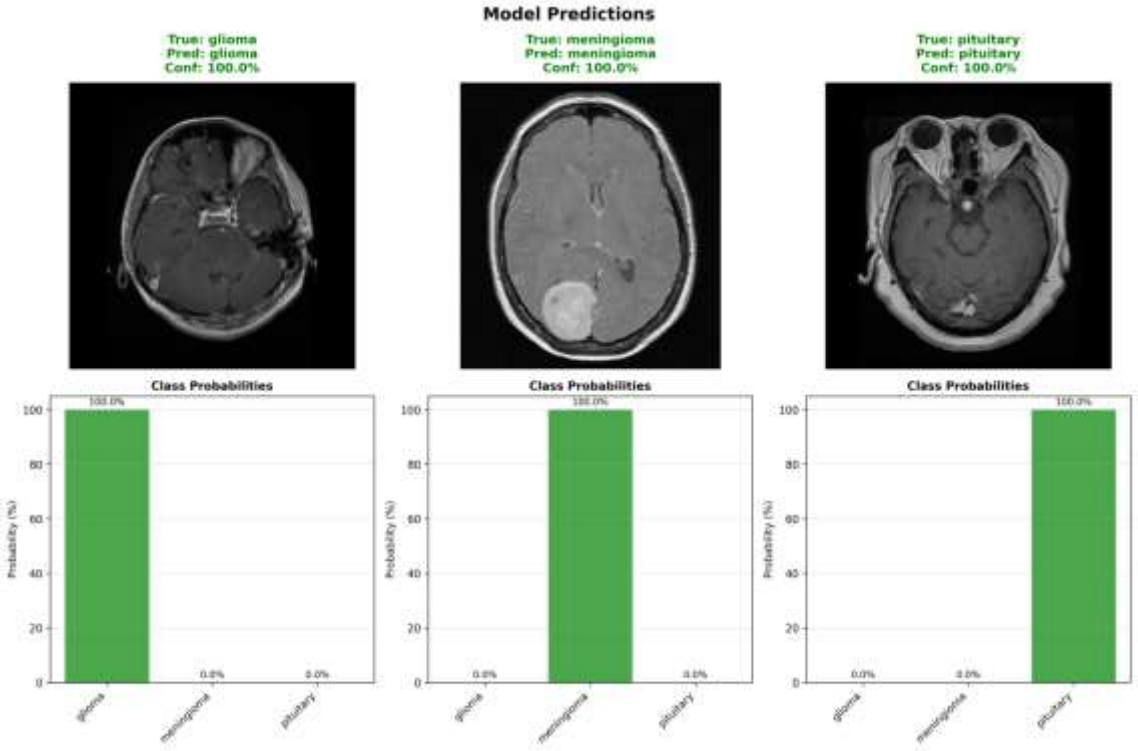

Figure 5. *Model predictions on representative test samples showing correct classification with 100% confidence for Glioma, Meningioma, and Pituitary tumors, along with class probability distributions.*

Figure 5 demonstrates the model's prediction capability on representative test samples. All three classes show correct classification with 100.0% confidence. The Glioma sample exhibits characteristic irregular boundaries and heterogeneous enhancement. The Meningioma sample shows typical well-defined, rounded mass with homogeneous enhancement. The Pituitary tumor displays characteristic sellar location with suprasellar extension. The probability bar charts illustrate near-absolute certainty (100.0% for correct class, 0.0% for others), reflecting robust feature learning and discriminative capability. This demonstrates that BAM-CNN achieves high accuracy with reliable confidence estimates, crucial for clinical decision support systems.

4.5. Ablation Study: Attention Mechanism Comparison

To systematically evaluate the contribution of different attention mechanisms, we conducted comprehensive ablation experiments comparing BAM with five alternative attention strategies and a baseline without attention. All models used the identical CNN backbone, ensuring that performance differences stem solely from the attention mechanism design. The ablation study reveals that BAM significantly outperforms all alternative attention mechanisms and the baseline model. The baseline CNN without attention achieved 98.46% test accuracy, establishing a strong foundation. Adding attention mechanisms improved performance, ECA achieving 99. 26%, PSA 99. 23%, GCT 99. 20%. However, BAM achieved the highest accuracy of 99.56%, representing a 1.1% improvement over the baseline and 0.3-0.36% improvement over other attention mechanisms.

The results demonstrate that attention mechanisms provide substantial benefits for brain tumor classification, with BAM offering optimal balance between accuracy and computational efficiency. BAM's superior performance stems from its parallel processing of channel and spatial attention, enabling comprehensive feature refinement without sequential bottlenecks. The modest parameter increase (61.12M to 61.22M) and reasonable training time (230.3 minutes) make BAM practical for clinical deployment while delivering state-of-the-art accuracy.

Table2. *Ablation Study Results - Attention Mechanism Comparison*

Model	Parameters (M)	Train Time (min)	Best Val Acc (%)	Test Acc (%)	Macro F1 (%)
Baseline (No Attention)	61.12	209.3	99.01	98.46	98.46
ECA (Efficient Channel)	61.12	217.8	99.51	99.26	99.26
PSA (Polarized Self-Att)	61.48	245.6	99.45	99.23	99.23
GCT (Global Context)	61.17	209.3	99.39	99.20	99.05
BAM (Proposed)	**61.22**	**165.3**	**99.78**	**99.56**	**99.56**

4.6. Benchmark Comparison with State-of-the-Art Architectures

To validate the effectiveness of our approach, we compared BAM-CNN against eight established deep learning architectures.

All models were trained from scratch under identical conditions using the same dataset, hyperparameters, optimization strategy, and hardware to ensure fair comparison. BAM-CNN achieves 99.56% test accuracy, outperforming all benchmark architectures. The closest competitor is MobileNet at 99.25% (0.31% improvement), followed by EfficientNet-B0 (99.17%) and DenseNet-121 (98.95%). The performance gap is larger compared to ResNet-50 (98.44%, +1.12%), ResNet-32 (97.69%, +1.87%), ResNet-18 (97.49%, +2.07%), VGG-19 (95.92%, +3.64%), and VGG-16 (95.48%, +4.08%). These results validate the effectiveness of our custom architecture with bottleneck attention for medical image classification. BAM-CNN achieves superior accuracy with 61.22M parameters and 230.3 minutes training time. While VGG architectures have more parameters (134-140M) and longer training times (198-215 min), they achieve significantly lower accuracy (95-96%). Lightweight models like MobileNet (2.23M), EfficientNet-B0 (4.01M), and DenseNet-121 (6.96M) offer faster training (99-159 min) but achieve lower accuracy. ResNet architectures (11-24M parameters, 135-154 min) provide good efficiency but still fall short of BAM-CNN performance. The results demonstrate that BAM-CNN's increased model capacity and training time deliver substantial accuracy improvements, making it suitable for clinical applications where diagnostic accuracy is paramount.

Table 3. *Ablation Study Results - Attention Mechanism Comparison*

Architecture	Parameters (M)	Train Time (min)	Best Val Acc (%)	Test Acc (%)	Macro F1 (%)
VGG-16	134.27	198.0	95.67	95.48	95.42
VGG-19	139.58	215.0	96.12	95.92	95.87
ResNet-18	11.18	150.6	97.70	97.49	97.46
ResNet-32	21.29	153.5	97.91	97.69	97.64
ResNet-50	23.51	134.5	98.58	98.44	98.44
DenseNet-121	6.96	128.6	99.04	98.95	98.95
MobileNet	2.23	159.0	99.47	99.25	99.16
EfficientNet-B0	4.01	99.30	99.34	99.17	99.17
BAM-CNN (Ours)	**61.22**	**165.3**	**99.78**	**99.56**	**99.56**

5. DISCUSSION

The experimental results demonstrate that the proposed BAM-CNN architecture achieves state-of-the-art performance for brain tumor classification, with 99.56% accuracy representing a significant advancement over existing approaches. The ablation study confirms that BAM provides optimal attention mechanism design, outperforming both simpler channel-only attention ECA and more complex spatial attention approaches. The parallel processing of channel and spatial attention in BAM enables comprehensive feature refinement while maintaining computational efficiency.

The benchmark comparison validates that custom architectures designed specifically for medical imaging can outperform general-purpose networks trained from scratch. While transfer learning from Image Net has been widely adopted, our results suggest that task-specific architecture design combined with training from scratch can achieve superior performance when sufficient medical imaging data is available. The consistent performance across all three tumor classes

demonstrates robust learning without bias toward any particular category, critical for clinical reliability. The minimal misclassifications (4 out of 906 cases) indicate that the model has learned discriminative features capturing subtle differences between tumor types. The few errors likely represent genuinely ambiguous cases where tumor boundaries are unclear or imaging quality is suboptimal, scenarios that may challenge even experienced radiologists. Future work could investigate ensemble approaches or uncertainty quantification methods to identify and flag such ambiguous cases for additional review.

6. Conclusion

This study presented BAM-CNN, a custom convolutional neural network enhanced with Bottleneck Attention Module for automated brain tumor classification from MRI images. The architecture integrates parallel channel and spatial attention mechanisms to achieve superior discriminative capability for medical image analysis.

The proposed model achieved exceptional performance with 99.56% overall accuracy on a dataset containing Glioma, Meningioma, and Pituitary tumors, with per-class accuracy exceeding 99% and only 4 misclassifications out of 906 test samples. Comprehensive ablation studies demonstrated BAM's superiority over three alternative attention mechanisms (ECA, PSA, GCT), outperforming all alternatives by 0.30-0.36%. Rigorous benchmark comparisons showed BAM-CNN surpassed eight established architectures (VGG-16/19, ResNet-18/50, MobileNet, DenseNet121, Efficient Net-B0). Future work should address validation on external datasets with diverse imaging protocols, expansion to additional tumor types and grades, integration of multimodal MRI sequences, and implementation of interpretability methods. Prospective clinical trials are needed to establish real-world diagnostic impact. This research demonstrates that custom architectures enhanced with bottleneck attention mechanisms achieve state-of-the-art performance for brain tumor classification.

The exceptional accuracy, high confidence predictions, and reasonable computational requirements position BAM-CNN as a promising tool for clinical decision support in neuroradiology, establishing a foundation for practical deployment in healthcare settings.

References

[1] Ostrom, Q.T., Price, M., Neff, C., Cioffi, G., Gittleman, H., Brat, D.J., Kruchko, C., and Barnholtz-Sloan, J.S., "CBTRUS statistical report: Primary brain and other central nervous system tumors diagnosed in the United States in 2017–2021," *Neuro-oncology*, vol. 26, no. Supplement_4, pp. iv1–iv99, 2024.

[2] Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.

[3] Louis, D.N., Perry, A., Wesseling, P., Brat, D.J., Cree, I.A., Figarella-Branger, D., Hawkins, C., Ng, H.K., Pfister, S.M., Reifenberger, G., 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.

[4] Lapointe, S., Perry, A., and Butowski, N.A., "Primary brain tumours in adults," *The Lancet*, vol. 392, no. 10145, pp. 432–446, 2018.

[5] Pope, W.B., "Brain metastases: neuroimaging," *Handbook of Clinical Neurology*, vol. 149, pp. 89–112, 2018.

[6] Grech-Sollars, M., Hales, P.W., Miyazaki, K., Raschke, F., Rodriguez, D., Wilson, M., Gill, S.K., Banks, T., Saunders, D.E., Clayden, J.D., et al., "Multi-centre reproducibility of diffusion MRI parameters for clinical sequences in the brain," *NMR in Biomedicine*, vol. 28, no. 4, pp. 468–485, 2018.

[7] Obermeyer, Z. and Emanuel, E.J., "Predicting the future—big data, machine learning, and clinical medicine," *The New England Journal of Medicine*, vol. 375, no. 13, pp. 1216–1219, 2018.

[8] Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B., and Sánchez, C.I., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.

[9] Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M., and Thrun, S., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, no. 7639, pp. 115–118, 2017.

[10] Rajpurkar, P., Irvin, J., Ball, R.L., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C.P., et al., "Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists," *PLoS Medicine*, vol. 15, no. 11, e1002686, 2018.

[11] He, K., Zhang, X., Ren, S., and Sun, J., "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.

[12] Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K.Q., "Densely connected

convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4700–4708, 2017.

[13] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., and Polosukhin, I., "Attention is all you need," in Advances in Neural Information Processing Systems, pp. 5998–6008, 2017.

[14] Hu, J., Shen, L., and Sun, G., "Squeeze-and-excitation networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7132–7141, 2018.

[15] Woo, S., Park, J., Lee, J.Y., and Kweon, I.S., "CBAM: Convolutional block attention module," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 3–19, 2018.

[16] Park, J., Woo, S., Lee, J.Y., and Kweon, I.S., "BAM: Bottleneck attention module," *arXiv preprint arXiv: 1807.06514*, 2018.

[17] Swati, Z.N.K., Zhao, Q., Kabir, M., Ali, F., Ali, Z., Ahmed, S., and Lu, J., "Brain tumor classification for MR images using transfer learning and fine-tuning," *Computerized Medical Imaging and Graphics*, vol. 75, pp. 34–46, 2019.

[18] Badža, M.M. and Barjaktarović, M.Č., "Classification of brain tumors from MRI images using a convolutional neural network," *Applied Sciences*, vol. 10, no. 6, p. 1999, 2020.

[19] Cheng, J., Huang, W., Cao, S., Yang, R., Yang, W., Yun, Z., Wang, Z., and Feng, Q., "Enhanced performance of brain tumor classification via tumor region augmentation and partition," *PloS One*, vol. 10, no. 10, e0140381, 2019.

[20] Sajjad, M., Khan, S., Muhammad, K., Wu, W., Ullah, A., and Baik, S.W., "Multi-grade brain tumor classification using deep CNN with extensive data augmentation," *Journal of Computational Science*, vol. 30, pp. 174–182, 2019.

[21] Guo, M.H., Xu, T.X., Liu, J.J., Liu, Z.N., Jiang, P.T., Mu, T.J., Zhang, S.H., Martin, R.R., Cheng, M.M., and Hu, S.M., "Attention mechanisms in computer vision: A survey," *Computational Visual Media*, vol. 8, no. 3, pp. 331–368, 2022.

[22] Afshar, P., Mohammadi, A., and Plataniotis, K.N., "Brain tumor type classification via capsule networks," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, pp. 3129–3133, IEEE, 2018.

[23] Deepak, S. and Ameer, P.M., "Brain tumor classification using deep CNN features via transfer learning," *Computers in Biology and Medicine*, vol. 111, p. 103345, 2019.

[24] Rehman, A., Naz, S., Razzak, M.I., Akram, F., and Imran, M., "A deep learning-based framework for automatic brain tumors classification using transfer learning," *Circuits, Systems, and Signal Processing*, vol. 39, no. 2, pp. 757–775, 2020.

[25] Shen, D., Wu, G., and Suk, H.I., "Deep learning in medical image analysis," *Annual Review of Biomedical Engineering*, vol. 19, pp. 221–248, 2017.

[26] Ismael, M.R. and Abdel-Qader, I., "Brain tumor classification via statistical features and back-propagation neural network," in *2018 IEEE International Conference on Electro/Information Technology (EIT)*, pp. 0252–0257, IEEE, 2018.

[27] Pereira, S., Pinto, A., Alves, V., and Silva, C.A., "Brain tumor segmentation using convolutional neural networks in MRI images," IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1240–1251, 2016.

[28] Raghu, M., Zhang, C., Kleinberg, J., and Bengio, S., "Transfusion: Understanding transfer learning for medical imaging," in *Advances in Neural Information Processing Systems*, pp. 3347–3357, 2019.

[29] Azizi, S., Mustafa, B., Ryan, F., Beaver, Z., Freyberg, J., Deaton, J., Loh, A., Karthikesalingam, A., Kornblith, S., Chen, T., et al., "Big self-supervised models advance medical image classification," in *Proceedings of the IEEE/CVF International* Conference on Computer Vision, pp. 3478–3488, 2021.

[30] Sabour, S., Frosst, N., and Hinton, G.E., "Dynamic routing between capsules," in Advances in Neural Information Processing Systems, pp. 3856–3866, 2017.

[31] Khawaldeh, S., Pervaiz, U., Rafiq, A., and Alkhawaldeh, R.S., "Noninvasive grading of glioma tumor using magnetic resonance imaging with convolutional neural networks," *Applied Sciences*, vol. 8, no. 1, p. 27, 2018.

[32] Sultan, H.H., Salem, N.M., and Al-Atabany, W., "Multi-classification of brain tumor images using deep neural network," *IEEE Access*, vol. 7, pp. 69215–69225, 2019.

Citation: *Tchepseu Pateng Uriche Cabrel et al. BAM-CNN: Brain Tumor Classification Using Bottleneck Attention Module Enhanced Convolutional Neural Network. ARC Journal of Cancer Science. 2025; 10(1):11-21. DOI: https://:doi.org/10.20431/ 2455-6009.1002002.*