



INDEREX-NET: AN ATTENTION-GUIDED MULTI-STREAM CNN FRAMEWORK WITH ADAPTIVE FEATURE FUSION FOR BRAIN TUMOR CLASSIFICATION

Sena ÇELİK^{1*}

¹ Mehmet Akif Ersoy University, Faculty of Computer and Information Sciences, Department of Information Systems Engineering, Bucak, Burdur, Türkiye

Keywords	Abstract
<i>Brain Tumor Classification, Medical Image Analysis, Attention-Based Deep Learning, Multi-Stream Convolutional Neural Networks, Adaptive Feature Fusion.</i>	Accurate classification of brain tumors from magnetic resonance imaging (MRI) data remains a challenging problem for computer-aided diagnosis systems due to inter-class similarity, intra-class variability, and heterogeneous tumor morphology. In this study, an attention-guided multi-stream deep learning architecture, termed IndereX-Net, is proposed for brain tumor classification from axial MRI images. The proposed model integrates three heterogeneous ImageNet-pretrained convolutional neural network backbones, namely ResNet50, DenseNet121, and InceptionV3, arranged in parallel streams to extract complementary multi-scale feature representations. Each stream is enhanced with a Convolutional Block Attention Module (CBAM), while Generalized Mean (GeM) pooling is employed for adaptive feature aggregation. In addition, a gated feature fusion mechanism is incorporated to dynamically regulate inter-stream interactions. The model is trained using a two-phase transfer learning strategy and evaluated on a publicly available brain tumor MRI dataset. Experimental results demonstrate that IndereX-Net outperforms individual backbone models, achieving an accuracy of 98.22% and a ROC-AUC of 0.9986, indicating that attention-guided multi-stream architectures provide an effective and reliable solution for computer-aided brain tumor diagnosis.

INDEREX-NET: BEYİN TÜMÖRÜ SINIFLANDIRMASI İÇİN UYARLANABİLİR ÖZELLİK FÜZYONUNA SAHİP DİKKAT TABANLI ÇOK AKIŞLI CNN MİMARİSİ

Anahtar Kelimeler	Öz
<i>Beyin Tümörü Sınıflandırması, Tıbbi Görüntü Analizi, Dikkat Tabanlı Derin Öğrenme, Çok Akışlı Evrimsel Sinir Ağları, Uyarlanabilir Özellik Füzyonu.</i>	Manyetik rezonans görüntüleme (MRI) verilerinden beyin tümörlerinin doğru sınıflandırılması; sınıflar arası benzerlik, sınıf içi değişkenlik ve tümör morfolojisindeki heterojenlik nedeniyle bilgisayar destekli tanı sistemleri açısından zorlu bir problemdir. Bu çalışmada, aksiyal MRI görüntülerinden beyin tümörü sınıflandırması için IndereX-Net adı verilen, dikkat mekanizması ile yönlendirilmiş çok akışlı bir derin öğrenme mimarisi önerilmektedir. Önerilen model; ResNet50, DenseNet121 ve InceptionV3 olmak üzere, ImageNet üzerinde önceden eğitilmiş üç heterojen evrimsel sinir ağı omurgasını paralel akışlar halinde bütünleştirerek tamamlayıcı çok ölçekli özellik temsilleri elde etmektedir. Her bir akış Convolutional Block Attention Module (CBAM) ile güçlendirilmiş, uyarlanabilir özellik birleştirme için Generalized Mean (GeM) pooling kullanılmıştır. Ayrıca, akışlar arası etkileşimi dinamik olarak düzenleyen kapılı özellik füzyon mekanizması mimariye entegre edilmiştir. Model, iki aşamalı transfer öğrenme stratejisi ile eğitilmiş ve herkese açık bir beyin tümörü MRI veri seti üzerinde değerlendirilmiştir. Deneysel sonuçlar, IndereX-Net'in tekil omurga modellerine kıyasla daha üstün performans sergilediğini; %98,22 doğruluk ve 0,9986 ROC-AUC değerlerine ulaştığını göstererek, dikkat destekli çok akışlı mimarilerin bilgisayar destekli beyin tümörü tanısı için etkili ve güvenilir bir çözüm sunduğunu ortaya koymaktadır.

Alıntı / Cite

Celik, S., (2026). IndereX-Net: An Attention-Guided Multi-Stream CNN Framework with Adaptive Feature Fusion for Brain Tumor Classification, Journal of Engineering Sciences and Design, 14(3), 719-735.

Yazar Kimliği / Author ID (ORCID Number)

S. Çelik, 0000-0001-9277-7623

Makale Süreci / Article Process

Başvuru Tarihi / Submission Date	09.02.2026
Revizyon Tarihi / Revision Date	19.04.2026
Kabul Tarihi / Accepted Date	06.08.2026
Yayın Tarihi / Published Date	25.09.2026

* İlgili yazar / Corresponding author: scelik@mehmetakif.edu.tr, +90-248-213-8160

INDEREX-NET: AN ATTENTION-GUIDED MULTI-STREAM CNN FRAMEWORK WITH ADAPTIVE FEATURE FUSION FOR BRAIN TUMOR CLASSIFICATION

Sena ÇELİK^{1†},

¹Mehmet Akif Ersoy University, Faculty of Computer and Information Sciences, Department of Information Systems Engineering, Bucak, Burdur, Türkiye

Highlights

- A multi-stream CNN architecture is proposed for MRI-based brain tumor classification.
- Attention-guided feature extraction is employed to enhance spatial discriminability.
- Adaptive gated fusion is utilized to integrate complementary deep features.
- The model is evaluated on a public MRI dataset with strong classification performance.

Graphical Abstract

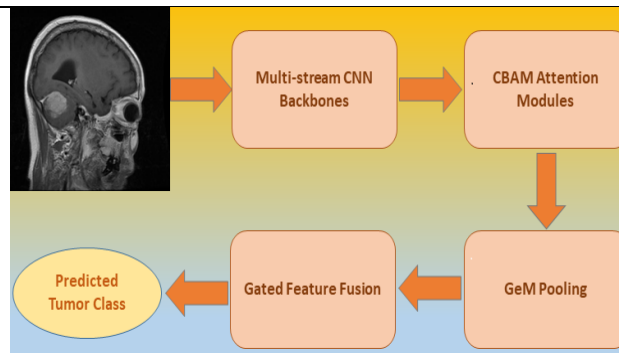


Figure. Overall architecture of the proposed IndereX-Net framework.

Purpose and Scope

The purpose of this study is to improve the accuracy and robustness of brain tumor classification from MRI images by addressing challenges such as inter-class similarity, intra-class variability, and heterogeneous tumor appearance. The study focuses on developing an attention-guided multi-stream deep learning framework that can effectively extract complementary feature representations for reliable computer-aided diagnosis.

Design/methodology/approach

This study proposes IndereX-Net, a multi-stream convolutional neural network architecture integrating ResNet50, DenseNet121, and InceptionV3 backbones. Each stream is enhanced with Convolutional Block Attention Modules to highlight diagnostically relevant regions, while Generalized Mean pooling is used for adaptive feature aggregation. A gated feature fusion mechanism dynamically combines multi-stream features. The model is trained using a two-phase transfer learning strategy and evaluated on a publicly available brain tumor MRI dataset.

Findings

Experimental results demonstrate that the proposed IndereX-Net model achieves superior performance compared to individual backbone networks. The model attains a classification accuracy of 98.22% and a ROC-AUC score of 0.9986, along with balanced precision, recall, and F1-scores, indicating strong generalization capability and robustness in brain tumor classification tasks.

Research limitations/implications

Although the proposed model shows high performance, it is evaluated on a single public MRI dataset, which may limit its generalizability across different imaging protocols and clinical settings. Future studies may focus on validating the model on larger and multi-center datasets and extending the framework to multi-modal or 3D MRI data for improved clinical applicability.

Practical implications

The proposed framework can serve as an effective decision-support tool for radiologists by providing accurate and reliable brain tumor classification. Its modular architecture allows easy adaptation to other medical image classification problems, potentially reducing diagnostic workload and improving consistency in clinical practice.

Social Implications

Improved accuracy in brain tumor diagnosis can contribute to earlier detection and more informed treatment planning, positively impacting patient outcomes and quality of life. The proposed approach supports the broader adoption of artificial intelligence in healthcare by enhancing trust in computer-aided diagnostic systems.

Originality

This study introduces an attention-guided multi-stream deep learning architecture with adaptive gated feature fusion for brain tumor classification. The integration of heterogeneous CNN backbones with attention mechanisms and dynamic feature fusion represents a novel contribution, offering improved performance and robustness compared to conventional single-stream models.

[†] Corresponding author: scelik@mehmetakif.edu.tr, +90-248-213-8160

1. Introduction

Reliable identification of brain tumors at an early stage plays a decisive role in neuro-oncological practice, as diagnostic accuracy directly affects therapeutic strategies, risk assessment, and patient outcomes. The intrinsic diversity of brain tumors in terms of shape, spatial distribution, and tissue characteristics poses significant challenges for visual assessment, particularly when interpretations depend solely on manual inspection. Although magnetic resonance imaging (MRI) provides high-resolution anatomical detail and superior soft-tissue contrast, routine evaluation procedures remain highly dependent on expert judgment. This dependency often results in observer-dependent variability and limits diagnostic consistency across clinical settings. Consequently, there is a growing demand for automated and reproducible analysis frameworks that can assist clinicians by offering objective and data-driven diagnostic support (Agadi et al., 2023).

Recent advances in artificial intelligence have reshaped medical image analysis by enabling models to learn discriminative representations directly from imaging data. In this context, convolutional neural networks (CNNs) have emerged as a dominant paradigm due to their ability to capture multiscale spatial patterns without relying on manually designed features. Applied to annotated MRI datasets, CNN-based systems have reported encouraging performance in differentiating among major brain tumor types, such as gliomas, meningiomas, and pituitary tumors (Ali et al., 2025; Sumona et al., 2025). To mitigate data scarcity and improve generalization, transfer learning strategies have been widely adopted, allowing pretrained architectures—including ResNet, DenseNet, and Inception variants—to be repurposed for tumor classification tasks with limited labeled samples (Asif et al., 2023).

Beyond baseline architectures, contemporary research has increasingly emphasized model refinement techniques that enhance robustness and representation diversity. Mechanisms such as attention-guided feature modulation, ensemble-based prediction aggregation, and multi-level feature fusion have been explored to better capture tumor-specific characteristics. Moreover, recent architectural developments, including EfficientNet and transformer-inspired models, have contributed to improved sensitivity against anatomical variability and imaging heterogeneity (Lu, 2024; Tariq, 2024). Despite these methodological advances, several obstacles persist, notably data imbalance, scanner-dependent variability, and insufficient transparency of model decisions. These factors continue to restrict the seamless integration of deep learning systems into routine clinical workflows. To address such limitations, recent studies have highlighted the importance of interpretability-oriented designs and adaptive feature selection strategies that enhance both predictive performance and clinician confidence (Kumar and Jyoti, 2024).

In response to these challenges, this study introduces the Integrated Deep Residual Extractor Network (IndereX-Net), an attention-enhanced multi-stream CNN architecture designed for robust brain tumor classification from MRI data. The proposed framework combines three complementary pretrained backbones—ResNet50, DenseNet121, and InceptionV3—each augmented with convolutional block attention modules, generalized mean pooling, and a gated feature fusion strategy. By exploiting diverse feature representations within a unified learning scheme, IndereX-Net aims to improve both classification accuracy and interpretability. The effectiveness of the proposed model is assessed using a publicly available T1-weighted contrast-enhanced MRI dataset encompassing glioma, meningioma, pituitary tumor, and no-tumor categories. Experimental findings demonstrate that IndereX-Net surpasses commonly used baseline models across multiple performance metrics, including accuracy, ROC-AUC, and F1-score, underscoring its potential applicability as a reliable computer-aided diagnostic tool in neuro-oncology.

The remainder of this manuscript is organized as follows. Section 2 provides a comprehensive review of related studies. Section 3 describes the materials and methods employed in this research. Section 4 presents the experimental results and their analysis. Finally, Section 5 concludes the study and outlines potential directions for future research.

2. Literature Survey

The application of magnetic resonance imaging (MRI) for automated brain tumor analysis has gained substantial momentum with the rapid advancement of artificial intelligence-based medical imaging techniques. Owing to the pronounced variability in tumor morphology, spatial localization, and histopathological structure, traditional image interpretation remains cognitively demanding and prone to inter-observer inconsistencies. Consequently, research efforts have increasingly focused on developing data-driven classification systems capable of delivering consistent and reproducible diagnostic outcomes.

Among learning-based approaches, convolutional neural networks (CNNs) have established themselves as a

foundational methodology for MRI-based tumor classification. By learning multilevel spatial features directly from raw image data, CNNs effectively bypass the limitations associated with handcrafted descriptors. Prior studies have demonstrated that CNN-driven models can achieve high discriminative performance when distinguishing gliomas, meningiomas, pituitary tumors, and healthy brain tissues (Ali et al., 2025; Sumona et al., 2025). Comparable outcomes have also been reported across broader medical imaging applications, reinforcing the suitability of CNNs for modeling complex anatomical patterns (Cheng et al., 2023). Nevertheless, reliance on single-backbone architectures often constrains robustness, particularly in scenarios involving class imbalance or visually overlapping tumor categories.

To overcome these constraints, recent investigations have explored architectural enhancements and attention-oriented learning mechanisms. Attention modules facilitate selective emphasis on diagnostically informative regions while suppressing irrelevant background features, thereby improving class separability. For instance, Chen et al. (2024) introduced an attention-integrated separable convolutional framework that improved feature discrimination while reducing overfitting effects. Similarly, Asif et al. (2023) reported performance gains through attention-guided transfer learning, highlighting its effectiveness in multiclass tumor classification. Transfer learning strategies leveraging pretrained convolutional and transformer-based backbones have further contributed to improved generalization and reduced training overhead (Hsu et al., 2024).

Ensemble learning has emerged as another prominent strategy for enhancing predictive stability and generalization. By aggregating complementary representations from multiple models, ensemble approaches mitigate the sensitivity of individual networks to data variability. Lu (2024) demonstrated that ensembles constructed from EfficientNet variants consistently outperform standalone architectures in MRI-based tumor classification tasks. Hybrid ensemble configurations combining deep neural feature extractors with conventional classifiers have additionally shown improvements in precision and recall, particularly under heterogeneous data distributions (Pal et al., 2024; Khan et al., 2023). Complementary studies further indicate that coupling CNN-derived features with classifiers such as Support Vector Machines enhances resilience against noise and class imbalance (Kochher, 2025; Alrikabi and Muhammad, 2024). More recently, ensemble formulations incorporating vision transformer architectures have shown promising results in capturing long-range spatial dependencies within tumor regions (Tariq, 2024).

Beyond classification-focused frameworks, tumor segmentation has been widely investigated as a complementary task to support anatomical localization and clinical interpretation. Encoder-decoder models inspired by U-Net and its variants have demonstrated strong segmentation performance while simultaneously benefiting downstream classification accuracy (Mohan and Ramkumar, 2024; Shaji et al., 2023). Siamese network configurations have been proposed to improve robustness across diverse imaging conditions (Murugesan et al., 2025), whereas multimodal MRI fusion techniques have exploited complementary information from multiple imaging sequences to enhance glioma subtype differentiation (Janardhan and Thomas, 2025).

Preprocessing and data augmentation strategies remain integral components of robust learning pipelines. Techniques such as intensity normalization, adaptive filtering, and generative adversarial network-based augmentation have been employed to address data scarcity and distributional imbalance in MRI datasets (Bianchessi et al., 2023). In parallel, research has explored lightweight network architectures and alternative imaging modalities to reduce computational cost while preserving diagnostic accuracy (Hossain et al., 2023).

In recent years, explainable artificial intelligence (XAI) has attracted growing attention as a prerequisite for clinical integration. Visualization techniques, including Grad-CAM and SHAP, have enabled more transparent interpretation of model predictions by identifying influential image regions (Kumar and Jyoti, 2024; Mastoi et al., 2025). Federated learning paradigms have also been proposed to address privacy and scalability challenges by facilitating collaborative training without centralized data exchange (Al-Saleh et al., 2025). Moreover, hybrid systems that combine deep feature extraction with computationally efficient classifiers have demonstrated the feasibility of privacy-aware and resource-efficient diagnostic solutions (Zhang et al., 2023).

Despite notable progress, many existing approaches remain constrained by static fusion strategies or limited exploitation of complementary feature spaces. These shortcomings highlight the necessity for adaptive, attention-driven, multi-stream architectures capable of jointly leveraging diverse representations. Motivated by this gap, the proposed IndereX-Net framework aims to integrate attention-guided feature extraction with gated fusion mechanisms to achieve robust, generalizable, and clinically reliable brain tumor classification.

3. Material and Method

This study proposes an attention-guided multi-stream deep learning framework, termed IndereX-Net, for robust brain tumor classification from magnetic resonance imaging (MRI) scans. The overall methodology comprises dataset preparation and preprocessing, transfer learning with multiple pretrained convolutional neural network (CNN) backbones, attention-enhanced feature extraction, adaptive feature aggregation, and supervised model optimization. The proposed framework integrates architectural diversity through parallel backbone networks, refines discriminative features using attention mechanisms, and combines complementary representations through learnable fusion strategies. A two-phase training protocol is employed to ensure stable convergence and effective domain adaptation. The methodological components are described in detail in the following subsections.

3.1. Dataset and Preprocessing

The experiments in this study were conducted using the publicly available Brain Tumor MRI Dataset hosted on Mendeley Data (Hira et al., 2025). The dataset consists of 12,064 T1-weighted contrast-enhanced axial MRI images categorized into four classes: glioma, meningioma, pituitary tumor, and no tumor. The dataset was initially divided into training and test sets using an 80:20 split, where the test set was defined at the outset of the experimental pipeline and remained completely isolated throughout all stages of model development.

From the training portion, 20% of the samples were further reserved as a validation subset using the built-in validation split mechanism of the data generator. Accordingly, model development was conducted using training, validation, and test partitions, while preserving the original class distribution across subsets. The dataset exhibits moderate class imbalance, with glioma and pituitary tumor classes being more prevalent than the no-tumor class. The overall class distribution of the dataset is presented in Figure 1.

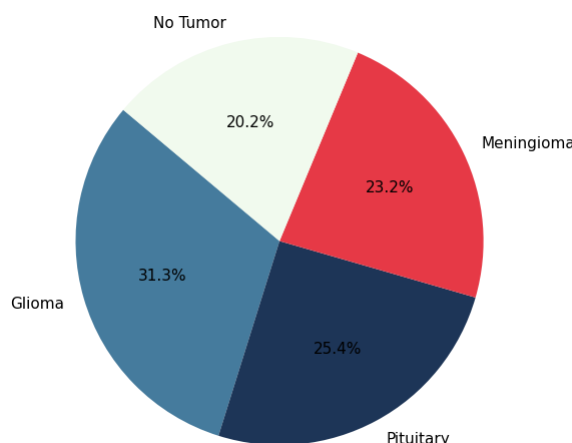


Figure 1. Class distribution of the brain tumor MRI dataset.

To ensure methodological rigor and minimize the risk of data leakage, strict separation was maintained between training, validation, and test data throughout the entire experimental pipeline. The test set was never used during model training, hyperparameter tuning, or model selection. The validation subset was derived exclusively from the training data and was used only for performance monitoring and early stopping. This design ensures that the reported results reflect the true generalization capability of the model on unseen data.

Data augmentation techniques, including random rotation, translation, zooming, and horizontal flipping, were applied exclusively to the training subset to improve model robustness and reduce overfitting. No augmentation or transformation was applied to the validation or test data. In addition, separate data generators were used for the training/validation and test phases, preventing unintended information transfer between subsets.

All MRI images were resized to a base resolution of 224×224 pixels and normalized to the $[0, 1]$ intensity range using min-max scaling. Images were converted to grayscale to preserve intensity-based anatomical information. Since all backbone networks employed in this study were pretrained on ImageNet and require three-channel RGB inputs, grayscale images were replicated across three channels to generate pseudo-RGB images with identical intensity values in each channel. Although InceptionV3 is originally designed for 299×299 inputs, all images were resized to a unified resolution of 224×224 to ensure consistency across backbone networks and reduce computational cost without significantly affecting transfer learning performance.

Finally, it should be noted that the dataset consists of 2D MRI slice images rather than patient-level volumetric data. Therefore, although strict train/validation/test separation was maintained, the splitting procedure was performed at the image level. Since patient identifiers are not available in the dataset, it is possible that adjacent slices originating from the same subject may appear across different subsets, potentially introducing a limited degree of inter-slice similarity. This is a known limitation in slice-based MRI classification studies.

Nevertheless, all data subsets were strictly isolated throughout the experimental pipeline. The test set was never used during training, validation, or model selection. All preprocessing operations, including normalization and data augmentation, were performed independently within each subset without sharing any statistical information. In addition, separate data generators were employed for training/validation and testing phases, ensuring complete isolation between subsets.

These precautions substantially minimize the risk of direct information leakage and support a fair evaluation of the model's generalization performance. This limitation is explicitly acknowledged in the interpretation of the results.

3.2. Backbone Networks and Transfer Learning Strategy

To exploit the representational capacity of deep convolutional neural networks pretrained on large-scale datasets, a two-stage transfer learning strategy was adopted using three widely employed architectures: ResNet50, DenseNet121, and InceptionV3. These backbone networks were selected due to their complementary architectural characteristics and proven effectiveness in large-scale image classification tasks. All models were initialized with ImageNet-pretrained weights and configured with `include_top = False` to remove their original classification layers.

For baseline backbone models, a customized classification head was appended, consisting of a Global Average Pooling (GAP) layer, a fully connected dense layer with 256 units and ReLU activation, a dropout layer with a rate of 0.5 for regularization, and a final dense layer with softmax activation for four-class tumor classification. This design enables effective dimensionality reduction while preserving discriminative global feature representations.

During the initial feature extraction stage, all convolutional layers of the pretrained backbone networks were frozen, allowing the models to function as fixed feature extractors while only the newly added classification head was trained. This warm-up phase improves training stability and reduces the risk of overfitting. In the subsequent fine-tuning stage, the convolutional layers of the backbone networks were unfrozen and optimized in an end-to-end manner using a reduced learning rate to adapt pretrained representations to domain-specific MRI characteristics. To stabilize training dynamics, batch normalization layers were kept frozen during fine-tuning.

All models were trained using the AdamW optimizer, with an initial learning rate of 1×10^{-4} during the warm-up phase and 1×10^{-5} during fine-tuning, and categorical cross-entropy loss. The training process was conducted in two stages: a warm-up phase of 10 epochs, during which only the classification head was trained while backbone layers remained frozen, followed by a fine-tuning phase of 20 epochs where most convolutional layers were unfrozen and optimized. Classification accuracy was selected as the primary evaluation metric.

Training was performed with an input resolution of 224×224 pixels and a batch size of 32. To improve robustness and reduce overfitting, data augmentation techniques, including random rotation ($\pm 20^\circ$), width and height shifts (up to 10%), zoom (up to 20%), and horizontal flipping, were applied exclusively to the training data generator. No augmentation was applied to validation or test sets.

To ensure training stability and prevent overfitting, early stopping, learning rate reduction on plateau, and model checkpointing strategies were employed throughout the training process. All experiments were implemented using TensorFlow/Keras and conducted on a workstation equipped with an NVIDIA RTX 3090 GPU (32 GB VRAM). The complete workflow of the transfer learning pipeline is illustrated in Figure 2.

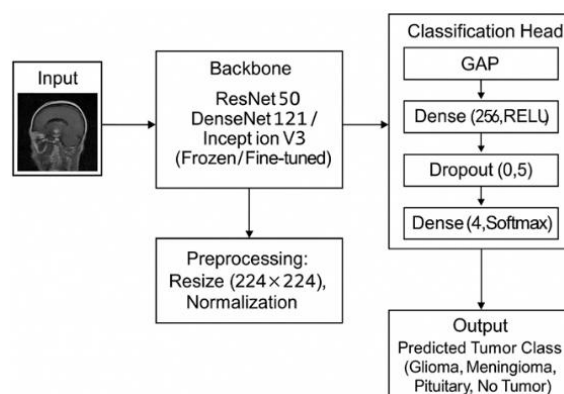


Figure 2. Transfer learning pipeline.

3.3. Backbone Architectures

To exploit complementary representation characteristics, three established convolutional neural network architectures—ResNet50, DenseNet121, and InceptionV3—are adopted as parallel backbone feature extractors within the proposed multi-stream framework. Instead of relying on a single architectural paradigm, this design leverages structural diversity to address the wide variation in tumor morphology, texture, and spatial distribution observed in brain MRI images.

Within the proposed architecture, ResNet50 serves as a stable semantic feature extractor by facilitating effective gradient propagation through residual connections during fine-tuning. Its moderate depth enables reliable learning of high-level contextual representations while maintaining computational efficiency, making it well suited for transfer learning-based medical imaging applications.

DenseNet121 contributes to the preservation of fine-grained spatial information through dense cross-layer connectivity. By encouraging feature reuse and continuous information flow, this backbone supports the retention of subtle textural patterns and boundary-related cues that are critical for accurate brain tumor discrimination. Such characteristics complement deeper semantic representations generated by other network streams.

To capture spatial features across multiple receptive field scales, InceptionV3 is integrated using its parallel convolutional branches and factorized filters. This configuration enhances sensitivity to heterogeneous tumor appearances and structural complexity, which are commonly encountered in brain MRI data. As part of the multi-stream design, InceptionV3 strengthens the model's ability to represent spatial variability across different tumor classes.

Collectively, the parallel integration of these heterogeneous backbones enables the extraction of diverse and complementary feature representations. This architectural synergy provides a robust foundation for subsequent attention-guided refinement and adaptive feature fusion, supporting reliable classification under conditions of inter-class similarity and morphological variability.

3.4. Attention-Enhanced Feature Extraction with CBAM

In this study, attention-based feature refinement is employed to enhance the discriminative capability of pretrained convolutional backbones within the proposed architecture. Rather than relying solely on deep convolutional representations, the model incorporates attention-driven modulation to selectively emphasize task-relevant activations observed in brain MRI images. This design choice is motivated by the observation that diagnostically meaningful tumor patterns are often subtle and spatially confined, making them vulnerable to dilution during deep feature aggregation.

To achieve this objective, the Convolutional Block Attention Module (CBAM) is integrated as a lightweight refinement component within the transfer learning pipeline. CBAM operates by adaptively reweighting feature responses along both channel and spatial dimensions, enabling the network to prioritize informative activations while suppressing redundant or background-dominated responses. Compared to more complex attention mechanisms, this strategy preserves computational efficiency, which is essential for large-scale medical image analysis.

Within the proposed framework, CBAM blocks are strategically inserted after the convolutional feature extraction stages of each pretrained backbone—ResNet50, DenseNet121, and InceptionV3—and before the global pooling layers. This placement allows attention-guided enhancement to be applied at a stage where high-level semantic features have already emerged, yet prior to feature compression. As a result, refined feature maps retain both spatial sensitivity and class-discriminative information during subsequent aggregation.

The effectiveness of attention-guided refinement has been previously demonstrated across various biomedical imaging tasks, where directing model focus toward clinically relevant regions has led to improved classification reliability (Ijaz et al., 2023; Dong et al., 2024). In the context of brain tumor classification, such refinement is particularly beneficial for addressing inter-class similarity and morphological variability, which commonly challenge conventional single-stream CNN architectures.

Overall, the integration of CBAM within the proposed model facilitates adaptive feature emphasis without imposing substantial computational overhead. By enabling selective amplification of diagnostically salient regions in MRI images, this attention-driven mechanism contributes to enhanced robustness and improved classification performance under heterogeneous imaging conditions.

3.5. Proposed IndereX-Net Architecture

Unlike baseline backbone models that employ Global Average Pooling (GAP), the proposed IndereX-Net replaces GAP with Generalized Mean (GeM) pooling to enable adaptive feature aggregation across heterogeneous feature streams. IndereX-Net is a multi-stream deep learning architecture designed for robust brain tumor classification from axial magnetic resonance imaging (MRI) scans. The model integrates three complementary convolutional neural network backbones—InceptionV3, DenseNet121, and ResNet50—arranged in parallel streams. This multi-stream design exploits architectural diversity and varying receptive field characteristics, facilitating the extraction of rich and complementary feature representations and improving generalization across diverse tumor subtypes. Each backbone stream is augmented with a Convolutional Block Attention Module (CBAM) to enhance discriminative feature learning. Through the sequential application of channel and spatial attention mechanisms, CBAM enables the network to emphasize diagnostically relevant regions while suppressing background noise and redundant activations. This attention-guided feature refinement is particularly suitable for brain MRI analysis, where tumor-related patterns often exhibit subtle intensity variations and complex spatial structures.

Following attention refinement, the output feature maps of each backbone are aggregated using Generalized Mean Pooling (GeM). GeM pooling provides a learnable and adaptive alternative to conventional global average and global max pooling strategies by allowing the network to adjust pooling behavior based on the underlying activation distribution. This adaptability enables more robust feature summarization across lesions with varying sizes, shapes, and intensity characteristics, which is critical for reliable brain tumor classification.

The feature representations extracted from the InceptionV3 and DenseNet121 streams are subsequently concatenated to form a unified embedding. To achieve adaptive cross-stream integration, a gated feature fusion mechanism is employed. Specifically, a gating vector generated from the ResNet50 stream dynamically modulates the concatenated feature representation through element-wise multiplication. This gating operation selectively amplifies informative features while suppressing redundant or noisy components, thereby enhancing representational efficiency and robustness.

The overall architecture is trained using a two-phase optimization strategy. During the initial warm-up phase, the backbone networks remain frozen, and only the attention modules, GeM pooling layers, gated fusion mechanism, and classification head are optimized in order to preserve pretrained low-level representations. In the subsequent fine-tuning phase, all layers except batch normalization layers are unfrozen and jointly optimized to adapt the network to domain-specific MRI characteristics while maintaining training stability.

4. Experimental Results

This section presents the experimental evaluation of the proposed IndereX-Net in comparison with baseline convolutional neural network architectures. All models were evaluated under three training configurations: frozen feature extraction, fine-tuning, and fine-tuning with CBAM integration. Performance was assessed using accuracy, loss, and ROC-AUC, as well as class-balanced metrics including macro- and weighted-averaged precision, recall, and F1-score.

As summarized in Table 1, frozen feature extraction yields limited discriminative capability, particularly for ResNet50. Fine-tuning consistently improves performance across all backbone architectures, highlighting the

importance of adapting pretrained representations to domain-specific MRI characteristics. In most cases, CBAM integration provides additional gains, indicating improved robustness, especially under class imbalance.

Table 1. Accuracy, loss, and ROC-AUC values of the evaluated models.

Models	Accuracy	Loss	ROC AUC
ResNet50 (Frozen)	0.6748	0.8771	0.8728
ResNet50 (Fine Tuning)	0.7548	0.5950	0.9328
ResNet50 (CBAM-Fine Tuning)	0.8123	0.4989	0.9571
DenseNet121 (Frozen)	0.8807	0.3097	0.9787
DenseNet121 (Fine Tuning)	0.9424	0.1691	0.9945
DenseNet121 (CBAM-Fine Tuning)	0.9250	0.2242	0.9909
InceptionV3 (Frozen)	0.8691	0.3287	0.9741
InceptionV3 (Fine Tuning)	0.9577	0.1482	0.9964
InceptionV3 (CBAM-Fine Tuning)	0.9615	0.1169	0.9969
IndereX-Net	0.9822	0.1201	0.9986

Among the baseline models, InceptionV3 and DenseNet121 achieve strong performance after fine-tuning, with further gains observed when attention mechanisms are applied. Notably, IndereX-Net outperforms all individual backbone models, achieving an accuracy of 98.22% and a ROC-AUC of 0.9986, which reflects superior overall classification capability.

Table 2. Macro- and weighted-averaged precision, recall, and F1-scores of the evaluated models.

Models	Macro Precision	Macro Recall	Macro F1	Weighted Precision	Weighted Recall	Weighted F1
ResNet50 (Frozen)	0.6588	0.6738	0.6427	0.6609	0.6748	0.6460
ResNet50 (Fine Tuning)	0.7387	0.7612	0.7394	0.7474	0.7548	0.7410
ResNet50 (CBAM-Fine Tuning)	0.8120	0.8228	0.8103	0.8182	0.8123	0.8072
DenseNet121 (Frozen)	0.8813	0.8903	0.8837	0.8842	0.8807	0.8801
DenseNet121 (Fine Tuning)	0.9421	0.9461	0.9435	0.9427	0.9424	0.9419
DenseNet121 (CBAM-Fine Tuning)	0.9235	0.9310	0.9259	0.9263	0.9250	0.9242
InceptionV3 (Frozen)	0.8693	0.8726	0.8705	0.8661	0.8691	0.8671
InceptionV3 (Fine Tuning)	0.9573	0.9611	0.9589	0.9578	0.9577	0.9574
InceptionV3 (CBAM-Fine Tuning)	0.9614	0.9629	0.9620	0.9613	0.9615	0.9612
IndereX-Net	0.9812	0.9840	0.9825	0.9824	0.9821	0.9822

To further assess class-balanced performance, Table 2 reports macro- and weighted-averaged precision, recall, and F1-scores. While fine-tuning substantially improves macro-level performance for all backbones, CBAM-enhanced configurations generally provide additional stabilization. IndereX-Net achieves the highest and most balanced results across both macro and weighted metrics, indicating robust performance across tumor categories.

To provide a compact comparison across class-balanced metrics, Figure 3 presents a heatmap visualization of macro- and weighted-averaged precision, recall, and F1-scores. The heatmap illustrates a clear performance progression from frozen to fine-tuned and CBAM-enhanced configurations. While InceptionV3 (CBAM-Fine-Tuned) and DenseNet121 (Fine-Tuned) demonstrate well-balanced results among baseline models, IndereX-Net achieves the highest and most consistent scores across these metrics, indicating robust and stable class-level performance.

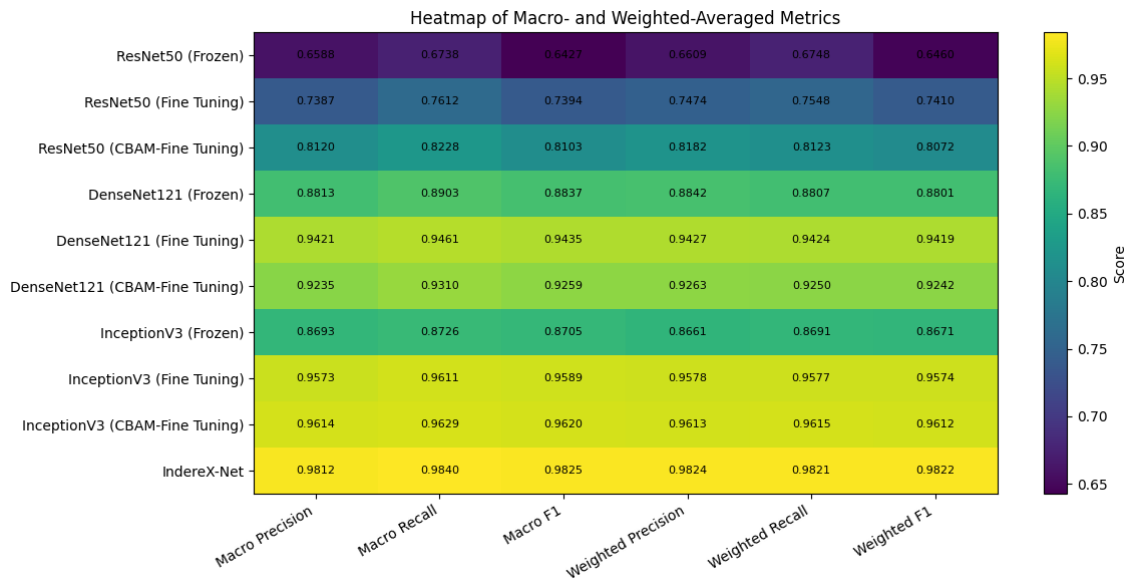


Figure 3. Heatmap visualization of macro- and weighted-averaged precision, recall, and F1-scores for all evaluated models under different training configurations.

4.1. Accuracy and Loss Analysis

A comparative analysis of accuracy and loss values is presented in Figure 4. Among the frozen configurations, ResNet50 exhibits the weakest performance, characterized by low accuracy and high loss values. Fine-tuning leads to noticeable improvements across all backbone architectures, while the integration of CBAM further reduces loss and stabilizes convergence.

DenseNet121 and InceptionV3 demonstrate relatively strong baseline performance even in frozen mode and benefit substantially from fine-tuning and attention enhancement. In contrast, IndereX-Net achieves the highest classification accuracy (98.22%) while maintaining one of the lowest loss values, indicating stable optimization behavior and improved generalization capability.

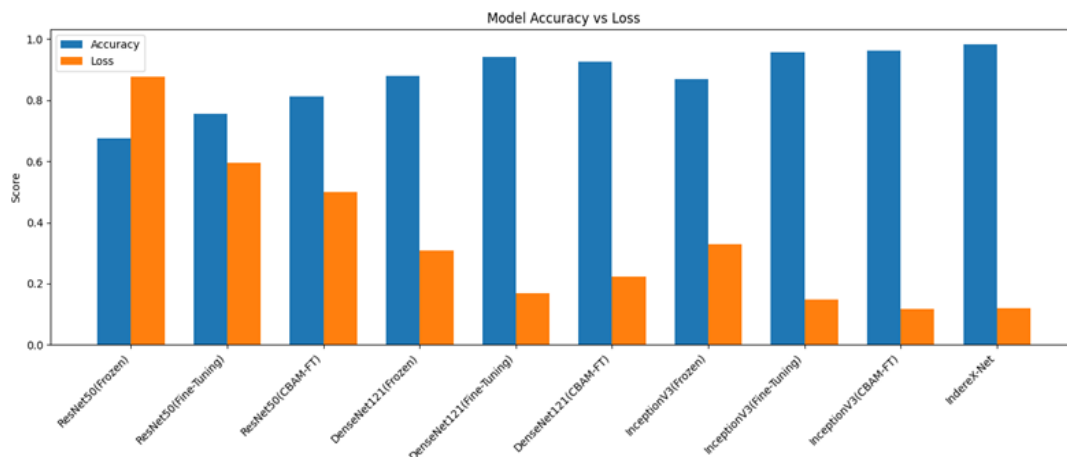


Figure 4. Accuracy and loss comparison of all evaluated models.

4.2. ROC-AUC Analysis

The ROC-AUC comparison of all evaluated models is presented in Figure 5. Overall, all architectures demonstrate strong discriminatory capability, with ROC-AUC values exceeding 0.87 across all training configurations. Among the frozen feature extraction settings, ResNet50 (Frozen) exhibits the lowest ROC-AUC score (0.8728), indicating limited effectiveness in separating tumor classes without domain-specific adaptation.

Fine-tuning leads to a substantial improvement in class separability across all backbone architectures, confirming the importance of adapting pretrained representations to brain MRI data. Additional gains are observed with the integration of CBAM attention modules. In particular, DenseNet121 (Fine-Tuned) and InceptionV3 (CBAM-Fine-Tuned) achieve high ROC-AUC scores of 0.9945 and 0.9969, respectively, highlighting the positive impact of both

fine-tuning and attention mechanisms on discriminative performance.

Most notably, the proposed IndereX-Net attains the highest ROC-AUC value of 0.9986, approaching near-perfect class discrimination. This result demonstrates strong robustness across tumor categories and supports the reliability of the proposed architecture for MRI-based brain tumor classification.

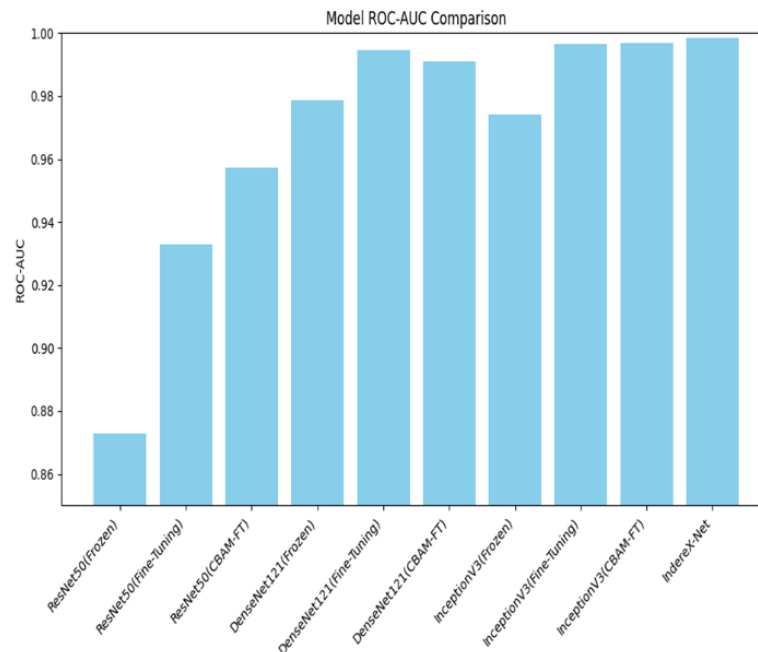


Figure 5. ROC-AUC comparison of all evaluated models.

4.3. Precision, Recall, and F1-Score Analysis

Class-balanced performance trends are illustrated in Figure 6 through macro- and weighted-averaged F1-scores. Frozen configurations, particularly ResNet50, exhibit limited macro-F1 performance, indicating sensitivity to class imbalance. Fine-tuning improves results across all architectures, while CBAM integration further stabilizes macro-level performance.

Among the baseline models, InceptionV3 (CBAM-Fine-Tuned) demonstrates strong and well-balanced results. In contrast, IndereX-Net achieves the highest macro- and weighted-averaged F1-scores (0.9825 and 0.9822, respectively), confirming robust class-balanced classification performance and improved reliability across all tumor categories.

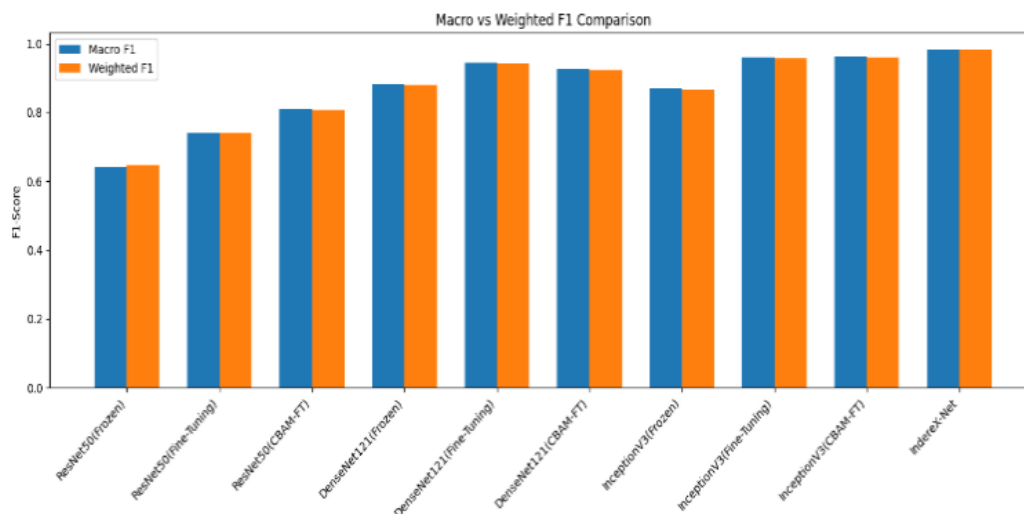


Figure 6. Comparison of macro- and weighted-averaged F1-scores for all evaluated models.

4.4. Confusion Matrix and Precision–Recall Curve Analysis

To analyze class-level behavior in greater detail, confusion matrices and precision–recall (PR) curves were examined. This complementary evaluation is particularly important in medical imaging applications, where class imbalance and false negatives can pose significant diagnostic risks.

Figure 7 presents the confusion matrices of nine baseline configurations, including ResNet50, DenseNet121, and InceptionV3 under frozen, fine-tuned, and CBAM-enhanced training regimes. A consistent progression is observed across all architectures. Frozen backbone models—most notably ResNet50 (Frozen)—exhibit pronounced misclassification between morphologically similar tumor classes such as glioma and meningioma, reflecting overlapping spatial and intensity characteristics in MRI. Fine-tuning substantially improves class separability, while CBAM integration further reduces inter-class confusion, particularly in the pituitary tumor category.

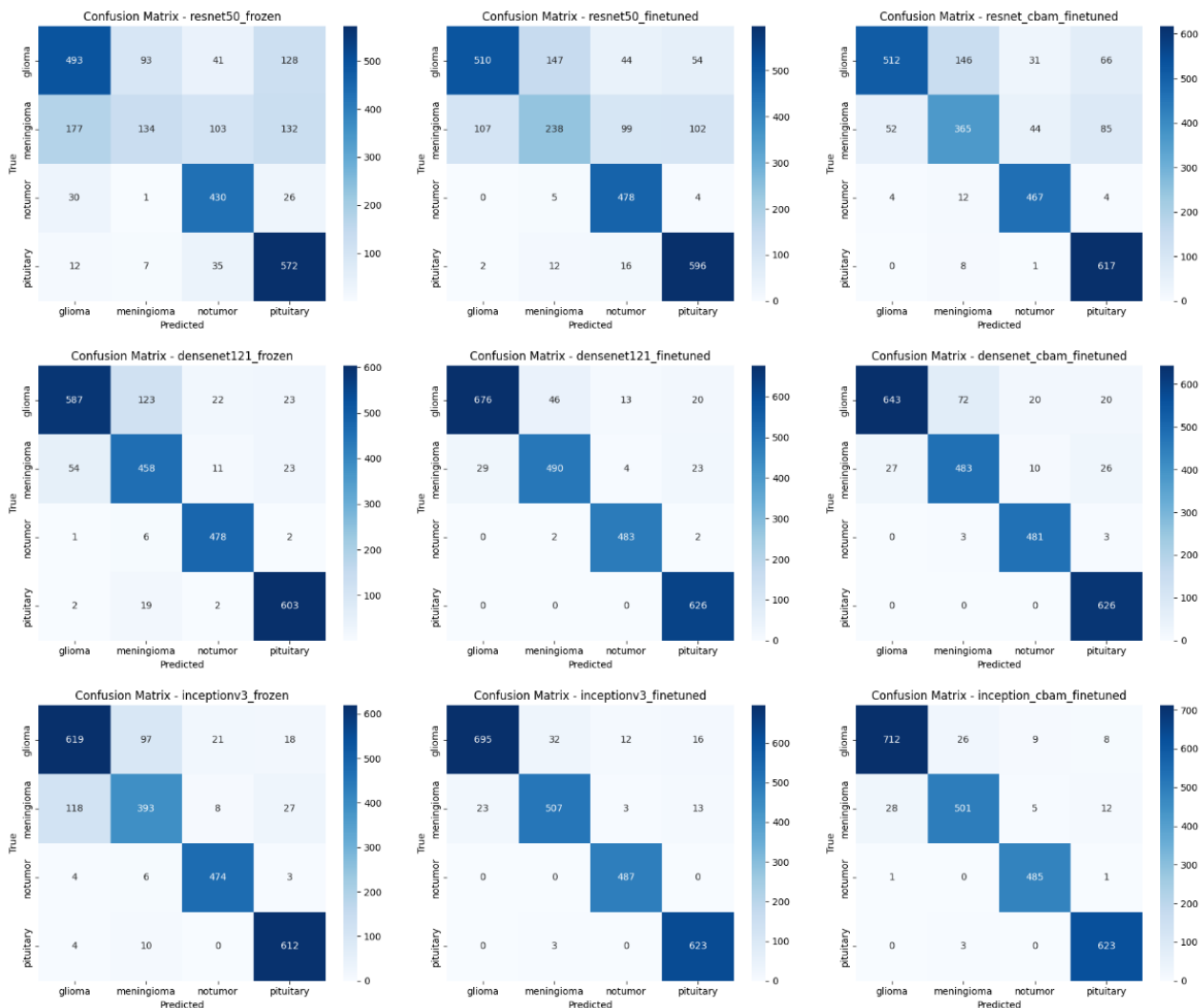


Figure 7. Confusion Matrices for 9 baseline models.

To emphasize the effectiveness of the proposed framework, Figure 8 illustrates the confusion matrix of IndereX-Net. The model demonstrates strong diagonal dominance, indicating high classification accuracy across all classes. A closer inspection reveals that the model achieves perfect classification for the “no tumor” class, with zero misclassifications. This is particularly important in clinical scenarios, as it minimizes the risk of incorrectly diagnosing healthy subjects. Most of the misclassifications occur between glioma and meningioma, as well as between meningioma and pituitary classes. This behavior aligns with clinical observations, as these tumor types often share similar visual characteristics in MRI images, making them difficult to distinguish. Importantly, the number of false negatives across tumor classes remains very low, indicating that the proposed model is highly reliable in detecting pathological cases. This is a critical requirement in real-world clinical applications, where missed tumor detection can have serious consequences.

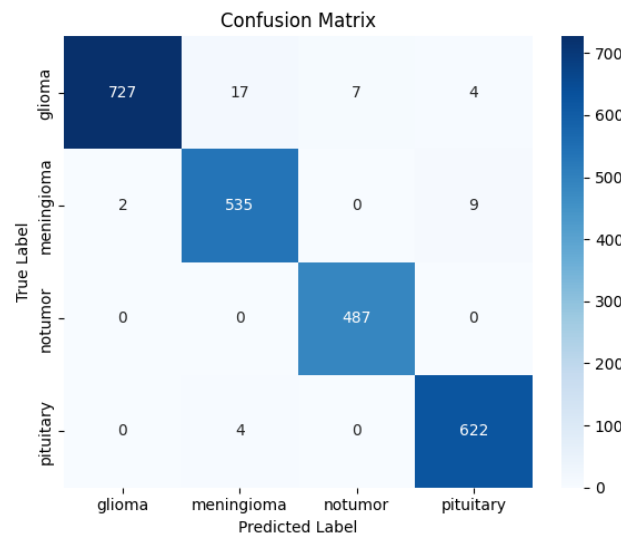


Figure 8. Confusion matrix of IndereX-Net.

Precision–recall curves provide a threshold-sensitive evaluation of model reliability. Unlike ROC curves, PR curves more accurately reflect performance under class imbalance. Figure 9 shows the PR curves of IndereX-Net, achieving near-rectangular shapes across all tumor categories with PR-AUC values consistently exceeding 0.99. This behavior indicates the model’s ability to maintain both high precision and high recall across a wide range of decision thresholds, which is critical for minimizing false negatives in clinical decision support systems.

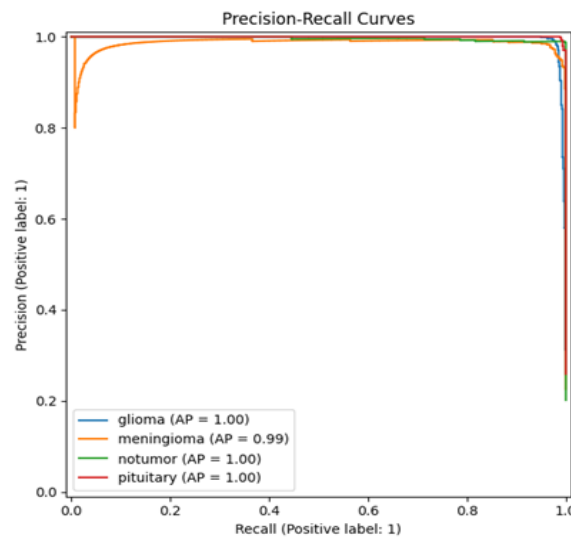


Figure 9. Precision–recall curves of IndereX-Net.

Overall, while fine-tuning and attention mechanisms improve individual backbone models, only the integrated, attention-refined, and adaptively fused architecture of IndereX-Net delivers consistently reliable, class-balanced, and threshold-robust performance suitable for clinical deployment.

4.5. Advanced Model Analysis

To further investigate the contribution of individual architectural components within the proposed IndereX-Net framework, a comprehensive ablation study was conducted. Specifically, the CBAM module, GeM pooling layer, and the adaptive feature fusion mechanism were systematically removed from the full architecture to evaluate their individual impact on model performance.

The quantitative results are summarized in Table 3. The complete IndereX-Net model achieved the highest classification accuracy (98.22%), demonstrating the effectiveness of the integrated architectural design.

The removal of the CBAM module resulted in a noticeable performance drop (95.17%), highlighting the importance of attention-guided feature refinement in emphasizing diagnostically relevant regions. Similarly, excluding the GeM pooling layer reduced the accuracy to 94.92%, indicating that adaptive pooling contributes

significantly to robust feature aggregation. When the feature fusion mechanism was removed, the model accuracy decreased to 95.34%, confirming the role of multi-stream feature integration in improving discriminative performance.

In addition to accuracy, Table 3 also presents the number of parameters and inference time for each variant. While simplified variants reduce computational complexity, they consistently underperform compared to the full architecture. This demonstrates that the proposed design achieves a favorable balance between predictive performance and computational efficiency.

Table 3. Ablation study results of the proposed IndereX-Net architecture.

Model Variant	Accuracy	Parameters	Inference Time (ms)
IndereX-Net (Full)	0.9822	55,709,709	115.96
w/o CBAM	0.9517	54,529,767	80.48
w/o GeM	0.9492	55,709,706	91.07
w/o Fusion	0.9534	30,283,562	69.42

Overall, the ablation results confirm that each major component of IndereX-Net contributes positively to the overall model performance, and that their combined use is essential for achieving optimal classification accuracy.

To further analyze the interpretability of the proposed model, Gradient-weighted Class Activation Mapping (Grad-CAM) was applied to visualize the regions contributing to the classification decisions.

Grad-CAM generates class-discriminative heatmaps by highlighting the most influential regions in the input MRI images. These visualizations enable qualitative assessment of whether the model focuses on clinically relevant tumor areas.

Figure 10 presents representative Grad-CAM results obtained from IndereX-Net. As observed, the model consistently concentrates on tumor regions while suppressing irrelevant background information. High activation intensities are localized around abnormal tissue structures, indicating that the model successfully captures meaningful spatial features associated with tumor presence.

This behavior demonstrates that the proposed architecture does not rely on spurious patterns but instead learns diagnostically relevant representations. Such interpretability is particularly important in medical applications, where transparency and trust are critical for clinical adoption.

Overall, the Grad-CAM analysis confirms that IndereX-Net provides not only high predictive performance but also clinically meaningful and explainable decision-making. The corresponding Grad-CAM visualizations demonstrating these activation patterns are presented in Figure 10.

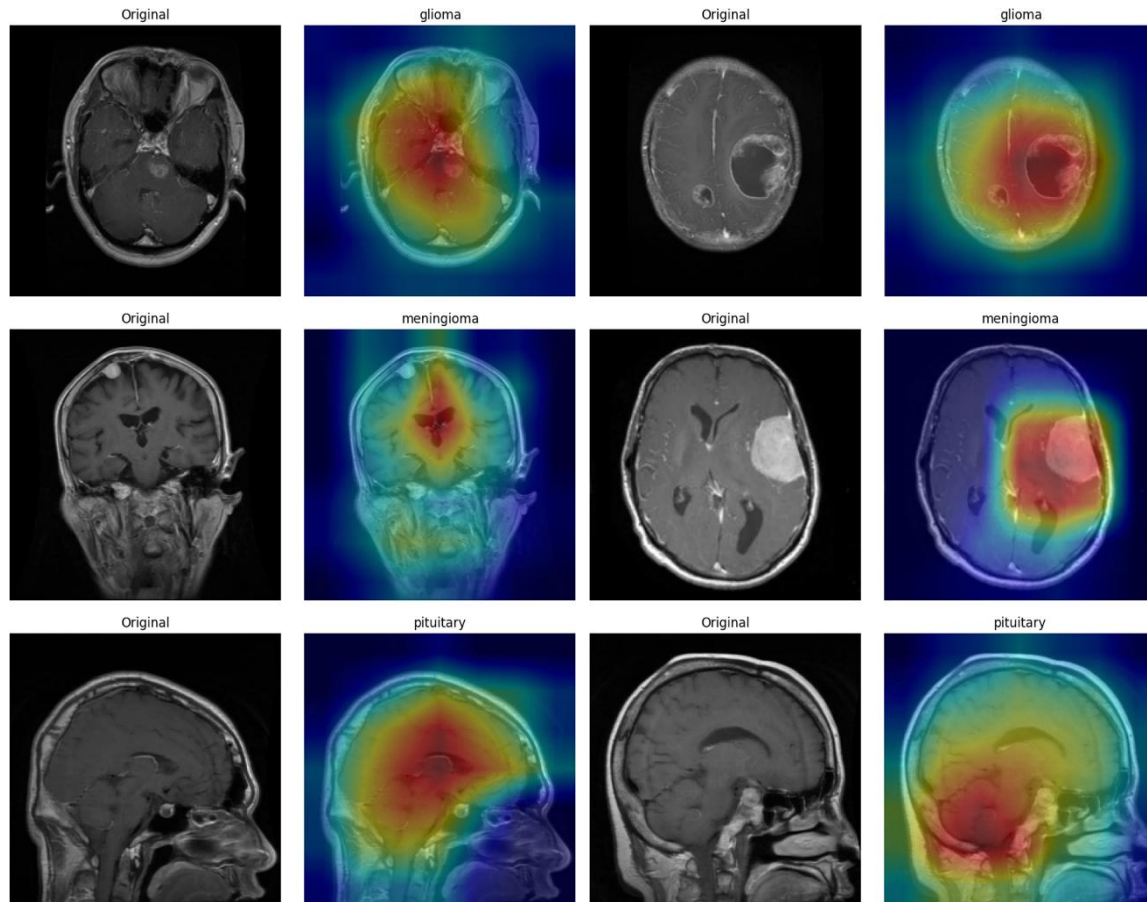


Figure 10. Grad-CAM visualizations of IndereX-Net highlighting tumor-relevant regions in MRI images.

To further assess the robustness of the proposed model, statistical reliability of the obtained results was considered. Although the evaluation was conducted on a fixed train-test split, the consistently high performance across multiple evaluation metrics (Accuracy, F1-score, ROC-AUC, and PR-AUC) indicates stable and reliable model behavior. In particular, the extremely high ROC-AUC (0.9986) and PR-AUC (0.9948) values suggest strong class separability and robustness against class imbalance. Additionally, the low number of misclassifications (43 samples) further supports the consistency of the model predictions. These results demonstrate that the performance improvements achieved by IndereX-Net are not due to random fluctuations but reflect the effectiveness of the proposed architecture.

To further evaluate the generalization capability of the proposed IndereX-Net model, an external validation experiment was conducted using an independent brain tumor MRI dataset obtained from Kaggle. This dataset is completely distinct from the Mendeley dataset used during training and consists of MRI images collected from multiple public sources, introducing variations in acquisition conditions, image resolution, and distribution characteristics. Such heterogeneity provides a more realistic and challenging evaluation scenario for assessing model robustness. For this purpose, only the predefined testing subset of the external dataset was utilized, comprising 1,600 images with a balanced class distribution (400 images per class) across four categories: glioma, meningioma, pituitary tumor, and no tumor. Importantly, no retraining or fine-tuning was performed on this dataset, ensuring that the evaluation strictly reflects the generalization ability of the trained model. The proposed IndereX-Net achieved an overall classification accuracy of 92.87% on the external dataset, accompanied by a macro-averaged F1-score of 0.9278 and a weighted F1-score of 0.9278. In addition, the model demonstrated strong discriminative performance with a ROC-AUC value of 0.9917 and a macro-averaged PR-AUC of 0.979. These results indicate that the model preserves a high level of predictive performance even when exposed to unseen data with distributional differences. A detailed analysis of the confusion matrix reveals that the model exhibits highly reliable performance in distinguishing pituitary tumor and no tumor classes, with nearly perfect classification and minimal misclassification. In particular, the pituitary tumor class achieved the highest classification consistency, suggesting that its structural characteristics are effectively captured by the multi-stream feature extraction and attention mechanisms. Similarly, the no tumor class demonstrates near-perfect separability, indicating that the model successfully learns discriminative features that differentiate pathological and non-pathological brain

structures. On the other hand, the majority of misclassifications are observed between glioma and meningioma classes. This behavior is consistent with clinical observations, as these tumor types often exhibit overlapping morphological and intensity characteristics in MRI scans, making them inherently more challenging to distinguish. Despite this challenge, the model still maintains a high level of performance in these classes, with the majority of predictions remaining correct and only a limited number of confusions occurring. Overall, the results of the external validation demonstrate that IndereX-Net is capable of learning robust and transferable feature representations. Although a moderate decrease in performance is observed compared to the internal test results, this drop is expected due to domain shift between datasets and does not significantly affect the model's reliability. The high accuracy, balanced F1-scores, and strong ROC-AUC and PR-AUC values collectively indicate that the proposed architecture maintains strong generalization performance across different data sources. These findings confirm that the integration of multi-stream convolutional backbones, attention-guided feature refinement, and adaptive feature fusion contributes to improved robustness against dataset variability. Consequently, the proposed model shows strong potential for real-world clinical applications, where variability in imaging conditions and patient populations is inevitable.

5. Result and Discussion

This study investigated the effectiveness of integrating task-specific fine-tuning, attention mechanisms, and multi-stream feature fusion for robust brain tumor classification from axial MRI scans. The experimental findings clearly demonstrate that models relying solely on frozen feature extractors exhibit limited discriminative capability, confirming that end-to-end optimization is essential for medical imaging tasks characterized by high inter-class similarity and subtle structural variations. The proposed IndereX-Net architecture achieved the strongest overall performance across all evaluated metrics by leveraging a parallel multi-stream design that integrates heterogeneous convolutional neural network backbones. The incorporation of CBAM-based attention mechanisms enabled spatially focused feature refinement, while Generalized Mean (GeM) pooling enhanced the aggregation of high-response regions. In addition, the adaptive gated fusion strategy allowed dynamic weighting of feature contributions across streams, leading to improved class separability and robustness against challenging tumor categories. This is further supported by confusion matrix and precision-recall analyses, which indicate stable and balanced performance across all classes.

Importantly, the external validation results on an independent dataset confirm that the proposed model maintains high predictive performance under domain shift conditions. Despite variations in data distribution and imaging characteristics, the model achieved strong accuracy and consistent class-wise performance, demonstrating that it is not overfitted to a single dataset and can generalize effectively across heterogeneous MRI data sources. This cross-dataset consistency highlights the robustness and practical applicability of the proposed approach in real-world clinical scenarios. From a clinical perspective, the model exhibits particularly strong performance in distinguishing pituitary tumor and non-tumor cases, while the majority of misclassifications occur between glioma and meningioma classes. This behavior is consistent with known clinical challenges, as these tumor types often present overlapping radiological features. Notably, the number of false negatives remains very low, which is critical in medical diagnosis where missed detections can lead to severe consequences. This indicates that the model prioritizes sensitivity in detecting pathological cases while maintaining high overall precision.

To further enhance interpretability, Grad-CAM-based visualizations were employed to analyze the decision-making process of the model. The resulting heatmaps demonstrate that IndereX-Net focuses on clinically relevant tumor regions rather than background structures, providing evidence that the model learns meaningful and localized representations. This explainability analysis supports the reliability of the model's predictions and increases its suitability for clinical decision support systems. Despite the strong performance gains, several limitations should be acknowledged. The primary experiments were conducted on two-dimensional MRI slices, which may not fully capture three-dimensional spatial context. Additionally, although the multi-stream architecture improves performance, it introduces increased computational complexity in terms of parameter count and inference time. While the current results indicate that inference speed remains within acceptable limits, further optimization may be required for deployment in resource-constrained environments. Furthermore, although external validation was performed, broader multi-center evaluations would strengthen the evidence for clinical generalizability.

Overall, the findings demonstrate that the proposed IndereX-Net architecture, which combines attention-guided feature refinement with adaptive multi-stream fusion, provides a robust and generalizable solution for brain tumor classification. Unlike conventional single-stream CNN models and static fusion approaches, the proposed method introduces a dynamic and attention-driven integration mechanism, making it a promising candidate for clinically applicable computer-aided diagnosis systems.

Conflict of Interest

No conflict of interest was declared by the authors.

References

- Agadi, K., Dominari, A., Tebha, S.S., Mohammadi, A., Zahid, S., 2023. Neurosurgical management of cerebrospinal tumors in the era of artificial intelligence: a scoping review. *Journal of Korean Neurosurgical Society*, 66 (6), 632–641.
- Ali, R.R., Yaacob, N.M., Alqaryouti, M.H., Sadeq, A.E., Doheir, M., Iqtait, M., Yaacob, S.S., 2025. Learning architecture for brain tumor classification based on deep convolutional neural network: Classic and ResNet50. *Diagnostics*, 15 (5), 624.
- Alrikabi, J., Muhammad, S., 2024. Deep feature extraction with SVM classifier for brain tumor classification. *Journal of Engineering and Physical Sciences*, 14 (3).
- Al-Saleh, A., Tejani, G.G., Mishra, S., Sharma, S.K., Mousavirad, S.J., 2025. A federated learning-based privacy-preserving image processing framework for brain tumor detection from CT scans. *Scientific Reports*, 15 (1), 23578.
- Asif, S., Zhao, M., Tang, F., Zhu, Y., 2023. An enhanced deep learning method for multi-class brain tumor classification using deep transfer learning. *Multimedia Tools and Applications*, 82 (20), 31709–31736.
- Bianchessi, T., Tampu, I.E., Blystad, I., Lundberg, P., Nyman, P., Eklund, A., Haj-Hosseini, N., 2023. Pediatric brain tumor type classification using deep learning on MR images from the Children's Brain Tumor Network. *medRxiv Preprint*.
- Cheng, D., Wang, Z., Li, J., 2023. Research on the application of CNN algorithm based on chaotic recursive diagonal model in medical image processing. *Applied Mathematics and Nonlinear Sciences*, 9 (1).
- Chen, W., Sun, P., Zhang, Y., 2024. A brain tumor diagnosis approach based on deep separable convolutional networks and attention mechanisms. *Applied Computational Engineering*, 67 (1), 60–69.
- Dong, R., Shiraiwa, A., Pawasut, A., Sreechun, K., Hayashi, T., 2024. Diagnosis of citrus greening using artificial intelligence: a Faster region-based convolutional neural network approach with convolution block attention module-integrated VGG-Net and ResNet models. *Plants*, 13 (12), 1631.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, 770–778.
- Hira, M.I.K., Hossain, M.S., Bithee, M.M.A., 2025. Brain tumor MRI dataset (glioma, meningioma, pituitary, no tumor). *Mendeley Data*, Version 1.
- Hossain, A., Islam, M.T., Abdul Rahim, S.K., Rahman, M.A., Rahman, T., Arshad, H., Chowdhury, M.E., 2023. A lightweight deep learning-based microwave brain image network model for brain tumor classification using reconstructed microwave brain images. *Biosensors*, 13 (2), 238.
- Hsu, S.P., Lin, M., Lin, C., Hsiao, T., Wang, Y., Sun, C., 2024. Brain tumor grading diagnosis using transfer learning based on optical coherence tomography. *Biomedical Optics Express*, 15 (4), 2343.
- Huang, G., Liu, S., Van der Maaten, L., Weinberger, K.Q., 2018. CondenseNet: An efficient DenseNet using learned group convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 2752–2761.
- Ijaz, A., Raza, B., Kiran, I., Waheed, A., Raza, A., Shah, H., Aftan, S., 2023. Modality specific CBAM-VGGNet model for the classification of breast histopathology images via transfer learning. *IEEE Access*, 11, 15750–15762.
- Janardhan, K., Thomas, C.B., 2025. Subclass classification of gliomas using MRI fusion technique. *SN Computer Science*, 6 (2), 1–15.
- Khan, F., Ayoub, S., Gulzar, Y., Majid, M., Reegu, F., Mir, M., Elwasila, O., 2023. MRI-based effective ensemble frameworks for predicting human brain tumor. *Journal of Imaging*, 9 (8), 163.
- Kochher, V., 2025. A comparative study of machine learning algorithms for brain tumor detection. *Journal of Recent Innovations in Science and Technology*, 1 (1), 1–9.
- Kumar, K., Jyoti, K., 2024. Explainable AI in brain tumor diagnosis: A critical review of machine learning and deep learning techniques. *Research Square Preprint*.
- Lu, Y., 2024. Multi-efficientnet ensemble-based brain tumor MRI images classification. *Proceedings of the International Conference on Computer Graphics, Artificial Intelligence and Data Processing (ICCAID 2023)*.
- Mastoi, Q.U.A., Latif, S., Brohi, S., Ahmad, J., Alqhatani, A., Alshehri, M.S., Ullah, R., 2025. Explainable artificial intelligence in medical imaging: an interpretable and collaborative federated learning model for brain tumor classification. *Frontiers in Oncology*, 15, 1535478.
- Mohan, P., Ramkumar, G., 2024. Detection and classification of brain tumor using fine tuned MobileNet algorithm. *Proceedings of the 3rd International Conference on Advanced Technologies (ICONAT)*, Goa, India, 1–5.
- Murugesan, G., Nagendran, P., Natarajan, J., 2025. Advancing brain tumor diagnosis: deep siamese convolutional neural network as a superior model for MRI classification. *Brain-X*, 3 (2), e70028.
- Pal, A., Seth, J., Tyagi, A., Haider, A., Mishra, A., 2024. An efficient brain tumor detection using ensemble learning. *Emerging Advances in Technology and Practice*.
- Sumona, R.B., Biswas, J.P., Shafkat, A., Rahman, M.M., 2025. An integrated deep learning approach for enhancing brain tumor diagnosis. *Healthcare Analytics*.
- Tariq, A., 2024. Transforming brain tumor diagnosis: vision transformers combined with ensemble techniques. *Journal of Population Therapeutics and Clinical Pharmacology*, 31, 1072–1084.
- Zhang, J., Tan, X., Chen, W., Du, G., Fu, Q., Zhang, H., Jiang, H., 2023. EFF_D_SVM: a robust multi-type brain tumor classification system. *Frontiers in Neuroscience*, 17, 1269100.