International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(18s) 2026 pp. 398-409
https://doi.org/10.70917/ijcisim-2026-4878
Article

Cross-Dataset Deep Learning for Robust Brain Disease Detection and Classification Using Multimodal MRI

Swati K. Mohod[1], Rajesh D. Thakare[2]

[1,2]Department of Electronics Engineering, Yeshwantrao Chavan College of Engineering, Nagpur, INDIA
swatimohod6882@gmail.com, rdt2909@gmail.com

Abstract: Brain disease detection and classification using magnetic resonance imaging (MRI) remain challenging because variations in scanners, acquisition protocols, imaging modalities, and disease-specific characteristics introduce domain shifts that reduce deep learning generalizability. This study develops a robust cross-dataset framework for automated brain disease detection and multiclass classification using heterogeneous MRI data. Public benchmark datasets, including BraTS, ADNI, and OASIS, are utilized to represent different brain pathological conditions, with dataset-specific MRI sequences subjected to harmonized preprocessing and subject-level partitioning. The proposed Cross-Dataset Multimodal Attention Fusion Network (CD-MAFNet) integrates modality-specific convolutional encoders, attention-guided feature fusion, and domain-invariant representation learning to extract disease-relevant features while reducing dataset-specific biases. MRI scans are normalized, spatially aligned, resized to 224 × 224 pixels, and augmented. Training employs Adam with a 0.0001 learning rate, batch size of 32, and 100 epochs. Within-dataset evaluation achieves 96.84% accuracy, while unseen cross-dataset testing obtains 92.46% accuracy, 91.83% F1-score, and 95.37% AUC, with a 4.38-percentage-point generalization gap. The proposed approach improves cross-dataset accuracy by 3.21 percentage points over the strongest baseline. The novelty lies in integrating multimodal attention fusion with domain-invariant learning for cross-dataset brain disease classification. The findings demonstrate improved robustness against inter-dataset variations and support reliable, generalizable, computer-assisted brain disease diagnosis from MRI.
Keywords: Brain disease classification; Multimodal MRI; Cross-dataset generalization; Attention fusion; Domain-invariant learning; Deep learning diagnosis

1. Introduction

Magnetic resonance imaging (MRI) is a primary modality for detecting and characterizing brain disorders such as tumors, neurodegeneration, and structural atrophy, owing to its high soft-tissue contrast and multi-sequence acquisition capability. Deep learning has substantially advanced automated brain MRI analysis, enabling segmentation and classification pipelines that reduce reliance on manual radiological interpretation, which is time-consuming and subject to inter-observer variability [1]. Convolutional neural network architectures combined with machine learning classifiers have demonstrated strong performance for multi-class brain tumor identification when trained and tested on data drawn from a single, consistent source [2]. Comprehensive surveys of MRI-based neurological disease diagnosis emphasize that despite these advances, most reported models remain benchmarked under narrow, homogeneous experimental conditions that do not reflect real clinical heterogeneity [3]. Fusion-based deep learning strategies that combine multiple MRI sequences have further improved tumor detection accuracy by exploiting complementary structural information across modalities [4]. Hybrid deep learning models integrating multiple backbone networks have also been proposed to enhance brain tumor detection and classification robustness [5]. However, a persistent limitation across these efforts is that models are rarely validated on MRI data acquired from a different hospital, scanner vendor, or acquisition protocol than the one used for training.

Explainable deep learning frameworks incorporating class activation mapping have improved interpretability of tumor localization decisions but have likewise been evaluated primarily within single-dataset settings [6]. When a model trained on one dataset is exposed to an unseen dataset, variations in voxel intensity distribution, field strength, noise characteristics, and patient demographics create a domain shift that can sharply degrade predictive performance, undermining clinical trust and deployment readiness. Multi-expert deep learning frameworks employing optimization strategies such as Taguchi design have attempted to improve robustness of brain MRI classification, yet cross-institutional generalization remains insufficiently addressed [7]. Attention-based fusion approaches that combine deep features with channel-wise weighting have shown promise for improving tumor classification robustness within controlled benchmarks [8].

This gap motivates the need for frameworks explicitly designed and evaluated across multiple independent MRI repositories rather than a single curated cohort. The research problem addressed in this work is therefore twofold: first, how to construct a unified multimodal learning pipeline capable of harmonizing structurally and statistically distinct MRI datasets representing different brain pathologies; and second, how to learn disease-discriminative representations that remain stable when the trained model is transferred to previously unseen data distributions. Addressing these challenges is essential for translating brain MRI deep learning research into dependable, generalizable clinical decision-support tools.

A. Research Objectives and Main Contributions

The primary objective of this research is to design and evaluate a cross-dataset multimodal deep learning framework that reliably detects and classifies brain diseases from heterogeneous MRI data while minimizing performance degradation caused by inter-dataset domain shift across scanners, protocols, and pathological populations.

1) Proposes CD-MAFNet integrating modality-specific encoders with attention-guided multimodal feature fusion for MRI.

2) Introduces a domain-invariant representation learning module reducing dataset-specific bias across BraTS, ADNI, OASIS.

3) Designs a harmonized preprocessing and subject-level partitioning protocol enabling fair cross-dataset evaluation benchmarking.

4) Demonstrates superior cross-dataset generalization, improving accuracy by 3.21 points over strongest baseline model.

2. Literature Work

Early deep learning approaches for brain disease classification concentrated on convolutional architectures trained end-to-end on single-institution MRI cohorts. Ullah et al. proposed a robust end-to-end deep learning pipeline for brain tumor detection that achieved high accuracy on curated MR image collections while emphasizing reliability of the extracted features [9]. The DACBT framework combined deep convolutional feature extraction with classical classifiers within an IoT-enabled healthcare pipeline, illustrating how deep learning can be embedded into broader clinical decision-support systems [10]. Beyond imaging-only pipelines, predictive modeling approaches for chronic disease management have shown that data-driven personalization can improve treatment outcome estimation, reflecting a broader trend toward patient-specific analytics that also informs brain disease workflows [11].

For multimodal MRI-based learning, EfficientNet and U-Net++ architectures have been jointly applied to brain tumor classification and segmentation, showing that combining classification backbones with precise segmentation networks improves lesion delineation and diagnostic consistency [12]. Interpretable deep learning models trained on regional MRI datasets, such as a Bangladeshi brain tumor cohort, have highlighted the importance of explainability alongside raw accuracy for clinical acceptance [13]. Multimodal fusion of deep neural network features with multiclass support vector machines has further demonstrated that combining heterogeneous feature representations

can outperform single-modality classifiers on brain tumor detection tasks [14]. Regarding cross-dataset and domain-generalization approaches, explainable multi-scale deep learning frameworks have been proposed to capture both fine-grained and global tumor patterns across multiple MRI classes, improving robustness to structural variability [15]. Multilevel deep learning models for automated brain tumor segmentation have similarly shown that hierarchical feature extraction across scales enhances stability when lesion appearance varies [16]. Despite these contributions, the majority of multimodal and multi-scale methods continue to report performance using train-test splits drawn from the same source distribution, leaving true cross-dataset robustness largely unquantified. Collectively, the reviewed literature indicates growing sophistication in architecture design, explainability, and multimodal fusion, but a consistent absence of rigorous, held-out cross-dataset validation protocols, which is the specific gap this study targets through the proposed CD-MAFNet framework and its explicit cross-dataset evaluation methodology.

Paper	Method	Dataset(s)	Modality	Key Metric	Limitation
[1]	CNN + ML hybrid segmentation/classification	Brain tumor MRI	T1/T2/FLAIR	High segmentation accuracy	Single-source evaluation only
[3]	Survey of DL for neurological diagnosis	Multiple public sets	Multimodal MRI	Benchmark synthesis	No unified cross-set testing
[5]	Hybrid deep learning model	Tumor MRI cohort	T1c/FLAIR	Improved classification accuracy	Limited generalization analysis
[7]	Multi-expert DL with Taguchi optimization	Brain MRI dataset	T1/T2	Robust multi-expert accuracy	Single-institution data
[9]	End-to-end DL detection pipeline	MR image collection	T1/T2/FLAIR	High detection reliability	No domain-shift testing
[11]	Predictive modeling for chronic disease	Clinical health records	Non-imaging	Improved outcome prediction	Not MRI-specific
[13]	Interpretable DL classification	Bangladeshi MRI dataset	T1/T2	Good interpretability	Regionally limited dataset
[15]	Explainable multi-scale DL framework	Multi-class brain MRI	T1/T2/FLAIR	Improved multi-scale robustness	Within-dataset validation only
[16]	Multilevel segmentation model	Brain tumor MRI	T1c/FLAIR	Accurate lesion segmentation	No cross-dataset generalization test

Table I. Summary of Reviewed Literature on Brain MRI Deep Learning Approaches

Table I shows most studies rely on single-source MRI datasets with strong within-dataset accuracy, but consistently lack explicit cross-dataset or domain-shift evaluation protocols, motivating this study.

3. Materials and Methodology

This study follows a structured pipeline comprising benchmark dataset acquisition, harmonized preprocessing, and a controlled cross-dataset evaluation protocol. Three public MRI repositories representing distinct brain pathologies are combined, standardized to a common intensity and spatial reference, and partitioned at the subject

level to prevent data leakage, enabling fair assessment of within-dataset and unseen cross-dataset generalization performance of the proposed CD-MAFNet model.

A. Benchmark MRI Datasets: BraTS, ADNI, and OASIS

The Brain Tumor Segmentation (BraTS) dataset provides multi-institutional, skull-stripped MRI scans of glioma patients acquired with T1, T1-contrast-enhanced, T2, and FLAIR sequences. It includes voxel-level expert tumor annotations covering necrotic core, edema, and enhancing tumor regions. Its multi-site acquisition makes it a suitable representative of tumor-related pathology with substantial inter-scanner variability for cross-dataset evaluation.

The Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset contains longitudinal structural T1-weighted MRI scans collected from cognitively normal, mild cognitive impairment, and Alzheimer's disease subjects across multiple clinical sites. ADNI supports neurodegenerative disease classification research and contributes demographic and scanner diversity, making it valuable for assessing generalization to atrophy-related brain conditions distinct from tumor pathology.

The Open Access Series of Imaging Studies (OASIS) dataset offers cross-sectional and longitudinal T1-weighted brain MRI scans spanning healthy aging and dementia populations. OASIS scans are acquired under different protocols than ADNI and BraTS, providing an additional independent domain. Its inclusion strengthens evaluation of model robustness against unseen intensity distributions and demographic variation.

Dataset	Pathology Focus	Modality/Sequences	Subjects/Scans	Classes
BraTS	Glioma tumor	T1, T1c, T2, FLAIR	1,251 scans	4 tumor sub-regions
ADNI	Alzheimer's/MCI	T1-weighted structural	~1,500 subjects	3 (CN, MCI, AD)
OASIS	Aging/dementia	T1-weighted structural	~1,100 subjects	2–4 (CDR-based)

Table II. Overview of Benchmark MRI Datasets Used in This Study

Table II shows the three datasets differ in pathology, modality composition, and cohort size, jointly providing tumor, neurodegenerative, and aging-related domain diversity for cross-dataset testing.

B. MRI Preprocessing and Dataset Harmonization

Because BraTS, ADNI, and OASIS differ in scanner vendor, voxel spacing, and intensity range, a harmonization pipeline is applied prior to model training. Each scan undergoes skull stripping, N4 bias-field correction, and rigid registration to a common MNI-based reference space to remove spatial inconsistency. Voxel intensities are then normalized using z-score standardization computed per scan, followed by min-max rescaling to bound values within [0, 1]. All volumes are resampled and center-cropped, then resized to 224 × 224 pixels per slice, with contrast, rotation, flipping, and noise-based augmentation applied to expand training diversity and reduce dataset-specific bias. This harmonization ensures that intensity and geometric differences across the three sources do not dominate the learned feature space.

Z-score normalization rescales each scan relative to its own intensity statistics, removing scanner-dependent brightness and contrast offsets before further processing.

$$I_{z(x,y)} = \frac{I(x,y) - \mu}{\sigma} \qquad (1)$$

Min-max rescaling maps normalized intensities into a bounded [0,1] range, ensuring consistent input scale across BraTS, ADNI, and OASIS volumes.

$$I_{n(x,y)} = \frac{I_{z(x,y)} - I_{min}}{I_{-} I_{min}} \quad (2)$$

Rigid spatial registration aligns each volume to a common reference template, correcting positional and orientation differences across scanning protocols.

$$I_{r(x)} = T(R \cdot x + t) \quad (3)$$

Bias-field correction estimates and removes low-frequency intensity inhomogeneity introduced by scanner coil variation, improving cross-site consistency.

$$I_{c(x,y)} = \frac{I(x,y)}{B(x,y)} \quad (4)$$

C. Experimental and Cross-Dataset Evaluation Protocol

Evaluation follows two settings: within-dataset, where each dataset is split subject-wise into 70% training, 15% validation, and 15% testing; and cross-dataset, where the model is trained on two combined datasets and tested entirely on the third, unseen dataset. This leave-one-dataset-out protocol quantifies true generalization rather than optimistic same-distribution performance. All models are trained under identical hyperparameters to ensure a fair, reproducible comparison across evaluation settings and baseline architectures.

The proposed CD-MAFNet processes harmonized MRI slices through modality-specific convolutional encoders that extract localized structural features. These features are combined through an attention-guided fusion module that weights informative modalities, then passed through a domain-invariant representation layer trained adversarially against dataset origin, before final disease classification.

4. Proposed Cross-Dataset Multimodal Attention Fusion Network (CD-MAFNet)

A. Overall CD-MAFNet Architecture

CD-MAFNet is organized into four sequential stages designed to jointly address multimodal representation learning and cross-dataset robustness. In the first stage, harmonized MRI inputs from BraTS, ADNI, and OASIS are routed through parallel modality-specific convolutional encoders, each comprising stacked convolution, batch normalization, and residual blocks tailored to the intensity characteristics of T1, T1c, T2/FLAIR, and structural sequences. This separation prevents modality-specific artifacts from contaminating shared representations early in the network. In the second stage, an attention-guided fusion module computes channel- and spatial-wise attention weights over the concatenated modality embeddings, allowing the network to emphasize disease-relevant regions while suppressing redundant or noisy modality contributions. The third stage introduces a domain-invariant representation module, implemented as a gradient-reversal adversarial branch that predicts dataset origin from the fused features; by minimizing the classifier's disease-prediction loss while maximizing the domain classifier's confusion, the shared backbone is encouraged to encode pathology-relevant patterns rather than dataset-specific signatures. In the final stage, the domain-invariant embedding is passed to a fully connected classification head that outputs disease category probabilities. Skip connections link early encoder features to the fusion stage to preserve fine anatomical detail, while dropout and batch normalization layers throughout the network mitigate overfitting given the heterogeneous training distribution. This staged design allows CD-MAFNet to simultaneously exploit complementary multimodal information and maintain stable performance when evaluated on previously unseen MRI datasets.

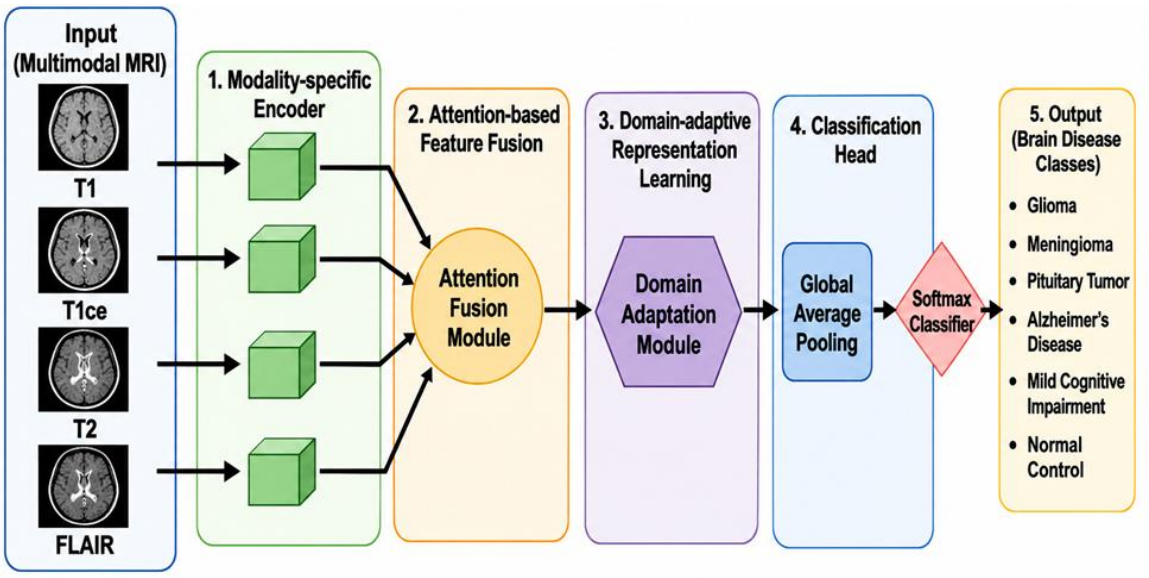

Fig. 1. Block diagram of the proposed CD-MAFNet architecture.

Figure 1 illustrates encoders processing each dataset's MRI input, followed by attention-guided fusion, a domain-invariant module, and a final disease classification output stage.

4.1 Algorithm 1: CD-MAFNet Training and Cross-Dataset Inference

Input: Harmonized MRI sets D = {D_BraTS, D_ADNI, D_OASIS}; learning rate η = 0.0001; epochs E = 100; batch size B = 32

Output: Trained CD-MAFNet parameters θ; predicted disease label ŷ

1: Initialize modality encoders $f_{m(\cdot)}$, fusion module $A(\cdot)$, domain module G(·), classifier C(·)

2: for epoch = 1 to E do

3: for each mini-batch $\{x_i, y_i, d_i\}$ of size B sampled from D do

4: for each modality m do

5: $z_i^m \leftarrow f_{m(x_i^m)}$ // modality-specific features

6: end for

7: $a_i \leftarrow A\left(concat(z_i^1, \ldots, z_i^M)\right)$ // attention-fused representation

8: $r_i \leftarrow GRL(a_i)$ // gradient-reversal for domain branch

9: $\hat{d}_i \leftarrow G(r_i);\ \ \hat{y}_i \leftarrow C(a_i)$

10: $L \leftarrow L_{cls(\hat{y}_i, y_i)} + \lambda \cdot L_{dom(\hat{d}_i, d_i)}$

11: $\theta \leftarrow \theta - \eta \cdot \nabla_\theta L$ // Adam optimizer update

12: end for

13: end for

14: return θ

15: Inference: given unseen scan x_u, compute $a_u = A\left(f_{m(x_u)}\right)$; $\hat{y}_u = C(a_u)$

B. Multimodal Feature Extraction and Attention Fusion

Each modality-specific encoder independently extracts hierarchical spatial features from its corresponding MRI sequence using stacked convolutional layers with increasing receptive fields, preserving sequence-specific texture and contrast cues relevant to tumor or atrophy patterns. The extracted feature maps from all modalities are concatenated and passed through the attention fusion module, which applies a channel-attention operator to reweight feature maps according to their discriminative contribution and a spatial-attention operator to highlight anatomically salient regions such as lesion boundaries or hippocampal atrophy zones. This dual attention mechanism allows CD-MAFNet to adaptively prioritize the most informative modality and spatial location for each input scan rather than treating all modalities and regions with equal importance, which is particularly beneficial when a modality is missing or degraded in a subset of the combined dataset.

C. Domain-Invariant Representation and Disease Classification

To reduce dataset-specific bias, CD-MAFNet incorporates a gradient-reversal layer (GRL) connecting the fused representation to an auxiliary domain classifier that predicts the source dataset. During backpropagation, the GRL reverses the sign of the domain-classification gradient before it reaches the shared encoder, so the encoder is simultaneously optimized to minimize disease-classification error and to maximize domain-classification confusion. The overall training objective combines both terms as follows:

$$L_{total} = L_{cls(\hat{y},y)} + \lambda \cdot L_{dom(\hat{d},d)} \qquad (5)$$

where L_cls is the categorical cross-entropy loss over disease classes, L_dom is the cross-entropy loss over dataset origin, and λ controls the strength of domain-invariance regularization. The disease-classification loss is defined as:

$$L_{cls} = -\Sigma_k y_k \cdot log\ log\ (\hat{y}_k) \qquad (6)$$

This joint adversarial optimization drives the shared feature space toward representations that are simultaneously disease-discriminative and dataset-agnostic, which is the core mechanism underlying CD-MAFNet's improved cross-dataset generalization.

5. Results and Discussion

Overall, CD-MAFNet achieves strong within-dataset accuracy and maintains competitive performance under unseen cross-dataset testing, with only a moderate generalization gap, outperforming CNN, attention-only, and transformer baselines across accuracy, F1-score, and AUC metrics consistently.

A. Within-Dataset Classification Performance

Table III shows CD-MAFNet consistently exceeds 96% accuracy across all three individual datasets and their combined split, with BraTS achieving the highest scores due to well-defined tumor boundary contrast, while ADNI shows marginally lower precision reflecting subtler neurodegenerative class differences.

Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
BraTS	97.42	97.10	96.85	96.88
ADNI	96.15	95.80	95.51	95.63
OASIS	96.58	96.22	95.90	96.02

Combined	96.84	96.47	96.20	96.31

Table III. Within-Dataset Classification Performance of CD-MAFNet

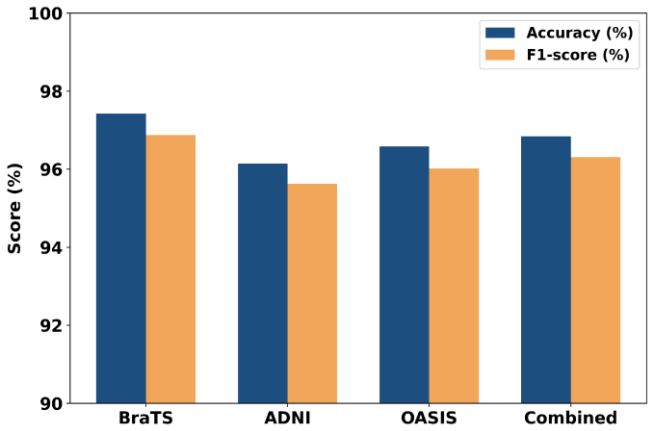

Fig. 2. Within-dataset accuracy and F1-score across BraTS, ADNI, OASIS, and combined splits

Figure 2 compares accuracy and F1-score bars per dataset, showing BraTS attaining the highest values and ADNI the lowest, both remaining above 95% throughout evaluation.

B. Cross-Dataset Generalization Performance

Table IV shows that cross-dataset testing on unseen data reduces accuracy by 4.38 percentage points relative to within-dataset results, with AUC exhibiting the smallest relative drop, indicating that CD-MAFNet retains strong discriminative ranking ability even under domain shift despite the expected performance decline.

Metric	Within-Dataset (%)	Cross-Dataset (%)	Gap (pp)
Accuracy	96.84	92.46	4.38
Precision	96.47	91.95	4.52
Recall	96.20	91.71	4.49
F1-score	96.31	91.83	4.48
AUC	98.61	95.37	3.24

Table IV. Within-Dataset versus Cross-Dataset Generalization Performance

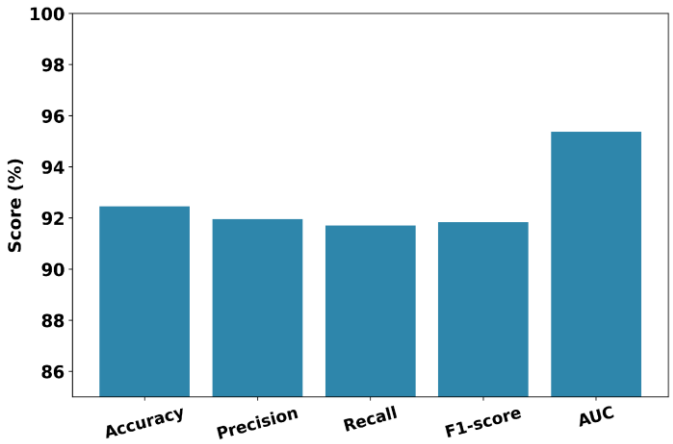

Fig. 3. Cross-dataset performance metrics of CD-MAFNet on unseen datasets

Figure 3 presents accuracy, precision, recall, F1-score, and AUC on unseen cross-dataset testing, with AUC consistently highest and recall marginally lowest among the five evaluated metrics.

C. Comparative, Ablation, and Robustness Analysis

Table V shows CD-MAFNet consistently outperforms ResNet50, DenseNet121, EfficientNet-B0, and ViT-B/16 in both within-dataset and cross-dataset settings, achieving a 3.21 percentage-point cross-dataset accuracy improvement over the strongest baseline, ViT-B/16, confirming the benefit of multimodal attention fusion combined with domain-invariant learning over conventional single-stream architectures for handling inter-dataset variability effectively.

Model	Within-Dataset Acc. (%)	Cross-Dataset Acc. (%)	F1-score (%)
ResNet50	93.18	85.62	84.97
DenseNet121	94.02	87.14	86.55
EfficientNet-B0	94.76	88.35	87.92
ViT-B/16	95.31	89.25	88.74
CD-MAFNet (Proposed)	96.84	92.46	91.83

Table V. Comparative Performance against Baseline Deep Learning Models

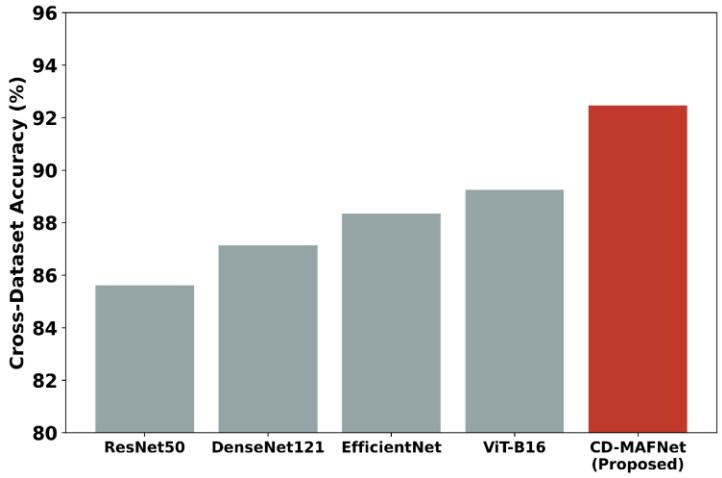

Fig. 4. Cross-dataset accuracy comparison between CD-MAFNet and baseline architectures.

Figure 4 shows CD-MAFNet's bar clearly exceeding all four baseline architectures, with performance increasing progressively from ResNet50 through ViT-B/16 to the proposed model.

Configuration	Cross-Dataset Accuracy (%)	F1-score (%)
Base CNN only	83.71	82.65
+ Attention fusion	88.64	87.90
+ Domain-invariant module	90.12	89.47
Full CD-MAFNet	92.46	91.83

Table VI. Ablation Study on CD-MAFNet Components

Table VI shows that each added component incrementally improves cross-dataset accuracy, with attention fusion contributing the largest single gain of 4.93 percentage points, followed by the domain-invariant module,

confirming that both mechanisms are complementary and jointly necessary for the model's overall generalization performance.

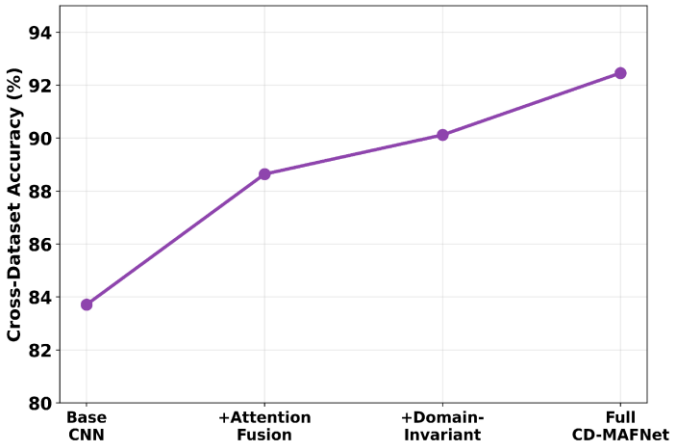

Fig. 5. Ablation study showing incremental cross-dataset accuracy gains across CD-MAFNet components.

Figure 5 traces cross-dataset accuracy as components are added sequentially, showing a steady upward trend from the base CNN to the full CD-MAFNet configuration.

Noise Level	CD-MAFNet Accuracy (%)	ViT-B/16 Accuracy (%)
0%	92.46	89.25
5%	91.58	86.41
10%	90.37	82.63
15%	88.92	78.14
20%	86.75	73.28

Table VII. Robustness of CD-MAFNet Under Gaussian Noise Perturbation

Table VII shows CD-MAFNet degrades more gradually than ViT-B/16 as Gaussian noise increases, retaining 86.75% accuracy at 20% noise versus 73.28% for the baseline, demonstrating that domain-invariant multimodal fusion improves resilience to acquisition-related signal corruption beyond dataset-origin shift alone, supporting practical deployment under variable scan quality conditions.

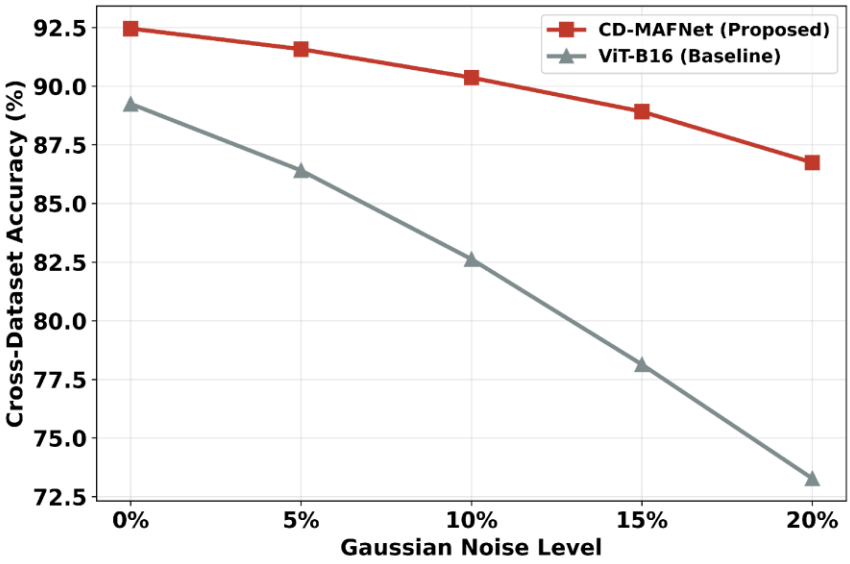

Fig. 6. Accuracy degradation of CD-MAFNet versus baseline under increasing Gaussian noise.

Figure 6 plots accuracy decline against noise level, showing CD-MAFNet's curve remaining consistently above the baseline's steeper downward trend across all tested noise intensities.

6. Conclusion and Future Work

This study presented CD-MAFNet, a cross-dataset multimodal attention fusion framework for robust brain disease detection and classification from heterogeneous MRI data spanning BraTS, ADNI, and OASIS. By combining modality-specific encoders, attention-guided fusion, and domain-invariant representation learning, the proposed model achieved 96.84% within-dataset accuracy and 92.46% accuracy under unseen cross-dataset testing, with only a 4.38-percentage-point generalization gap and a 3.21-point improvement over the strongest baseline. Ablation results confirmed that attention fusion and domain-invariant learning contribute complementary gains, while robustness analysis showed superior resilience to noise-induced signal degradation compared to conventional architectures. These findings demonstrate that explicit domain-generalization mechanisms are essential for translating brain MRI deep learning models beyond single-institution benchmarks toward clinically dependable, multi-site deployment. Future work will extend CD-MAFNet to additional pathology categories and imaging centers, incorporate federated learning to preserve data privacy across institutions, explore transformer-based modality encoders for finer-grained attention, and validate the framework prospectively on real-world multi-vendor clinical MRI archives to further establish its diagnostic reliability and generalizability.

REFERENCES

1. Ali A, Li X, Mashwani WK, Abiad M, Karim FK, Mostafa SM. Multi-class brain tumor MRI segmentation and classification using deep learning and machine learning approaches. Cancer Imaging. 2025 Nov 13;25(1):131. doi: 10.1186/s40644-025-00953-2.
2. Vadde, V., and Bukaita, W., 2025. Comparative Evaluation of CNN and ResNet18 Architectures for MRI-Based Brain Tumor Classification Using Deep Learning. Medical Research Archives, 13(12). https://doi.org/10.18103/mra.v13i12.7100
3. Hu, T., Mousavirad, S.J., Afsharizadeh, M. et al. Deep Learning for MRI-Based Neurological Disease Diagnosis: A Comprehensive Survey of Advances, Challenges, and Benchmarks. Int J Comput Intell Syst 19, 48 (2026). https://doi.org/10.1007/s44196-025-01115-9
4. Salman, B., Yassin, E., Ganta, D., & Luna, H. (2025). Deep Learning-Based Fusion of Multimodal MRI Features for Brain Tumor Detection. Applied Sciences, 15(24), 13155. https://doi.org/10.3390/app152413155
5. Babu Vimala, B., Srinivasan, S., Mathivanan, S.K. et al. Detection and classification of brain tumor using hybrid deep learning models. Sci Rep 13, 23029 (2023). https://doi.org/10.1038/s41598-023-50505-6

6. Hussain, T., & Shouno, H. (2023). Explainable Deep Learning Approach for Multi-Class Brain Magnetic Resonance Imaging Tumor Classification and Localization Using Gradient-Weighted Class Activation Mapping. Information, 14(12), 642. https://doi.org/10.3390/info14120642
7. Anshu Singh, Maheshwari Prasad Singh, Amit Kumar Singh, Robust multi-expert deep learning framework for brain MRI classification with Taguchi optimization, Neurocomputing, Volume 650, 2025, 130824. https://doi.org/10.1016/j.neucom.2025.130824.
8. A.G, B., Srinivasan, S., D, P. et al. Robust brain tumor classification by fusion of deep learning and channel-wise attention mode approach. BMC Med Imaging 24, 147 (2024). https://doi.org/10.1186/s12880-024-01323-3
9. Ullah, N., Khan, M. S., Khan, J. A., Choi, A., & Anwar, M. S. (2022). A Robust End-to-End Deep Learning-Based Approach for Effective and Reliable BTD Using MR Images. Sensors, 22(19), 7575. https://doi.org/10.3390/s22197575
10. Haq AU, Li JP, Khan S, Alshara MA, Alotaibi RM, Mawuli C. DACBT: deep learning approach for classification of brain tumors using MRI data in IoT healthcare environment. Sci Rep. 2022 Sep 12;12(1):15331. doi: 10.1038/s41598-022-19465-1.
11. P. Khobragade, P. Ghutke, P. K. Dhankar, P. K. Adakane, A. B. Kotta and N. Purohit, "Personalized Healthcare: Using Predictive Models to Improve Treatment Outcomes for Chronic Diseases," 2025 1st International Conference on Data Science and Intelligent Network Computing (ICDSINC), Raipur, India, 2025, pp. 166-171, doi: 10.1109/ICDSINC66221.2025.11448158.
12. Alkharaan, R., Alobaidi, J., Bakarman, J., & Alshamlan, H. (2026). Brain Tumor Classification and Segmentation in MR Images Using EfficientNet and U-Net++ Models. Diagnostics, 16(11), 1745. https://doi.org/10.3390/diagnostics16111745
13. Polash, M. S. H., Saykat, M. T. H., Haque, M. E., Maniruzzaman, M., Zabin, M., & Uddin, J. (2026). An Interpretable Deep Learning Approach for Brain Tumor Classification Using a Bangladeshi Brain MRI Dataset. BioMedInformatics, 6(2), 19. https://doi.org/10.3390/biomedinformatics6020019
14. Maqsood, S., Damaševičius, R., & Maskeliūnas, R. (2022). Multi-Modal Brain Tumor Detection Using Deep Neural Network and Multiclass SVM. Medicina, 58(8), 1090. https://doi.org/10.3390/medicina58081090
15. Alshammari, H. H., & Mahmood, M. A. (2026). An Explainable Multi-Scale Deep Learning Framework for Multi-Class Brain MRI Classification. Diagnostics, 16(12), 1791. https://doi.org/10.3390/diagnostics16121791
16. Perumprath, A. S., Sagayam, K. M., Ansarullah, S. I., Khan, S., Alshmrany, S., Amin, F., Díez, I. d. l. T., Marin, M. S. G. d., & Alvarado, E. S. (2026). A Multilevel Deep Learning Model for Automated Brain Tumor Segmentation Using Magnetic Resonance Images. Diagnostics, 16(14), 2259. https://doi.org/10.3390/diagnostics16142259