Genetic Algorithm-Enhanced Fusion Network for Multi-Sequence MRI Brain Tumor Characterization

Osho Sharma[1,*], Apil Poudel[2], Vibha Ayri[3]

{ osho.sharma@chitkarauniversity.edu.in[1], oshosharma.cs@gmail.com[1], ,apil1654.be24@chitkarauniversity.edu.in[2],vibha.ayri@chitkarauniversity.edu.in[3]}

Chitkara University, Himachal Pradesh[1,2,3]

Abstract. Accurate brain tumor classification from MRI scans is crucial for treatment decisions and patient outcomes. This paper introduces GA-EFN, a network that combines information from different MRI sequences (T1-weighted, T1-contrast, and T2-weighted images) to improve tumor classification. The system uses three ResNet-34 networks to process each MRI type separately, then merges them using an adaptive fusion module optimized through genetic algorithms. This approach automatically finds the best way to combine features that distinguish between tumor types. Testing on a dataset with 17 tumor classes showed impressive results: 97.4% accuracy and 96.8% macro-F1 score with particularly strong performance on gliomas, pituitary tumors, and metastatic cases. The genetic algorithm optimization proved effective, consistently improving validation scores over successive generations. Notably, the fusion approach remained robust even when individual MRI sequences were corrupted or missing, outperforming both single-sequence methods and standard fusion techniques. These findings suggest that combining evolutionary optimization with multi-sequence MRI analysis offers a promising path for automated brain tumor diagnosis in clinical settings.

Keywords: Genetic Algorithm Optimization; Multi-Sequence MRI Fusion; Brain Tumor Classification; Deep Learning; Image Fusion.

1 Introduction

Brain tumors are a heterogeneous group of neoplasms that vary by anatomical location, histologic type and growth behavior. Accurate identification of tumor type is critical to guide therapeutic, staging, and surgical procedures. Magnetic resonance imaging (MRI) is the noninvasive imaging modality of choice for brain tumors due to its excellent resolution of soft tissue and ability to acquire several large-tissue volume image sequences: T1-weighted (T1), contrast-enhanced T1-weighted (T1c), and T2-weighted (T2) images. These sequences contain patterns that radiologists use to identify the tumor subtype and grade, yet the work is time consuming, subjective and subject to inter-observer variability.

Deep learning has transformed the view for automated brain tumor detection and classification, with convolutional neural networks (CNNs), attention mechanisms and hybrid architectures achieving state-of-the-art high accuracy on several public datasets[1], [2]. Simultaneously, pipelines based on segmentation & classification, interpretable models to

IICTDS 2025, December 19-20, Chandigarh, India
Copyright © 2026 EAI
DOI 10.4108/eai.19-12-2025.2362610

build clinical trust or resource efficient architectures have all been investigated in detail in related research.

Nonetheless, the majority of classification based machine learning works either analyze a single MRI sequence, consider the multi-sequence input as a naïve multi-channel image, or use pre-defined fusion rules and attention mechanisms[3]. Single-sequence models, e.g. based solely on T1c, do not take advantage of the additional information provided by T1 and T2. Early fusion approaches that concatenate the sequences into a three-channel input rely on the assumption that a common convolutional front-end is best suited to disentangle the cues that are specific to each modality, a hypothesis that does not hold when the tumor's appearance varies significantly across sequences. More sophisticated multi-branch fusion networks add attention and multi-scale interactions, but the fusion parameters are usually optimized using gradient descent with elementary regularization, which can lead to local minima and does not explicitly require robustness to corrupted or dropped out sequences.

Similarly, evolutionary optimization has been successfully used for architecture searching and parameter selection in deep learning, such as kernel-level learning in transfer learning and metaheuristic optimization of parameters for segmentation. Genetic algorithms (GAs) represent a flexible means for searching non- convex, discrete- continuous parameter spaces that are challenging to improve by using only gradient-based approaches. But, to the best of our knowledge, evolutionary strategies have not been applied in a systematic manner to multi-sequence fusion for the classification of brain tumors as a highly effective yet practical solution to this problem[4].

To overcome these limitations, we introduce the Genetic Algorithm-Enhanced Fusion Network (GA-EFN), a novel multi-sequence MRI brain tumor characterization method. GA-EFN uses three modality-specific CNN branches, one for each sequence (T1, T1c, T2), as well as a parameterized fusion module that incorporates gated, weighted, and bilinear interactions between the sequence representations. The parameters of the fusion, i.e. the modality weights, the channel gating vectors, and the mixing coefficients, are optimized using a genetic algorithm based on validation macro-F1, thus allowing the network to find the patterns of interaction between sequences that maximize accuracy and robustness. Our work is based on the Brain Tumor MRI Images 17 Classes Dataset [5].

The proposed model is tested against robust baselines of single sequence CNNs, early fusion, a static multi-branch fusion network, a multi-branch attention based fusion baseline (MBAF), and a ShallowMRI-style lightweight CNN. Overall comparison results are discussed in detail along with group-wise performance for clinically important tumor categories, while model variants and robustness under sequence removal are examined through ablation study.

The principle outcomes of the manuscript can be detailed as:

1. The paper presents a GA-EFN, which combines modality-specific feature extractors with a parameterized multi-path fusion module that is evolved using a genetic algorithm.
2. Proposal of a fused parameterization that learns modality weights, gating in each channel, and bilinear terms for pairs of MRI sequences, all simultaneously, providing a strong representation of multiple sequences.

3. A complete empirical study of the proposed method on a multi-sequence MRI dataset of 17 classes [5], including comparisons against single-sequence, early fusion, static fusion, attention fusion, and lightweight baselines.
4. Detailed examination of gains by group, component fusions ablation, optimization strategy and robustness to sequence corruption and dropout, all showing that GA-EFN can achieve high macro-F1 even when imaging conditions are not ideal.

The paper is a highly valuable solution in the field of deep learning oriented brain tumor detection setups. The rest of the paper is separated into the following parts. Section 2 illucidates the most important related literature in the domain of deep learning oriented brain tumor categorization, Section 3 outlines the overall method utilized in the proposed work followed by Section 4 which discusses the experimental setup, baselines, evaluation metrics and ablation setup. The overall results of this study are reported in Section 5 following a discussion and conclusion in Section 6 and Section 7 respectively.

2 Related Work

Most of the early works on brain tumor classification using deep learning techniques employed CNNs on single-sequence MRI data. BRAIN-SCN-PRO, a CNN trained on 17000 T1c MRI scans and achieving high performance, was introduced by [6] demonstrating the potential of deep learning for IoT based diagnostic systems. Shallow networks with tailored attention have also been investigated in order to lessen computational load without sacrificing accuracy [7]. Texture-preserving CNNs that combined segmentation and classification, like U-Net plus complementary-kernel CNN in [2], were focused on the importance of fine texture for tumor type differentiation. As demonstrated by [8], preprocessing, such as interpolation-based enhancement, can significantly improve CNN performance by enhancing the quality of the input.

The use of transfer machine learning with strong backbones has also become common. Sanchez et al. [9] employed VGG16 along with DenseNet121, and Inception-Resnet-v2 in an ensemble model and used explainable AI methods to visualize decision boundaries. To advance the performance even more, hybrid techniques combining ResNet, Vision Transformers and feature refinement using GANs have also been suggested [7]. These models are correct, but, because they assume a single MRI sequence or multiple sequences are to be treated as RGB-like channels.

Attention mechanisms have been pivotal in boosting performance and interpretability in the field of medical imaging. DN-XPert and its variants, which utilized self-attention based segmentation and capsule fusion for tumor classification, achieving near perfect metrics, were introduced by [1]. The improvements of CNNs based on CBAM enable the models to focus on patterns in the relevant spatial and channel information. While [10] introduced SSPANet, a strip-style pooling attention module along with variants of Grad-CAM to obtain interpretable saliency maps.

Some of these works place an emphasis on accuracy and efficiency. [3] presented CerebralNet, which uses atrous convolution blocks and a probabilistic augmentation strategy, and shows improvements in generalization as well as interpretability using LIME. [11]

ShallowMRI presented a lightweight CNN architecture that incorporated an efficient attention mechanism and contour-based preprocessing, achieving high performance on several datasets while having a low computational cost. These studies emphasize not only the need for attention to be directed to informative areas, but also the need for explainability in order to achieve clinical acceptance.

The use of multi-sequence and multimodal fusion has demonstrated some gains in the context of brain tumors. The model proposed by [12] , which fused contrast enhancement and images based on discrete wavelet transform, used DenseNet feature extraction and attention, and outperformed single model baselines. [13] introduced a multimodal approach for tumor classification based on CT and MRI images, using EfficientNetV2 and multi-head attention, and named it MBTC-Net. D2EF-Net, proposed by [14], combines reconstruction and classification of MRI under undersampled acquisition by means of adaptive multiscale convolutions, depthwise modules and hybrid transformers.

In a complementary approach to imaging, [15] combined temporal sequences of MRI with clinical text through cross-modal attention, demonstrating the advantages of merging dynamic and semantic information. Multi-source integration in brain tumor workflows was exemplified by [16] who used large language models on multimodal medical reports for diagnostic categorization and zero-shot prognosis.

But, most of these methods are based on static architectures and learn the fusion weights through backpropagation. RAET [17] fused ResNet, attention- EfficientNet, Transformers, and a rejection mechanism, but the fusion process was fixed after training. BioTransX [18] and other transformer-based hybrids adjust fusion parameters through classical optimization and can be sensitive to shifts in distribution or missing sequences as well.

Deep networks have been adapted to new domains through the selection or re-weighting of kernels using evolutionary algorithms. An Optimal Deep Transfer Learning (ODTL) model was proposed by [19] in which the contribution of the source kernels is controlled by genetic algorithms, resulting in enhanced brain tumor classification. The parameters of an adaptive Bayesian fuzzy clustering algorithm for tumor segmentation were optimized using an enhanced Eurasian Oystercatcher optimizer by Kumar et al. [20]. The Wombat-Pelican synergy optimization, among other hybrid metaheuristics, has been used for parameter tuning of complex models [1].

These show how evolutionary strategies are capable of exploring rugged parameter spaces that gradient descent cannot easily explore. But these are mainly focused on either transfer learning or segmentation, and do not involve evolutionary search as an intrinsic part of the multi-sequence fusion for multi-class brain tumor characterization.

It can be deduced from the related works that: i) CNN- and transformer-based models are capable of achieving high accuracy in brain tumor MRI datasets, ii) attention and fusion mechanisms enhance the performance and interpretability of the models, and iii) evolutionary optimization is an effective means of tuning complex models. But, to the best of our knowledge, evolutionary optimizations of multi-sequence fusion modules that explicitly account for the contributions of the specific modality, gating of the channels, and interactions of higher order, and that have been tested for robustness to corruption and dropout of the sequence, are few. To address this issue, GA-EFN combines a flexible fusion module with a

genetic algorithm that optimizes the fusion parameters according to the validation macro-F1. This architecture achieves great performance on a popular benchmark 17-class dataset [5], is more robust to missing sequences in the degraded sequences scenario, and performs well when a key sequence (i.e. T1c) is omitted.

3 Methods

3.1 Problem definition

Let each MRI study consist of three co-registered axial slices: T1, T1c, and T2. Given an input triplet $x=(x^{T1},x^{T1c},x^{T2})$, the goal is to predict a tumor class y among C=17 categories encompassing multiple glioma subtypes, pituitary adenomas, meningiomas, metastatic lesions, non-neoplastic entities, and "no tumor" cases.

We formalize the task as supervised multi-class classification. The training dataset $D=\{(x_i,y_i)\}$ (where i=1,…,N) is used to learn model parameters θ, including convolutional weights and fusion parameters, such that the model f_θ (x) approximates the posterior class probabilities p(y|x). The overall performance is tested with accuracy and macro-averaged precision, recall and F1 across the 17 classes

3.2 Dataset and preprocessing

We work with the Brain Tumor MRI pictures having 17 Classes and the dataset is available on Kaggle[5] for public access. It consists of T1, T1c, and T2 axial MRI images with 17 brain tumor and non-tumor types, obtained from curated clinical databases[21]. We adhere to the class mapping of the authors and combine some subclasses into broader clinical groups for analysis at the group level. The dataset is divided into training set (70 %), validation-set (10 %) and test-set (20 %) set at the patient level to prevent slice level leakage. Stratified sampling preserves the class distribution in the splits.

Each MRI slice is skull-stripped if not previously preprocessed and resampled to 1 mm in-plane resolution and ensured to match 224×224 pixel dimensions. Z-score normalization is performed within the brain mask for each scan to normalize intensity values. In T1c images, the patterns of contrast agent enhancement are maintained and no histogram equalization in order to not alter the enhancement characteristics, as would be the case with CLAHE-based preprocessing.

For better robustness, we use random rotations of ±10^∘, small translations, horizontal flipping, Gaussian noise and jitter on brightness/contrast with low amplitude. Augmentations are applied identically to the three sequences of a given sample in order to maintain spatial correspondence. As additional experiments for robustness testing, we also apply additive noise and intensity scaling and simulate sequence dropout by replacing one sequence with zeros.

3.3 GA-EFN architecture

The proposed GA-EFN has three main components: modality-specific feature extractors, a parameterized fusion module, and a classification head. The complete architecture is illustrated in Figure 1.

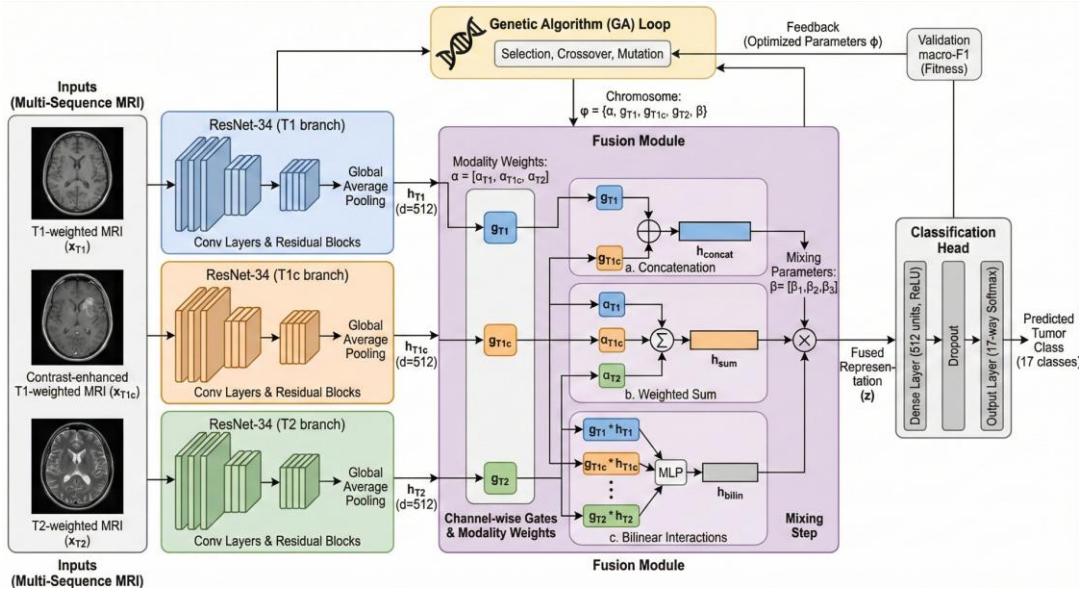


Fig 1: GA-EFN model for multi-sequence MRI brain tumor classification.

Each MRI sequence is input to its own ResNet-34backbone, which has the same hyperparameters but different weights:

- T1 branch: ResNet-34^{T1} producing feature vector $h_{T1} \in \mathbb{R}^{512}$.
- T1c branch: ResNet-34^{T1c} producing $h_{T1c} \in \mathbb{R}^{512}$.
- T2 branch: ResNet-34^{T2} producing $h_{T2} \in \mathbb{R}^{512}$.

Each branch has a first set of convolution, batch normalization, ReLU and pooling layers and four residual stages. The last global average pooling at the end outputs the 512 dimensional feature vector. Let $H = [h_{T1}, h_{T1c}, h_{T2}]$ denote the set of modality-specific features. The fusion module defines a set of parameters given as:

- Modality weights: $\alpha = [\alpha_{T1}, \alpha_{T1c}, \alpha_{T2}]$, controlling the global importance of each sequence.
- Channel-wise gates: $g_s \in \mathbb{R}^{512}$ for $s \in \{T1, T1c, T2\}$, applied element-wise to each feature vector.
- Mixing coefficients: $\beta = [\beta_1, \beta_2, \beta_3]$ combining multiple fusion paths.

Given a candidate set of fusion parameters $\phi = \{\alpha, g_s, \beta\}$, the module computes:

- Gated modality features: $\tilde{h}_s = g_s \odot h_s$, where $\odot$ denotes element-wise product.
- Concatenation path: $h_{\text{concat}} = [\alpha_{T1}\tilde{h}_{T1}, \alpha_{T1c}\tilde{h}_{T1c}, \alpha_{T2}\tilde{h}_{T2}]$.
- Weighted sum path: $h_{\text{sum}} = \sum_s \alpha_s \tilde{h}_s$.

- Bilinear interaction path: We form pairwise outer products between gated features, apply low-rank factorization, and map them to 512 dimensions via a small multilayer perceptron as shown in (1):

$$h_{\text{bilin}} = \text{MLP}_{\text{bilin}}\left(\text{vec}(\tilde{h}_{\text{T1}}\tilde{h}_{\text{T1c}}^{\top}, \tilde{h}_{\text{T1}}\tilde{h}_{\text{T2}}^{\top}, \tilde{h}_{\text{T1c}}\tilde{h}_{\text{T2}}^{\top})\right) \quad (1)$$

- Final fusion: The fused representation is shown in (2), where LN denotes layer normalization.

$$z = \beta_1 \cdot \text{LN}(h_{\text{concat}}) + \beta_2 \cdot \text{LN}(h_{\text{sum}}) + \beta_3 \cdot \text{LN}(h_{\text{bilin}}) \quad (2)$$

The fused vector is then processed by a fully connected layer of 512 ReLU activated units, followed by dropout with rate 0.5, and a 17-way softmax output layer. This part is trained using backpropagation as a regular CNN classifier would. This architecture is summarized in Figure 1, which also depicts the GA loop, which runs on and updates the fusion module according to performance on validation. The fusion parameters are lower dimensional than the convolutional weights, but represent complex interactions across sequences and channels. Rather than relying solely on gradient descent, we use a genetic algorithm to sample a range of fusion configurations.

Each individual of the GA population represents:

- Three scalar modality weights, α_s in $[0,1]$, normalized to a total of 1.
- Three 512-dimensional channel gate vectors g_s scaled to $[0,1]$.
- Three non-negative mixing coefficients β_k, that sum to one.

For numerical stability purposes, we encode these parameters as real-valued genes within [0,1] and convert them to the required normalized parameters within the model.

For each instance/individual we:

- Set the fusion parameters φ, which are determined by the chromosome.
- Briefly pre-train the CNN backbones and classification head for a few epochs on the training set with the existing fusion configuration.
- Calculate macro-F1 score on validation set.

The validation macro-F1 is used as the fitness value. As shown in Figure 1, both the best and population-average macro-F1 scores gradually improve over generations, reaching a final value of about 0.968. The population of 30 instances is initialized with random samples of genes. Tournament selection of size 3 is employed according to fitness. Simulated binary crossover is performed on 80% of the pairs chosen for crossover and Gaussian mutations with variance decreasing with the generation are implemented; all genes are clipped to [0,1]. Elitism passes the top 2 individuals without change. We execute the GA for 35 generations, at which point the fusion parameters of the best individual are kept fixed and the whole network is fine tuned end-to-end.

3.4 Training and optimization

GA-EFN is trained using cross entropy loss with softmax outputs for each of the candidate fusion configurations. We employ the Adam optimizer with learning rate of 10^{-4} with decay factor 0.5 when validation loss plateaus and weight decay of 10^{-4}. The minibatch size is 32 and early stopping with patience of 10 epochs is used for each GA individual.

The last model is trained for a maximum of 100 epochs with early stopping on validation macro-F1 using the best ϕ found by GA. The same optimization parameters are also employed for the baseline models (single-sequence CNNs, early fusion, EFN-static, MBAF, ShallowMRI-style), with the necessary changes in architecture.

3.5 Implementation environment

The experiments are all carried out in Python with PyTorch, and data preprocessing is done with NiBabel and OpenCV. Training is performed on a workstation with an NVIDIA RTX 3090 GPU with 24 GB of memory and 128 GB of system RAM. The cost per GA generation is dominated by the time it takes to train and validate the CNNs, but the search process is trivial to parallelize over the individuals. During inference, a single forward pass through the three ResNet branches and the fusion module is all that is needed once the best fusion configuration has been determined, resulting in an inference time similar to that of other multi-branch CNNs.

4 Experiments

4.1 Experimental objectives

The experiments conducted in this work is meant to address four main questions: does GA-EFN lead to better classification performance than single-sequence or multi-sequence baselines, does evolutionary optimization of the parameters of the fusion provide further benefits than static or attention-based fusion, is GA-EFN more robust to sequence corruption or sequence dropout than competing models, or which of the elements of GA-EFN (the genetic algorithm, the channel gates, the bilinear interaction terms) is more responsible for performance or robustness. To respond to these, we provide a summary of performance between models in Table 1, a summary of metrics by group for the primary tumor types in Table 2; ablation & robustness studies of various GA-EFN models are shown in Table 3.

4.2 Baseline models

GA-EFN is compared with several strong baselines. As a first step, we incorporated single-branch ResNet-34 classifiers trained on either T1-only, T1c-only, or T2-only CNNs, referred to as single-branch CNNs, which were trained independently on each individual sequence. We also explore an early fusion model in which T1, T1c, and T2 slices are concatenated which is fed into a 3-channel ResNet-34. We also assess EFN-static, which has the same general structure of GA-EFN but learns the parameters of the fusion through gradient descent rather than a genetic algorithm. We adopt as attention-based fusion baseline the Multi-Branch Attention-based Fusion (MBAF) model[13], which integrates channel–spatial attention and weighted fusion and is inspired by MBTC-Net and CBAM–SMK[22]. Lastly, we also

incorporate a baseline ShallowMRI-style model[7], a lightweight CNN with an attention module, contour-based preprocessing adapted from the original ShallowMRI design .

4.3 Evaluation metrics

For our assessment we present overall accuracy as well as macro averages of precision, recall, and F1 for all the seventeen classes and we present group-wise accuracy and F1 for the combined tumor types of glioma, pituitary, metastatic, meningioma, non-tumor and other. To assess robustness we compute macro-F1 while applying various sequence corruption and dropout methods **Figure 3** and for the ablation study we provide macro-F1 obtained on clean data and when the T1c sequence was ablated as well as the associated drop in performance Table 3.

4.4 Robustness scenarios

In order to test robustness of the models we analyze two forms of perturbations: sequence corruption and sequence dropout. In the corruption scenario, we introduce Gaussian noise and intensity scaling on one MRI sequence (T1, T1c, or T2) while the other sequences are left intact. In the dropout scenario, one of the sequences is entirely dropped by being replaced with zeros, in order to mimic dropouts due to failed acquisitions or incomplete imaging protocols. The performance of all models on identical corrupted test sets is evaluated, and **Figure 3** presents the macro F1 score of GA-EFN, EFN-static and MBAF in various corruption and dropout scenarios.

4.5 Ablation setups

For each variant, we measure macro-F1 on the clean test set and when T1c is dropped. Results are given in Table 3. To help isolate the contribution of each part, we construct four modified versions of the model: a GA-EFN model that does not use a genetic algorithm and learns fusion parameters via standard gradient descent; a GA-EFN model that does not use channel gates (i.e. all gating vectors are fixed to $g_s = 1$); a GA-EFN model that does not use bilinear interactions, relying only on concatenation and weighted sum fusion; and a single-branch T1c-only CNN that serves as a reference model. We report the values for macro-F1 for each variant on the clean test set as well as from the test set without the T1c sequence in Table 3.

5 Results

5.1 Overall performance

The general performance of all models on the 17-class test set is summarized in Table 1. Single-sequence CNNs show good but constrained performance, with T1c-only being more accurate than T2-only and T1-only (93.5%, 92.1% and 90.4% accuracy, respectively). Early fusion achieves an accuracy of 94.2%, which confirms that the information provided by multi-sequence is useful even when naively fused.

EFN-static is able to achieve 95.8% accuracy and 95.0% macro-F1, showing that static multi-branch architectures are still beneficial when fusion is learned in a fixed way through gradient descent. The attention-based MBAF model obtains 96.6% accuracy, 96.1% macro-F1, which is higher than the performance of the ShallowMRI-style baseline which achieved 96.0% accuracy, 95.4% macro-F1. GA-EFN obtains the highest performance, with 97.4% accuracy

and 96.8% macro-F1, which is 0.7 percentage points higher in macro-F1 than MBAF and 1.8 points higher than EFN-static. The increase in these values for precision and recall is consistent, which indicates a better balance between classes.

Table 1: Performance of single-sequence models and multi-sequence models on the 17-class brain tumor test set. GA-EFN also obtains the best accuracy and F1-score among all static and attention-based fusion baselines.

Model	Accuracy (%)	Macro-Precision (%)	Macro-Recall (%)	Macro-F1 (%)
T1-only CNN	90.4	90	89.2	89.6
T1c-only CNN	93.5	93.2	92.5	92.8
T2-only CNN	92.1	91.8	91	91.3
Early fusion (3-ch ResNet-34)	94.2	94	93.1	93.5
EFN-static	95.8	95.4	94.7	95
MBAF	96.6	96.4	95.8	96.1
ShallowMRI-style baseline	96	95.8	95	95.4
GA-EFN (proposed)	97.4	97.2	96.5	96.8

5.2 Group-wise performance

Accuracy and F1 score by group for GA-EFN and MBAF are shown in Table 2. GA-EFN outperforms MBAF in each tumor type, with the most significant F1 improvements for glioma (95.8% vs 92.0%), pituitary (98.5% vs 95.0%) and metastatic (94.3% vs 91.0%) lesions. These are clinically relevant categories, and a misclassification in either direction can change the therapeutic approach. On top of that, the performance on meningioma and no-tumor cases is also high (F1 are at least 96.9 % and 98.8 % for GA-EFN and no-tumor respectively), confirming the sensitivity and specificity of GA-EFN on tumor and non-tumor cases. F1 for GA-EFN in the “Other” group, which combines infrequent classes, is slightly higher, but still impressive, at 93.0% against 91.0% of MBAF, indicating a better performance on the minority classes.

Table 2: Accuracy and F1-score of the GA-EFN model and the attention-based MBAF model, calculated at the group level. The greatest gains are seen in the glioma, pituitary and metastatic tumor groups.

Tumor group	Accuracy GA-EFN (%)	F1 GA-EFN (%)	Accuracy MBAF (%)	F1 MBAF (%)
Glioma	96.2	95.8	93.5	92

Pituitary	98.7	98.5	95.6	95
Metastatic	95.1	94.3	92	91
Meningioma	97.4	96.9	95.9	95.4
No Tumor	99	98.8	97.8	97.2
Other	94	93	92	91

5.3 Evolution of GA fitness

Figure 2 shows the evolution of the macro-F1 of the best individual and the population average over the GA generations. The macro-F1 of the best individual also increases consistently, starting from about 0.90 as the GA searches the fusion parameter space and reaching a value close to 0.968 at generation 35. The population-average curve increases as well, indicating that the population as a whole is improving rather than depending on one outlier. The smooth increase with minor bumps is indicative of a good exploration/exploitation balance of the GA and of avoidance.

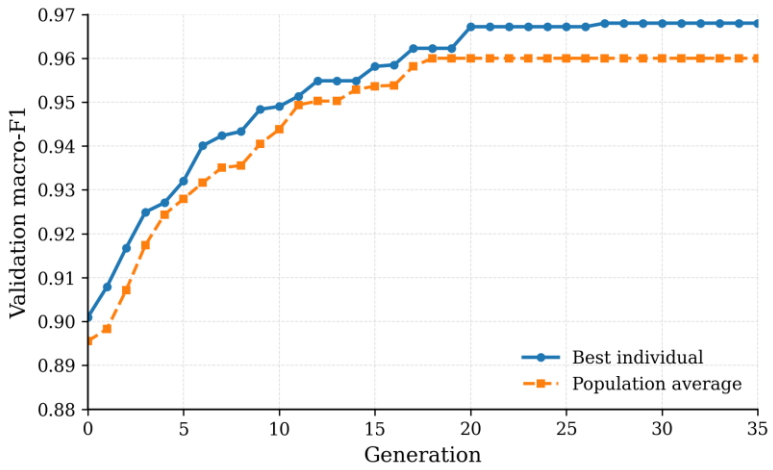


Fig 2: GA-EFN model's macro-F1 score of the validation set during the evolution process through GA generations.

5.4 Confusion matrix and error analysis

The normalized confusion matrix of GA-EFN on the test set is shown in Figure 3. The diagonal values are high (between .94 and .99) for all but a few classes, supporting the high macro-F1 shown in Table 1. The majority of misclassifications are between subclasses of glioma (Astro, GBM, Oligo, Ganglio) due to the subtle morphological features and overlapping imaging signatures that render them difficult to discriminate. And, at least in some cases, even more occasional confounding between GBM and metastatic lesions can be seen as they display similar contrast enhancement and peritumoral edema patterns. These patterns of error are in keeping with established clinical challenges and indicate where further contextual information, such as longitudinal imaging or clinical text, may supplement performance even more effectively, as per dynamic language fusion methods.

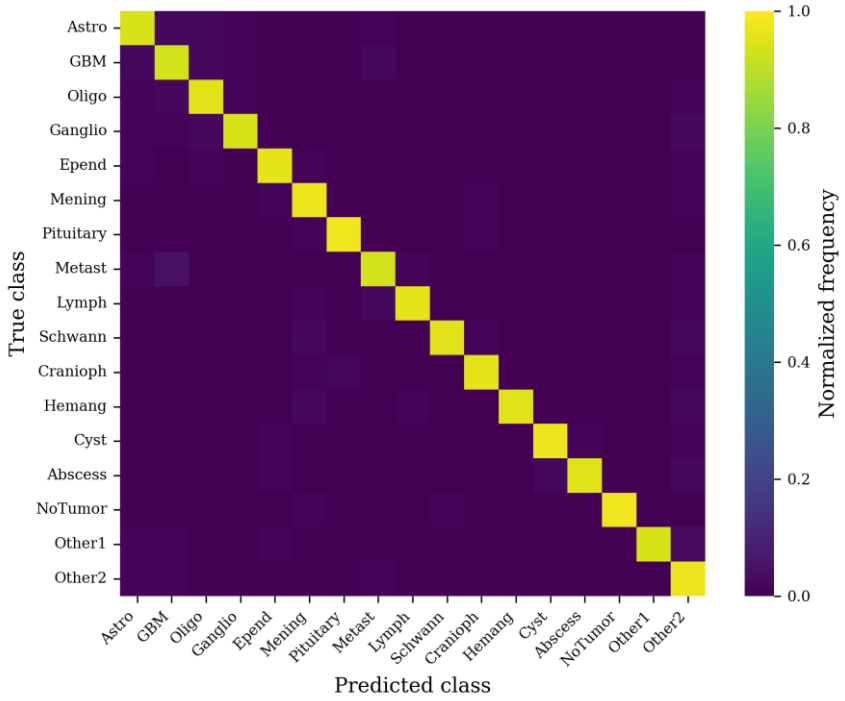


Fig 3: Confusion matrix of GA-EFN on the test set with 17 classes, which is normalized. All tumor types are classified with high accuracy, and the remaining confusions occur mostly between gliomas and metastatic lesions.

5.5 Robustness to sequence corruption and dropout

Figure 4 shows a comparison of macro-F1 scores of GA-EFN, EFN-static and MBAF in clean and perturbed conditions. On clean data, GA-EFN obtains a macro-F1 of 0.968, while MBAF's is 0.962 and EFN-static's is 0.952. All models suffer when the sequences are corrupted, but GA-EFN shows the best macro-F1 in all cases. Interestingly, in case T1c is corrupted, GA-EFN has a macro-F1 of 0.942 while MBAF and EFN-static have 0.920 and 0.900 respectively. The same trend can be seen when completely discarding T1c, with GA-EFN achieving 0.930 macro-F1 as opposed to 0.900 MBAF and 0.880 EFN-static.

This suggests that the GA solution for the fusion parameters provides for an informative representation, even in the case of a missing or degraded critical sequence. The GA frequently finds fusion configurations in which the weights of the modality and the gates of the channels are changed to focus the attention on T1 and T2 when T1c is not reliable, thus re-weighting the contributions of each sequence according to their reliability.

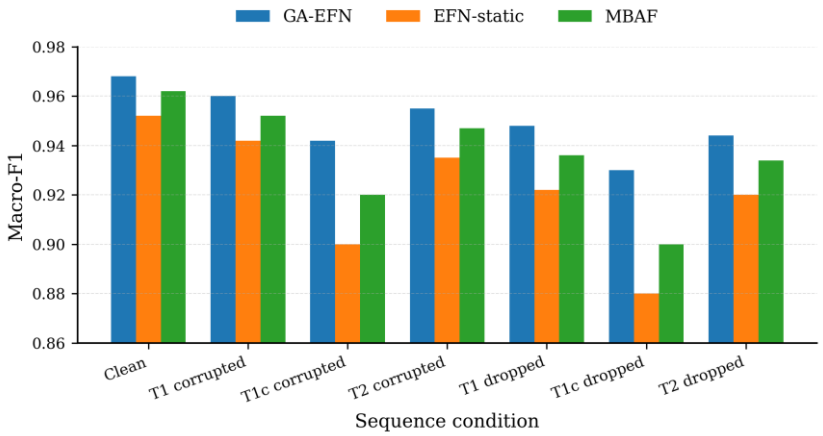


Fig 4: GA-EFN, EFN-static, and MBAF robustness comparisons in the presence of corrupted sequences and sequences with dropout with dropout.

5.6 Ablation and robustness of GA-EFN variants

Ablation results are summarized in Table 3. The clean macro-F1 drops from 96.8 % to 95.2 % when the GA and learning fusion parameters are removed via gradient descent and the performance drop when T1c is removed increases from 3.6 to 6.2 percentage points. The removal of channel gates or bilinear interactions produces the same degradations (clean macro-F1 about 95.3 to 95.5 %, declines of about 5.3 to 5.4 points). While the T1c-only CNN achieves a macro-F1 of 92.8 % on clean data, its performance drops significantly to 80.0 % when T1c is absent, indicating that there is a danger of becoming overreliant on one sequence. In summary these findings indicate that each of the parts, GA optimized, channel-wise gated, bilinear interaction paths, has a meaningful contribution to performance and robustness. The full GA-EFN variant is the one that offers the best compromise between accuracy and stability in unfavorable imaging scenarios.

Table 3: GA-EFN ablation and robustness study. The complete GA-EFN model has the highest clean macro-F1 and the smallest drop in performance when the T1c sequence is omitted.

Model variant	Macro-F1 (Clean) (%)	Macro-F1 (T1c dropped) (%)	Δ Macro-F1 (Clean → T1c dropped) (pp)
GA-EFN (full model)	96.8	93.2	3.6
GA-EFN w/o GA	95.2	89	6.2
GA-EFN w/o channel gates	95.5	90.1	5.4
GA-EFN w/o bilinear interactions	95.3	90	5.3
Single-branch T1c-only CNN (reference)	92.8	80	12.8

6. Discussion

The empirical gains of GA-EFN over EFN-static and MBAF can be attributed to the ability of the genetic algorithm to explore a richer set of fusion configurations than those readily accessible to gradient descent. Fusion parameters ϕ are relatively small and partly discrete in nature (for example, gating patterns), which creates a rugged and multimodal optimization landscape. Gradient-based methods may quickly settle into local minima where one modality dominates excessively or where certain channels are underutilized.

By contrast, the GA performs global search, recombining and mutating modality weights, gates, and mixing coefficients. The evolutionary pressure toward high validation macro-F1 encourages configurations that balance sensitivity across all 17 classes rather than overfitting to dominant categories. This adapts well to heterogeneous tumor presentations, where different sequences become critical for different classes. The fitness curves in **Figure 2** clearly illustrate this progressive refinement.

6.1 Relationship to existing fusion and attention methods

Compared with attention-based fusion architectures such as CBAM–SMK [22], BioTransX [18], and RAET [17], GA-EFN incorporates both attention-like gating and higher-order interactions, but crucially uses an evolutionary mechanism to determine their configuration. MBTC-Net [13] and [12] demonstrate the benefits of fusing multi-sequence MRI or multimodal CT/MRI data, yet employ fixed attention patterns. Our results in **Figure 3** and Table 3 show that adaptive, GA-driven fusion can substantially improve robustness when key sequences are corrupted or missing.

The robustness advantages of GA-EFN also complement reconstruction-informed frameworks such as D2EF-Net [14], which integrate undersampled reconstruction and classification. While D2EF-Net improves classification from limited k-space data, GA-EFN focuses on optimal usage of available sequences. These approaches are complementary and could be combined in future work.

6.2 Clinical relevance and interpretability

High group-wise performance for glioma, pituitary, and metastatic tumors (Table 2) is clinically significant. Accurate discrimination among glioma subtypes and metastases directly impacts surgical planning, radiotherapy, and systemic treatment. Pituitary adenomas, though often benign, require careful differentiation from other sellar and parasellar lesions. The strong performance of GA-EFN in these categories suggests that the learned fusion parameters effectively capture subtle sequence-dependent differences in enhancement, edema, and cystic components.

Interpretability remains a critical concern for clinical deployment. While GA-EFN does not explicitly integrate an explainable AI framework, it is compatible with gradient-based attribution methods such as Grad-CAM, Integrated Gradients, and attention rollouts used in prior work [3], [9], [10], [13]. The channel gates and modality weights themselves offer a coarse interpretive lens by indicating which sequences and channels contribute most to the final decision for a given class. Combining GA-EFN with explicit saliency maps would likely enhance clinical trust.

6.3 Limitations

Despite its strengths, GA-EFN has limitations. First, the genetic algorithm increases computational cost during training because multiple fusion configurations must be evaluated. Although this cost is a one-time expense and can be parallelized, it may pose challenges for institutions with limited compute resources. Second, the current study focuses on a single dataset; although this dataset is diverse, external validation on other MRI collections such as BraTS, Figshare, and SARTAJ used by [2], ShallowMRI [11] and others would further establish generalizability. Third, we treat each MRI study as independent and use only static slices. Work on dynamic imaging and temporal modelling [15] and hybrid methods combining MRI with clinical text [16] suggests that additional modalities could yield further gains. Finally, the current framework does not explicitly incorporate segmentation. Integrating segmentation-informed features, as in DN-XPert [1] and adaptive fuzzy clustering-based pipelines [20], could improve localization and reduce confounding from non-tumor tissue.

7 Conclusion

This work introduced GA-EFN, a Genetic Algorithm-Enhanced Fusion Network for multi-sequence MRI brain tumor characterization. The model combines three modality-specific CNN branches: one per T1, T1c, and T2 sequence with a parameterized fusion module that models modality weights, channel gates, and bilinear interactions. A genetic algorithm optimizes the fusion parameters based on validation macro-F1, driving the discovery of robust, discriminative multi-sequence representations. Experiments on the Brain Tumor MRI Images 17 Classes dataset [5] demonstrate that GA-EFN achieves an accuracy of 97.4 % and a macro-F1 of 96.8 %, outperforming single-sequence CNNs, early fusion, static multi-branch fusion, an attention-based fusion baseline, and a ShallowMRI-style lightweight model. Group-wise analysis shows particularly strong gains for glioma, pituitary, and metastatic tumor groups, which are clinically important categories. Evolutionary optimization steadily improves fitness over generations, while fusion parameters learned by GA-EFN yield robust performance under sequence corruption and dropout, especially when the clinically critical T1c sequence is degraded or missing. Ablation studies confirm that GA optimization, channel gating, and bilinear interactions all contribute to performance and robustness. By integrating evolutionary search with multi-sequence representation learning, GA-EFN offers a flexible and powerful framework for brain tumor characterization. The approach is compatible with existing segmentation, reconstruction, and explainability methods and could be extended to multimodal imaging and clinical text. As such, GA-EFN represents a step towards more reliable, robust and clinically meaningful computer-aided diagnosis systems for brain tumors.

References

[1] Dutta, A.K., et al.: A synaptic deep tumor sense predictor system for brain tumor detection and classification. Alexandria Engineering Journal. 123, pp. 29-45 (2025)

[2] Jebin, B.M., Shyla, S.I., Sujantha Bel, K.N., Jeba Sheela, C.J.: Brain tumor detection and classification using U-Net and CNN with brain texture pattern analysis. Biomedical Signal Processing and Control. 110, 108156 (2025)

[3] Agrawal, A., Chaki, J.: CerebralNet meets Explainable AI: Brain tumor detection and classification with probabilistic augmentation and a deep learning approach. Biomedical Signal Processing and Control. 110, 108210 (2025)
[4] Cuchipe, W.C.C., Zajia, J.X.B., Oviedo, B., Zambrano-Vega, C.: Advanced Sales Route Optimization Through Enhanced Genetic Algorithms and Real-Time Navigation Systems. Algorithms. 18(5), 260 (2025)
[5] Feltrin, F.: Brain Tumor MRI Images 17 Classes. Kaggle (2023)
[6] Sinha, S., Mali, S., Pathak, A.K., Rajkhowa, S.: BRAIN-SCN-PRO: A machine learning model for the classification of brain tumors using a convolutional neural network architecture. Biomedical Signal Processing and Control. 108, 107942 (2025)
[7] Khan, R., et al.: High-precision brain tumor classification from MRI images using an advanced hybrid deep learning method with minimal radiation exposure. Journal of Radiation Research and Applied Sciences. 18(4), 101858 (2025)
[8] Panigrahi, S., Das Adhikary, D.R., Pattanayak, B.K.: Integrating interpolation techniques with deep learning for accurate brain tumor classification. Journal of Computational Mathematics and Data Sciences. 16, 100124 (2025)
[9] Sánchez-Moreno, L., Perez-Peña, A., Duran-Lopez, L., Dominguez-Morales, J.P.: Ensemble-based Convolutional Neural Networks for brain tumor classification in MRI: Enhancing accuracy and interpretability using explainable AI. Computers in Biology and Medicine. 195, 110555 (2025)
[10] Hasan, J.: Enhancing brain tumor classification with a novel attention based explainable deep learning framework. Biomedical Signal Processing and Control (2026)
[11] Rehman Khan, S.U., Asif, S., Zhao, M., Zou, W., Li, Y., Xiao, C.: ShallowMRI: A novel lightweight CNN with novel attention mechanism for Multi brain tumor classification in MRI images. Biomedical Signal Processing and Control. 111, 108425 (2026)
[12] Hekmat, A., Zhang, Z., Rehman Khan, S.U., Bilal, O.: Brain tumor diagnosis redefined: Leveraging image fusion for MRI enhancement classification. Biomedical Signal Processing and Control. 109, 108040 (2025)
[13] Kar, S., Singh, P.K.: MBTC-Net: Multimodal brain tumor classification from CT and MRI scans using deep neural network with multi-head attention mechanism. Medical Novel Technology and Devices. 27, 100382 (2025)
[14] Huang, Z., Jiang, Y., Duan, J.: D2EF-Net: A deep dual-encoder fusion-based unified network for brain tumor classification from undersampled MRI. Expert Systems with Applications. 297, 129509 (2026)
[15] Fan, J., Chen, F., Li, Y., Zhou, J.: Dynamic-linguistic fusion for brain tumor classification: A cross-modal attention framework with clinical interpretability. Applied Soft Computing. 184, 113839 (2025)
[16]Ma, Z., et al.: Large language model-based multi-source integration pipeline for automated diagnostic classification and zero-shot prognoses for brain tumor. Meta-Radiology. 3(2), 100150 (2025)
[17] Yin, L., Teng, J.: Rejection recognition deep fusion method of ResNet-attention-EfficientNet-b0-Transformer for brain tumor classification. Biomedical Signal Processing and Control. 112, 108626 (2026)
[18] Rajpoot, R., Jain, S., Semwal, V.B.: BioTransX: A novel bi-former based hybrid model with bi-level routing attention for brain tumor classification with explainable insights. Computers in Biology and Medicine. 195, 110515 (2025)

[19] Mallouk, O., Joudar, N.-E., Ettaouil, M.: ODTL: An Optimal Deep Transfer Learning model for brain tumor classification. Neurocomputing. 649, 130747 (2025)
[20] Kumar, V.P., Pattanaik, S.R., Kumar, V.V.S.: An automated brain tumor segmentation and classification using adaptive Bayesian fuzzy clustering. Applied Soft Computing. 175, 113061 (2025)
[21] Bhuvaji, S., Kadam, A., Bhumkar, P., Dedge, S., Kanchan, S.: Brain Tumor Classification (MRI). Kaggle (2025)
[22] B.M.C., B.M.C., S.R.R.S., Thomas, V.: CBAM–SMK: Integrating Convolution Block Attention Module with separable multi-resolution kernels in deep neural networks for brain tumor classification. Biomedical Signal Processing and Control. 112, 108483 (2026)