A Hybrid Particle Swarm--Genetic Algorithm Framework for U-Net Hyperparameter Optimization in High-Precision Brain Tumor MRI Segmentation

Shoffan Saifullah *, Rafał Dreżewski *, Anton Yudhana, Radius Tanone, Andiko Putro Suryotomo

Posted Date: 6 March 2026

doi: 10.20944/preprints202603.0466.v1

Keywords: brain tumor segmentation; U-Net optimization; Particle Swarm Optimization (PSO); Genetic Algorithm (GA); hybrid metaheuristics; medical image analysis

 doi:10.20944/preprints202603.0466.v1

Article

A Hybrid Particle Swarm–Genetic Algorithm Framework for U-Net Hyperparameter Optimization in High-Precision Brain Tumor MRI Segmentation

Shoffan Saifullah [1,2,*], Rafał Dreżewski [1,3,*], Anton Yudhana [4], Radius Tanone [5] and Andiko Putro Suryotomo [2]

[1] Faculty of Computer Science, AGH University of Krakow, Krakow 30-059, Poland
[2] Department of Informatics, Universitas Pembangunan Nasional Veteran Yogyakarta, Yogyakarta 55283, Indonesia
[3] Artificial Intelligence Research Group (AIRG), Informatics Department, Faculty of Industrial Technology, Universitas Ahmad Dahlan, Yogyakarta 55181, Indonesia
[4] Department of Electrical Engineering, Universitas Ahmad Dahlan, Yogyakarta 55191, Indonesia
[5] Artifificial Intelligence Research Center, Universitas Kristen Satya Wacana, Salatiga 50711, Indonesia
* Correspondence: shoffans@upnyk.ac.id (S.S.); drezew@agh.edu.pl (R.D.)

Abstract

Accurate and robust brain tumor segmentation remains a critical challenge in medical image analysis due to high inter-patient variability, complex tumor morphology, and modality-specific noise in MRI scans. This study proposes PSO-GA-U-Net, a novel hybrid deep learning framework that integrates Particle Swarm Optimization (PSO) and Genetic Algorithms (GA) to optimize the U-Net architecture for enhanced segmentation performance and generalization. PSO dynamically tunes the learning rate to accommodate modality-specific variations, while GA adaptively regulates dropout to improve feature diversity and reduce overfitting. The model is evaluated on three benchmark datasets—FBTS, BraTS 2021, and BraTS 2018—using five-fold cross-validation. PSO-GA-U-Net achieves Dice Similarity Coefficients (DSC) of 0.9587, 0.9406, and 0.9480, and Jaccard Index (JI) scores of 0.9209, 0.8881, and 0.9024, respectively, consistently outperforming state-of-the-art models in both overlap accuracy and boundary delineation. Statistical tests confirm that these improvements are significant across folds ($p < 0.05$). Visual heatmaps further illustrate the model's ability to preserve structural integrity across tumor types and modalities. These results indicate that metaheuristic-guided deep learning offers a promising and clinically applicable solution for automatic tumor segmentation in radiological workflows.

Keywords: brain tumor segmentation; U-Net optimization; Particle Swarm Optimization (PSO); Genetic Algorithm (GA); hybrid metaheuristics; medical image analysis

1. Introduction

Brain tumors pose significant clinical and radiological challenges due to their aggressive progression [1], diverse histopathological characteristics [2], and heterogeneous presentation in magnetic resonance imaging (MRI) [3]. Accurate and timely segmentation of brain tumors is essential for treatment planning, surgical navigation, and post-operative assessment [4–6]. Manual delineation by expert radiologists, although precise, is labor-intensive and subject to intra- and inter-observer variability [7–9], particularly in multimodal MRI where subtle boundary differences and modality-specific artifacts frequently occur [2,10,11]. This has propelled the development of automated segmentation systems, with deep learning methods—especially convolutional neural networks (CNNs)—emerging as the de facto standard for high-throughput tumor delineation [12,13].

Among CNN-based architectures, U-Net and its variants have demonstrated remarkable performance in biomedical segmentation tasks due to their encoder-decoder structure with skip connections that preserve spatial context [14–17]. However, despite their widespread adoption, U-Net models

 doi:10.20944/preprints202603.0466.v1

often exhibit limited generalization across datasets due to static architecture configurations, sensitivity to hyperparameter settings, and vulnerability to overfitting [18,19]—especially when exposed to heterogeneous tumor morphology or imbalanced multimodal inputs [20,21]. These challenges are further exacerbated in real-world clinical datasets where MR acquisitions may vary in contrast, noise levels, and tumor size.

To overcome these limitations, recent studies have explored metaheuristic optimization strategies such as Particle Swarm Optimization (PSO) and Genetic Algorithms (GA) to adaptively tune deep network hyperparameters. PSO is known for its global convergence ability in continuous optimization problems [22–24], while GA [25–27] offers robust exploration through population-based dropout and structural tuning. However, these methods are often applied independently and fail to harness their complementary strengths. Moreover, previous works lack an integrated framework that simultaneously optimizes both learning dynamics and architectural regularization to improve boundary precision and robustness across diverse types of tumors.

In this work, we propose a novel hybrid optimization-based deep learning framework, named PSO-GA-UNet, which leverages PSO for dynamic learning rate adaptation and GA for dropout regulation within the U-Net architecture. This dual-strategy optimization enhances the model's ability to capture fine-grained tumor boundaries while preserving global contextual integrity. The model is extensively evaluated on three benchmark datasets—FBTS (Figshare Brain Tumor Segmentation), BraTS 2021, and BraTS 2018—across multiple tumor classes and MRI modalities. Experimental results demonstrate that PSO-GA-U-Net consistently outperforms state-of-the-art (SOTA) models in the Dice Similarity Coefficient (DSC), Jaccard Index (JI), and boundary preservation. Statistical significance tests confirm the reliability of improvements across five-fold cross-validation settings, while qualitative heatmaps further illustrate the structural fidelity of the model in complex tumor regions.

The main contributions of this study are summarized as follows:

- We design a novel hybrid metaheuristic optimization framework (PSO-GA-U-Net) that integrates PSO and GA to jointly optimize learning rate dynamics and dropout rates in U-Net for brain tumor segmentation.
- We implement a robust training pipeline validated across three challenging datasets (FBTS, BraTS 2021, BraTS 2018), ensuring generalization across tumor types, MRI modalities, and anatomical variations.
- We conduct extensive performance evaluations using both quantitative metrics (DSC, JI, HD, ASSD) and qualitative heat map analyses, alongside statistical validation using paired t-tests on cross-validation folds.
- We present a comprehensive comparison with existing SOTA models, demonstrating superior performance in terms of volumetric overlap, boundary precision, and robustness to modality-specific noise.

The remainder of the paper is structured as follows: Section 2 presents an overview of related works, highlighting recent developments in deep learning and metaheuristic optimization for medical image segmentation. Section 3 details the proposed PSO-GA-U-Net architecture, the hybrid optimization strategy, and the experimental setup. Section 4 presents the quantitative and qualitative results, comparative evaluations, and statistical analyses. Finally, Section 5 concludes the paper with a summary of the findings, discussion of limitations, and directions for future research.

2. Related Works

The field of automated brain tumor segmentation has evolved significantly with the rise of deep learning methods, particularly convolutional neural networks (CNNs) [8,28]. Among these, the U-Net architecture has become a foundational backbone due to its encoder-decoder structure with skip connections that enable multiscale feature fusion [29–31]. However, traditional U-Net variants often suffer from rigidity in parameter selection and limited adaptability to complex tumor morphologies and multimodal MRI inputs [15,32,33].

 doi:10.20944/preprints202603.0466.v1

To enhance segmentation performance, numerous studies have focused on architectural modifications to the U-Net. Attention-based models such as U-Net-AG [30,34–36] integrate spatial attention gates to guide the model's focus toward salient tumor regions. Residual and dense connectivity strategies, as employed in Residual-Attention-U-Net [8,37–40], aim to improve gradient flow and multi-scale representation. However, these models are often manually tuned and sensitive to overfitting when confronted with high variability in tumor intensity and shape. Recent works, such as DeepLabV3+ [41,42] and SPPNet [43], introduce atrous spatial pyramid pooling to address multiscale features; however, they lack adaptive learning dynamics.

To address hyperparameter sensitivity, researchers have turned to bio-inspired optimization techniques [44]. Particle Swarm Optimization (PSO), inspired by collective behavior in swarms, has been applied to optimize learning rates [24,33], loss function weights [45,46], and even thresholding in preprocessing pipelines [24,44,47]. PSO has demonstrated success in the efficient tuning of continuous value parameters, but can suffer premature convergence in high-dimensional search spaces.

On the other hand, Genetic Algorithms (GA) excel in exploring diverse architectural variants through population-based evolution strategies [48–50]. In segmentation tasks, GA has been used to evolve filter configurations [51,52], dropout rates, and layer depths [53,54]. Despite its global search capability, GA's performance can degrade when used alone due to a lack of convergence guidance or redundancy in feature activation [55,56].

Several studies have explored hybrid optimization to mitigate these individual shortcomings. Works such as PSO-U-Net [24] and GA-U-Net [53,57] demonstrate that integrating optimization algorithms into deep learning pipelines can yield better generalization. However, most prior methods employ these techniques in isolation or apply them to static configurations, without a coordinated mechanism to optimize both training dynamics (e.g., learning rate schedules) and regularization factors (e.g., dropout schemes) simultaneously.

In the context of benchmark datasets, models such as ViT-self-attention [58] and AWA-VGG-19 [59] have been applied to the BraTS dataset with competitive results. However, Transformer-based models often require extensive training data and struggle with local feature refinement, particularly in edematous or necrotic tumor subregions. Cascaded architectures such as 2C-U-Net [60] and U-Net-ASPP-EVO [61] improve multi-resolution learning, but are computationally intensive and sensitive to modality imbalance.

The proposed PSO-GA-U-Net differs from existing approaches in several key ways:

- In contrast to previous PSO or GA models, it introduces a dual-level optimization pipeline where PSO governs adaptive learning rate scheduling during training, while GA dynamically evolves the dropout rate to prevent overfitting and promote feature diversity.
- The architecture remains U-Net-based for its proven spatial fidelity, but is optimized across cross-validation folds and datasets, ensuring both performance and generalization.
- The evaluation covers diverse tumor types (meningioma, glioma, pituitary, HGG, LGG) and MRI modalities (T1CE, FLAIR, T1, T2), with consistent superiority in DSC and JI compared to transformer-based, residual, and attention-enhanced models.
- Statistical significance tests ($p < 0.05$) are used to validate the robustness of improvements across datasets and tumor classes.

In summary, while various strategies have been employed to improve tumor segmentation—ranging from architectural innovations to standalone optimizers—our work uniquely combines adaptive learning and architectural regularization into a single, hybrid metaheuristic framework. This positions PSO-GA-U-Net as a robust and scalable solution for clinical-grade segmentation in MRI-based tumor analysis.

 doi:10.20944/preprints202603.0466.v1

3. Methods

3.1. Dataset Description

To comprehensively evaluate the robustness, adaptability, and generalization capability of the proposed PSO-GA-U-Net framework, three publicly available brain tumor segmentation datasets were used: the Figshare Brain Tumor Segmentation (FBTS) dataset, the BraTS 2021 dataset, and the BraTS 2018 dataset. These datasets vary in tumor types, imaging modalities, and annotation granularity, providing a diverse and challenging experimental setup. All datasets were preprocessed to a uniform spatial resolution of 256×256 pixels and normalized to the range $[0, 1]$ prior to training.

3.1.1. Figshare Brain Tumor Segmentation (FBTS) Dataset

The FBTS dataset consists of T1-weighted contrast-enhanced (T1CE) axial MRI slices sourced from a curated clinical dataset [62]. It contains manually annotated tumor masks for three types of tumors: meningioma, glioma, and pituitary adenoma. Each slice is associated with a binary mask that delineates the tumor region.

- **Meningioma**: 3,060 slices;
- **Glioma**: 708 slices;
- **Pituitary Tumor**: 930 slices.

Tumor sizes and shapes vary significantly across classes, with gliomas exhibiting irregular, infiltrative structures, and pituitary tumors generally appearing as small, well-circumscribed masses. Each image is a single-channel but expanded to three channels during preprocessing to meet the input requirements of the deep learning model. The binary masks were thresholded and rescaled to ensure compatibility with the sigmoid output of the segmentation network.

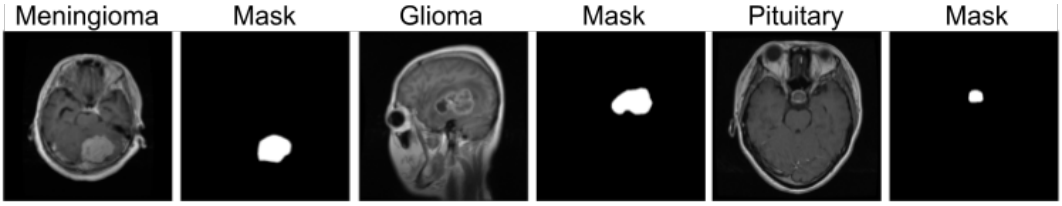


Figure 1. Representative examples from the FBTS dataset: Meningioma, Glioma, and Pituitary tumors with their respective binary masks.

3.1.2. BraTS 2021 Dataset

The BraTS 2021 dataset comprises axial MRI slices acquired using four distinct modalities: T1, T2, T1CE (post-contrast), and FLAIR [63]. Each sample is annotated with a label representing the tumor region, which in our study is converted into a binary whole tumor mask. This dataset is particularly relevant for evaluating modality-specific performance, as each MRI sequence highlights different pathological characteristics:

- T1 emphasizes anatomical structures;
- T2 highlights fluid-filled regions;
- T1CE captures active tumor enhancement;
- FLAIR is sensitive to edema and non-enhancing tumor cores.

All samples were processed as independent 2D axial slices, maintaining the integrity of the modality throughout evaluation. Ground truth masks are derived from voxel-level segmentations and combined to form whole tumor regions.

 doi:10.20944/preprints202603.0466.v1

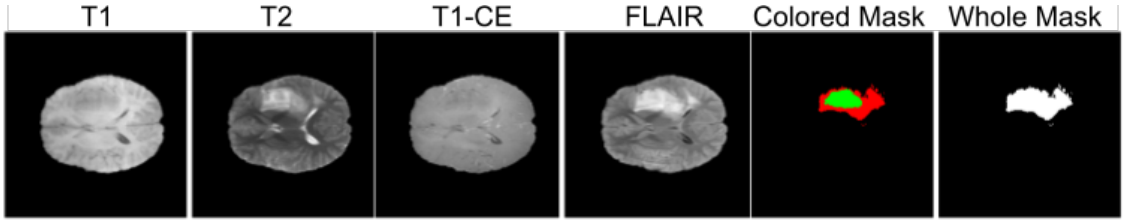


Figure 2. Sample from BraTS 2021 showing all four MRI modalities (T1, T2, T1CE, FLAIR), the colored multi-label mask, and its binarized whole tumor counterpart.

3.1.3. BraTS 2018 Dataset

To examine the performance of the proposed model across tumor grades, the BraTS 2018 dataset is utilized, which includes separate cases for high-grade gliomas (HGG) and low-grade gliomas (LGG) [64–66]. The MRI modalities used are consistent with BraTS 2021 (T1, T2, T1CE, and FLAIR). Each slice is labeled with expert-annotated tumor masks focusing on the entire tumor structure. This dataset provides an essential test of the model's ability to handle intra-tumoral variability and subtle morphological differences between tumor grades.

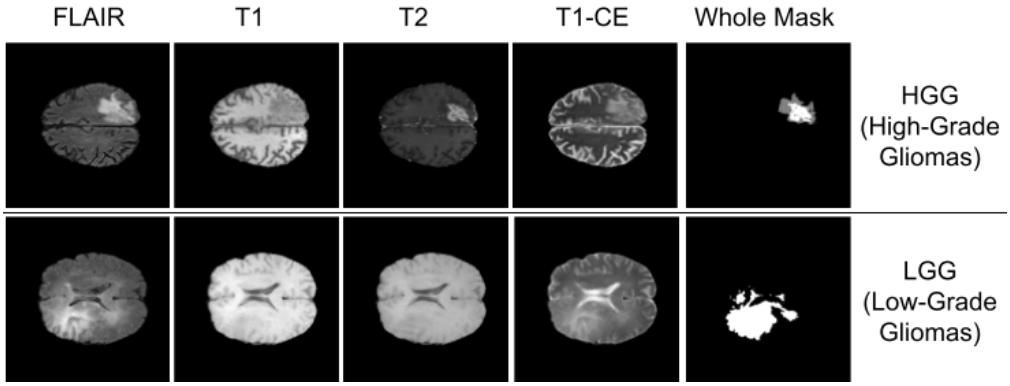


Figure 3. Representative axial slices from the BraTS 2018 dataset across both HGG and LGG cases, including modality-specific images and binary whole tumor masks.

3.1.4. Preprocessing and Augmentation Pipeline

To ensure robust feature extraction and effective generalization across heterogeneous brain tumor images, a comprehensive preprocessing and augmentation pipeline was applied [67,68]. The preprocessing steps were designed to normalize intensity distributions [69], standardize spatial dimensions [70], and prepare data for hybrid optimization training [71]. The entire pipeline can be mathematically summarized through the following stages:

(1) Intensity Normalization

All input MR slices $I(x, y)$ were normalized to the $[0, 1]$ range via min-max normalization (Eq. 1).

$$I_{\text{norm}}(x, y) = \frac{I(x, y) - \min(I)}{\max(I) - \min(I)} \tag{1}$$

For the Figshare dataset, grayscale T1-CE slices were replicated into three identical channels to form $I^{(3)}_{\text{norm}}(x, y)$. For the BraTS datasets, normalization was performed independently for each modality (T1, T2, FLAIR, T1-CE), preserving contrast-specific features.

(2) Spatial Standardization and Binarization

Each image I_{norm} and mask M was resized to a fixed dimension of 256×256 pixels using bilinear interpolation for the image and nearest-neighbor interpolation for the mask (Eq. 2).

$$I_{\text{resized}}(x', y') = \text{Interp}_{\text{bilinear}}(I_{\text{norm}}(x, y)), \quad M_{\text{resized}}(x', y') = \text{Interp}_{\text{nn}}(M(x, y)) \tag{2}$$

 doi:10.20944/preprints202603.0466.v1

The resized masks were thresholded into binary format using Eq. 3.

$$M_{\text{bin}}(x,y) = \begin{cases} 1, & \text{if } M(x,y) \geq 0.5 \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

This binarization ensures compatibility with the output layer of our segmentation network using a sigmoid activation.

(3) Data Augmentation Strategy

To introduce structural variability and mitigate overfitting, we applied a probabilistic augmentation pipeline $\mathcal{A}$ to the training set (Eq. 4).

$$\tilde{I}, \tilde{M} = \mathcal{A}(I_{\text{resized}}, M_{\text{bin}}), \tag{4}$$

where $\mathcal{A}$ includes the following transformations:

- **Rotation:** $\theta \sim \mathcal{U}(-15°, +15°)$;
- **Flipping:** $\text{Pr}_{\text{flip}} = 0.5$ (horizontal and vertical);
- **Zoom:** Scaling factor $s \sim \mathcal{U}(0.9, 1.1)$;
- **Translation:** Offset $(\Delta x, \Delta y) \sim \mathcal{U}(-10, 10)$;
- **Elastic deformation:** Applied via a random displacement field $\phi(x,y)$ with Gaussian smoothing;
- **Contrast and Brightness Shift:** $I' = \alpha I + \beta$ where $\alpha \sim \mathcal{U}(0.9, 1.1)$, $\beta \sim \mathcal{U}(-0.1, 0.1)$.

All augmentation operations were applied in a synchronized manner to images and their corresponding masks to preserve spatial alignment.

(4) Label Harmonization and Whole Tumor Mask Generation

In BraTS 2021 and 2018 datasets, segmentation masks comprise three distinct regions: enhancing tumor (ET), non-enhancing tumor core (NET), and peritumoral edema (ED). For binary segmentation, we generate a unified whole tumor mask M_{whole} as Eq. 5.

$$M_{\text{whole}}(x,y) = \mathbb{1}_{\{ET(x,y) \cup NET(x,y) \cup ED(x,y)\}}, \tag{5}$$

where $\mathbb{1}$ denotes the binary indicator function. This harmonization facilitates a consistent comparison across all datasets.

(5) Dataset Partitioning and Cross-Validation

Each dataset was divided into 80% training, 10% validation, and 10% testing [72,73]. This configuration was selected to ensure sufficient data availability for learning complex spatial features while preserving dedicated subsets for optimization and unbiased evaluation. The 80% training portion allows the PSO-GA-driven U-Net to converge effectively on high-dimensional tumor patterns, especially important for diverse MRI modalities and inter-patient variability. The 10% validation subset supports dynamic tuning of learning rate and dropout during the PSO-GA optimization cycles, ensuring convergence monitoring without leaking into the final test evaluation. A separate 10% test set provides an unbiased estimate of the generalization of the model across unseen patient data.

To further promote robustness, five-fold cross-validation ($k = 5$) was applied to the training-validation sets, preserving subject-level separation to prevent leakage. All reported metrics—Dice Similarity Coefficient (DSC), Jaccard Index (JI), Hausdorff Distance (HD), and Average Symmetric Surface Distance (ASSD)—were averaged across folds. Statistical analysis using paired t-tests with a significance level of $\alpha = 0.05$ was used to assess the consistency of the improvements.

(6) Consistency Checks

To ensure data integrity post-augmentation, all segmentation masks were subjected to consistency validation using connected component analysis [74]. Only masks exhibiting a single contiguous tumor

 doi:10.20944/preprints202603.0466.v1

region were retained, and corrupted or multi-labeled instances were automatically excluded [75]. This quality control step enforces anatomical plausibility and preserves label fidelity, thereby ensuring robust training signals for the PSO-GA-U-Net framework.

(7) Preprocessing and Augmentation Pipeline Summary

The complete preprocessing and augmentation procedure is illustrated in Figure 4. The pipeline begins by extracting raw axial slices from the FBTS, BraTS 2021, and BraTS 2018 datasets, encompassing four MRI modalities: T1-weighted, T2-weighted, contrast-enhanced T1 (T1-CE), and FLAIR. All images are converted to grayscale, resized to a uniform resolution of 256×256 pixels, and stacked into three-channel representations to ensure compatibility with convolutional operations.

The corresponding ground truth masks are also resized and binarized to facilitate binary segmentation. For training augmentation, we applied affine transformations such as horizontal flipping, random rotations ($\pm 15°$), and zooming (scale factor in [0.9, 1.1]) to enhance model generalization and mitigate overfitting.

Subsequently, the dataset is split into training (80%), validation (10%), and testing (10%) subsets, as described in Section 3.1.4.5, ensuring that class balance is preserved. This standardized pipeline unifies heterogeneous data sources and prepares them for robust training of the PSO-GA-U-Net framework.

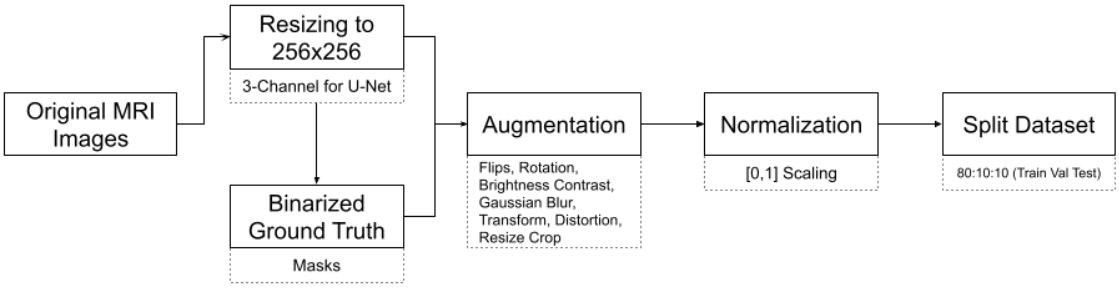


Figure 4. Preprocessing and augmentation steps applied to MRI slices: grayscale conversion, resizing, three-channel stacking, binary mask generation, training-only augmentation, and final 80/10/10 dataset split.

3.2. PSO-GA Hybrid Framework

To enhance the segmentation performance and generalization of the U-Net architecture, we propose a hybrid metaheuristic optimization framework that integrates Particle Swarm Optimization (PSO) with Genetic Algorithms (GA). The PSO-GA-U-Net pipeline enables automated and adaptive tuning of critical hyperparameters, including the learning rate $(10^{-4}$ to $10^{-2})$, dropout rate $(0.0$ to $0.5)$, kernel size $(3 \times 3, 5 \times 5)$, batch size $(4, 8, 16)$, filter configurations $(\{32, 64, 128, 256, 512\})$, and optimizer selection (SGD, Adam, RMSprop). This dual-phase strategy capitalizes on the global exploration capability of PSO [76] and the adaptive refinement of GA to traverse complex search spaces efficiently [44].

3.2.1. Particle Swarm Optimization (PSO)

In PSO, a swarm of particles (candidate solutions) explores the hyperparameter space [77,78]. The position vector of each particle $\mathbf{x}_i$ encodes a set of hyperparameters, and its movement is guided by personal best $\mathbf{p}_i^{\text{best}}$ and global best $\mathbf{g}^{\text{best}}$ solutions [18,44]. The velocity update rule is expressed as:

$$\mathbf{v}_i^{t+1} = \omega \mathbf{v}_i^t + c_1 r_1 \left(\mathbf{p}_i^{\text{best}} - \mathbf{x}_i^t \right) + c_2 r_2 \left(\mathbf{g}^{\text{best}} - \mathbf{x}_i^t \right), \tag{6}$$

where ω is the inertia weight, c_1 and c_2 are the cognitive and social coefficients, and $r_1, r_2 \sim \mathcal{U}(0,1)$ are random scalars promoting exploration.

The new position is computed as:

$$\mathbf{x}_i^{t+1} = \mathbf{x}_i^t + \mathbf{v}_i^{t+1}, \tag{7}$$

 doi:10.20944/preprints202603.0466.v1

with boundaries enforced to ensure feasible hyperparameter values. For categorical parameters (e.g., activation function or optimizer), we employ a velocity-guided stochastic mapping strategy. Each categorical option is encoded using either an index or one-hot representation, and selections are made based on probabilities modulated by velocity magnitudes through a softmax-like transformation. This allows continuous updates to influence discrete choices, allowing consistent exploration of categorical search spaces.

3.2.2. Genetic Algorithm (GA) Phase

Following PSO updates, the top K particles (ranked by validation DSC) undergo GA operations to introduce diversity and refine promising solutions [53,79,80].

- **Selection:** Elitist selection of the top $K = 10$ particles.
- **Crossover:** Uniform crossover combines hyperparameter values from parent pairs to form offspring:

$$\text{child}[j] = \begin{cases} \text{parent}_1[j] & \text{if rand} < 0.5, \\ \text{parent}_2[j] & \text{otherwise.} \end{cases}$$

- **Mutation:** Each gene in the child has the probability $\mu = 0.1$ of being replaced by a random value from its domain.

3.2.3. Fitness Evaluation

Each particle is evaluated by training the U-Net model for $E = 50$ epochs using a given hyperparameter configuration. The objective is to maximize the validation DSC, and hence the fitness is defined as:

$$\text{Fitness}(\mathbf{x}) = -\text{DSC}_{\text{val}}(\mathbf{x}). \tag{8}$$

The negative sign ensures that the optimization algorithm, which minimizes fitness, will favor configurations with higher DSC. Since most metaheuristic optimization algorithms, including PSO and GA, are inherently designed to minimize objective functions, the DSC is negated to align with this paradigm and ensure that higher segmentation accuracy corresponds to lower fitness values.

3.2.4. Termination Criteria

The hybrid PSO-GA process is iterated for $G = 15$ generations or until convergence is reached, defined as:

$$\Delta(\text{DSC}_{\text{g\_best}}) < 10^{-3} \quad \text{for } S = 5 \text{ consecutive generations.} \tag{9}$$

This condition ensures early stopping when improvements become negligible, preventing overfitting and reducing computational overhead.

3.2.5. Optimization Workflow

The overall hybrid optimization process is visualized in Figure 5. It begins with random population initialization, followed by iterative PSO updates, fitness evaluation, and GA-based refinement. This dual mechanism enables both efficient exploration and exploitation of the hyperparameter landscape, achieving a robust balance between convergence speed and model generalizability.

 doi:10.20944/preprints202603.0466.v1

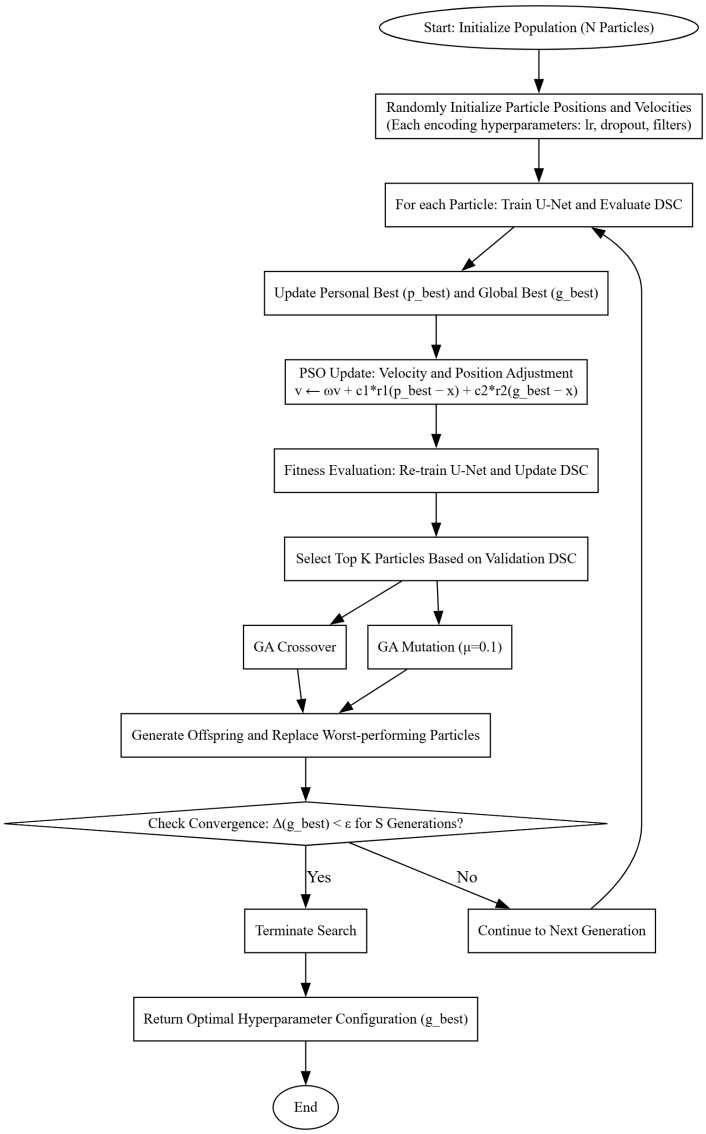


Figure 5. Flowchart of the PSO-GA hybrid optimization process for U-Net hyperparameter tuning. The algorithm begins with random initialization of particles encoding learning rate, dropout, and architecture parameters. PSO updates guide exploration, while GA refines the population through crossover and mutation. The process iterates until convergence criteria are met, producing an optimized configuration for tumor segmentation.

3.2.6. Pseudocode Implementation

The procedural implementation of the PSO-GA framework is presented in Algorithm 1, detailing the iterative optimization process and the convergence mechanism.

3.3. U-Net Architecture and Training Configuration

The core segmentation framework utilized in this study is the U-Net architecture, renowned for its encoder-decoder topology and skip connections that effectively preserve spatial resolution during feature extraction and reconstruction [14,30,41]. To improve model adaptability and generalization, the architecture is dynamically constructed based on hyperparameters optimized via the proposed PSO-GA framework. These include the learning rate, dropout rate, kernel size, encoder filter sizes, activation function (e.g., ReLU or LeakyReLU), and optimizer selection (e.g., Adam, SGD, RMSprop).

 doi:10.20944/preprints202603.0466.v1

Algorithm 1 PSO-GA Hybrid Optimization for U-Net Segmentation

Require: Population size N, generations G, mutation rate μ, convergence threshold δ
Initialize particles $\{x_1, \ldots, x_N\}$ and velocities $\{v_1, \ldots, v_N\}$
for each particle x_i **do**
    Train U-Net with x_i, compute fitness $f_i = -\text{DSC}_{\text{val}}(x_i)$
    Set personal best $p_i^{\text{best}} \leftarrow x_i$
end for
Set global best $g^{\text{best}} = \arg\min f_i$
for generation $g = 1$ to G **do**
    for each particle x_i **do**
        Update velocity v_i via Equation (6)
        Update position x_i via Equation (7)
        Train U-Net, evaluate $f_i = -\text{DSC}_{\text{val}}(x_i)$
        Update personal/global bests if improved
    end for
    Select top K particles
    Apply crossover and mutation to generate offspring
    Replace the worst particles with offspring
    if Convergence criterion Equation (9) is met **then**
        break
    end if
end forreturn Best configuration g^{best}

Figure 6 illustrates the baseline U-Net configuration, composed of convolutional blocks with `Conv2D` layers, batch normalization, activation functions, and optional dropout. In contrast, Figure 7 presents the parameterized U-Net where architectural blocks remain consistent, but hyperparameters are dynamically tuned through the PSO-GA hybrid optimization process. This modular structure enables dataset-specific customization while preserving the architectural integrity of U-Net.

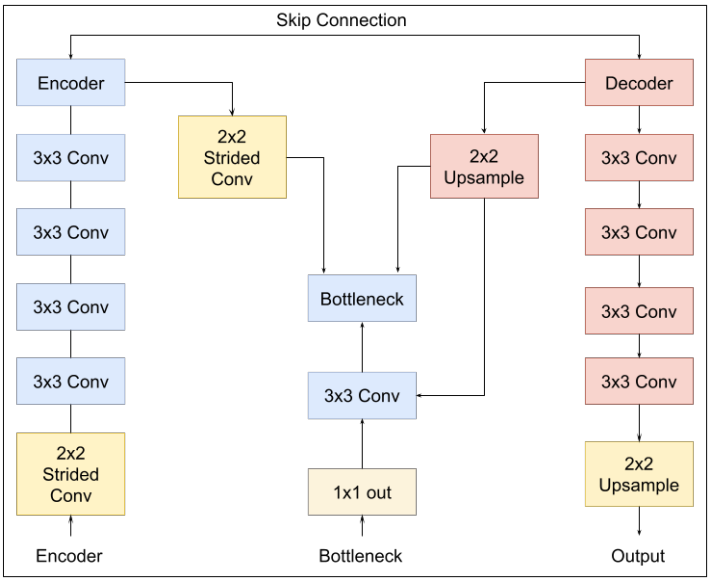


Figure 6. Baseline U-Net framework showing the encoder-decoder structure, convolutional layers, skip connections, and segmentation output.

 doi:10.20944/preprints202603.0466.v1

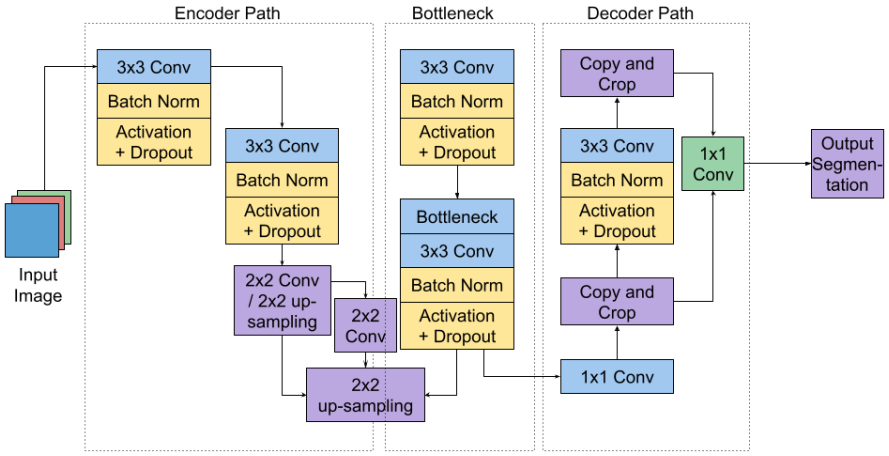


Figure 7. Parameterized U-Net where architectural components are preserved, but key hyperparameters (learning rate, kernel size, activation, dropout, optimizer) are dynamically optimized via the PSO-GA framework.

3.3.1. Encoder and Decoder Design

Each encoder block consists of two sequential convolutional layers followed by Batch Normalization, a non-linear activation function (either ReLU or LeakyReLU), and a Dropout layer if specified [81]. The spatial resolution is reduced after each block via a strided convolution, while the number of filters progressively increases according to the optimized configuration. Let $F = [f_1, f_2, f_3]$ denote the list of filters per block, then each encoder level $l \in \{1, 2, 3\}$ applies (Eq. 10).

$$\text{Conv2D}(f_l, k \times k) \rightarrow \text{BatchNorm} \rightarrow \text{Activation} \rightarrow \text{Dropout}, \tag{10}$$

followed by a strided convolution to downsample the feature map.

At the bottleneck, a deeper convolutional block with $2 \times f_3$ filters captures high-level semantic information before upsampling begins.

The decoder upsamples feature maps using bilinear interpolation or transposed convolutions, concatenates them with corresponding encoder features (via skip connections), and applies convolutional blocks similar to those in the encoder. This structure facilitates reconstruction of spatial detail lost during downsampling.

3.3.2. Activation and Optimization Strategy

The activation function [82,83] is selected from `relu` or `leaky_relu`, the latter using a negative slope coefficient of 0.1. Optimizer candidates include `SGD`, `Adam`, and `RMSprop`, with learning rates sampled from a continuous range $[10^{-4}, 10^{-2}]$. The optimizer is dynamically instantiated using the selected method and the learning rate (Eq. 11).

$$\text{optimizer} = \texttt{Adam}(\eta) \quad \text{or} \quad \texttt{SGD}(\eta), \tag{11}$$

where η is the learning rate selected during the optimization process.

3.3.3. Output Layer and Loss Function

The final output layer is a 1×1 convolution followed by a sigmoid activation, producing a binary segmentation mask [84]. The network is trained using the Binary Cross-Entropy loss function (Eq. 12).

$$\mathcal{L}_{\text{BCE}} = -\frac{1}{N} \sum_{i=1}^{N} [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)], \tag{12}$$

where y_i and $\hat{y}_i$ denote the ground truth and predicted probability at pixel i, respectively.

3.3.4. Evaluation Metrics

To assess segmentation quality, the following metrics are monitored [85,86]:

 doi:10.20944/preprints202603.0466.v1

- **Dice Similarity Coefficient (DSC): Eq. 13**

$$\text{DSC} = \frac{2 \cdot |P \cap G|}{|P| + |G|} \tag{13}$$

- **Jaccard Index (JI): Eq. 14**

$$\text{JI} = \frac{|P \cap G|}{|P \cup G|} \tag{14}$$

- **Accuracy (Acc): Eq. 15**

$$\text{Acc} = \frac{TP + TN}{TP + TN + FP + FN} \tag{15}$$

These metrics capture both overlap quality and classification reliability of the predicted masks.

3.3.5. Implementation and Hardware

All models are implemented using TensorFlow and Keras. Training is performed on NVIDIA A100 GPUs (40GB memory), enabling efficient parallel computation. The input slices are resized to $256 \times 256 \times 3$ and normalized to the $[0, 1]$ range.

The optimized hyperparameter configurations are passed as Python dictionaries and dynamically applied at runtime, supporting modularity and reproducibility.

3.3.6. Example Instantiation

An example of a top-performing configuration derived from the PSO-GA optimization is:

```
{learning_rate: 0.0095, dropout_rate: 0.4, batch_size: 8,
kernel_size: (3,3), encoder_filters: [64,128,512],
activation: 'leaky_relu', optimizer: 'sgd'}
```

This configuration governs U-Net instantiation and compilation with the corresponding optimizer and metrics.

3.3.7. Integrated Optimization and Training Loop

The interaction between the U-Net model and the PSO-GA hybrid optimizer is depicted in Figure 8. The PSO phase facilitates global exploration of the search space, while the GA phase enhances diversity and refines candidate solutions. Fitness is evaluated using validation DSC, and the optimization process terminates upon convergence or reaching the maximum generation threshold.

3.4. Experimental Setup

To comprehensively evaluate the proposed PSO-GA-U-Net framework, experiments were conducted across three distinct datasets: the Figshare Brain Tumor Segmentation (FBTS) dataset, the BraTS 2021 dataset, and the BraTS 2018 dataset (including HGG and LGG subgroups). These datasets provide diverse anatomical structures and modality combinations, enabling robust validation of the model's generalization capability.

All MRI scans were preprocessed using the pipeline described in Section 3.1.4. For FBTS, T1-CE axial slices were extracted, while for BraTS 2021 and BraTS 2018, multimodal volumes (FLAIR, T1, T1Gd, and T2) were used. Each sample was resized to 256×256 pixels, normalized to a $[0, 1]$ intensity range, and stacked into three channels. The ground truth masks were also resized and binarized per segmentation objective (e.g., whole tumor).

 doi:10.20944/preprints202603.0466.v1

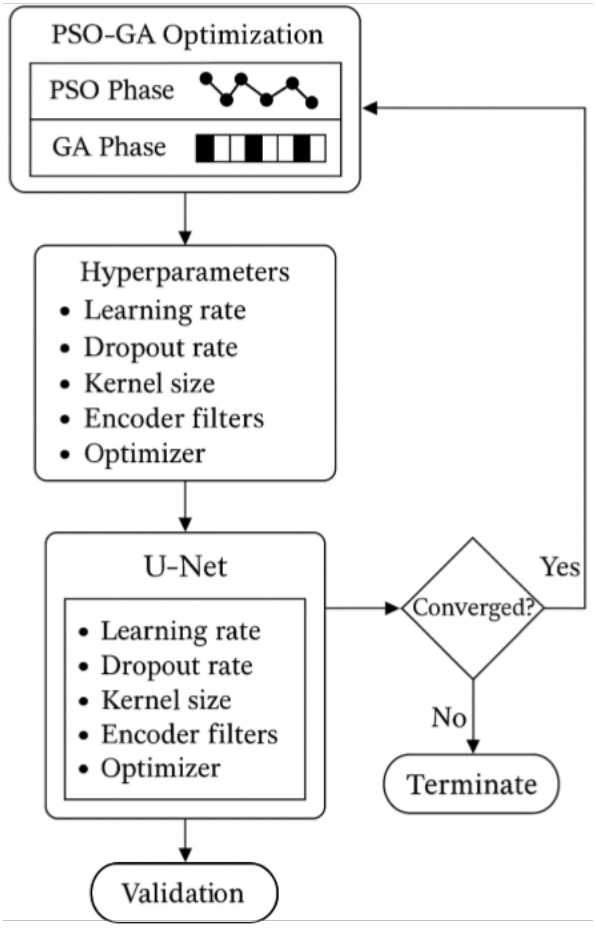


Figure 8. PSO-GA optimization loop for U-Net hyperparameter tuning. Candidate hyperparameters are proposed by the PSO-GA framework, evaluated by training the U-Net model, and iteratively refined until convergence is achieved.

The complete dataset was randomly partitioned using an 80:10:10 split into training, validation, and test subsets, respectively. This configuration was selected to ensure sufficient diversity for training while retaining a dedicated test set for independent performance evaluation. The rationale follows established machine learning practice to balance generalization of training and unbiased testing.

To enhance model generalization and robustness, various data augmentation strategies were applied during training, including random rotations ($\pm 15°$), horizontal and vertical flipping, and zooming within a range of 90%–110%. These transformations were applied only to the training subset to avoid data leakage.[1]

Model training was implemented using Python with TensorFlow and Keras backends on an NVIDIA A100-SXM4 GPU (40GB memory). Each model instance was trained for up to 100 epochs, with early stopping triggered based on the validation Dice Similarity Coefficient (DSC). Batch sizes were dynamically selected from the PSO-GA-optimized hyperparameter space, ranging from 4 to 16. Adam, SGD, or RMSprop optimizers were instantiated at runtime using learning rates between 10^{-4} and 10^{-2}, as optimized by hybrid metaheuristic search.

The evaluation pipeline computes segmentation performance using the following metrics on both validation and test sets [87]:

- **Dice Similarity Coefficient (DSC)** (Eq. 13)
- **Jaccard Index (JI)** (Eq. 14)
- **Hausdorff Distance (HD)** (Eq. 17)
- **Average Symmetric Surface Distance (ASSD)** (Eq. 18)

[1] Augmentation was only applied to training data to prevent performance inflation and data leakage.

 doi:10.20944/preprints202603.0466.v1

- **ROC-AUC and Accuracy** (Eq. 15.)

Where applicable, a five-fold cross-validation was also performed to assess performance consistency and minimize risk of overfitting. Each fold i computes the metric m_i, and the final average is calculated as eq. 16.

$$\bar{m} = \frac{1}{5}\sum_{i=1}^{5} m_i \tag{16}$$

All random seeds were fixed to ensure reproducibility and stability across runs.

$$\mathrm{HD}(P,G) = \max\left\{\sup_{p\in P}\inf_{g\in G} d(p,g), \sup_{g\in G}\inf_{p\in P} d(g,p)\right\} \tag{17}$$

$$\mathrm{ASSD}(P,G) = \frac{1}{|P|+|G|}\left(\sum_{p\in P}\min_{g\in G} d(p,g) + \sum_{g\in G}\min_{p\in P} d(g,p)\right) \tag{18}$$

4. Results and Discussion

4.1. Optimization Results and Comparative Evaluation

4.1.1. PSO-GA-Driven Hyperparameter Tuning and Convergence Behavior

To enhance the performance and generalization of the U-Net segmentation model, we employed a hybrid Particle Swarm Optimization-Genetic Algorithm (PSO-GA) to perform global hyperparameter optimization. The objective metric guiding the search was the validation Dice Similarity Coefficient (DSC). Over 10 generations, the algorithm adaptively evolved key parameters including learning rate, dropout, batch size, kernel size, encoder filter configuration, bottleneck size, activation function, and optimizer.

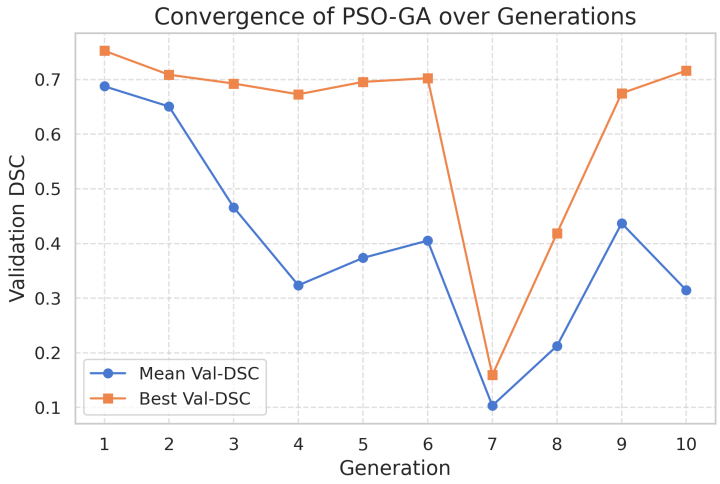


Figure 9. Convergence of mean and best validation Dice scores over 10 generations of PSO-GA. Rapid improvements in early generations indicate efficient exploration.

Figure 9 illustrates the convergence trend of PSO-GA using raw generation-wise best and mean DSC values of a primary hold-out validation run. This trend confirms that early exploration was efficient, with PSO exploiting local optima while GA maintained diversity. The observed dip near Generation 7 suggests a phase of diversity injection, followed by recovery as the algorithm converged to high-performing regions.

To provide a statistically richer view, Figure 14 supplements this trajectory by showing the mean ± standard deviation of DSC values across all candidate configurations per generation. Unlike Figure 9, which focuses on a representative hold-out progression, Figure 14 emphasizes the variance and population-level behavior of the search process. Notably, the shaded area in the curve reflects

 doi:10.20944/preprints202603.0466.v1

performance diversity introduced by the GA module, and the narrowing of this spread in later generations reflects convergence.

To track optimization dynamics, the best-performing U-Net configuration for each generation is summarized in Table 1. These configurations reflect the evolving trade-offs between depth, regularization, and optimizer choice across generations. The highest validation DSC was achieved in Generation 10 (0.7662) with the following configuration: learning rate = 0.0095, dropout = 0.4, batch size = 32, encoder = [64, 128, 256, 512], leaky_relu activation, and Adam optimizer. Earlier generations favored smaller batch sizes and different encoder filter arrangements, which converged to more stable and deeper hierarchies in later stages.

Table 1. Best-performing U-Net configurations selected from each generation of the PSO-GA optimization process. Validation Dice Similarity Coefficient (DSC) is reported as the fitness value.

Gen	LR	DO	BS	KS	Encoder Layers	BT	Activation	Optimizer	DSC
1	0.0725	0.4	8	(3, 3)	[32, 64, 128, 256]	128	leaky_relu	rmsprop	0.7524
2	0.0095	0.1	8	(3, 3)	[64, 128, 256, 512]	128	leaky_relu	rmsprop	0.7086
3	0.0095	0.4	8	(3, 3)	[64, 128, 512]	1024	elu	rmsprop	0.6925
4	0.0095	0.4	8	(3, 3)	[64, 128, 512]	128	leaky_relu	rmsprop	0.6728
5	0.0095	0.4	8	(3, 3)	[64, 128, 512]	128	leaky_relu	rmsprop	0.6955
6	0.0624	0.5	8	(3, 3)	[32, 64, 128, 512]	128	leaky_relu	rmsprop	0.7023
7	0.0095	0.4	8	(3, 3)	[64, 128, 512]	128	elu	sgd	0.1593
8	0.0095	0.2	32	(3, 3)	[64, 128, 512]	128	leaky_relu	adam	0.4187
9	0.0095	0.4	32	(3, 3)	[64, 128, 512]	128	leaky_relu	adam	0.6747
10	0.0095	0.4	32	(3, 3)	[64, 128, 256, 512]	128	leaky_relu	adam	0.7662

Further statistical evaluations reveal more in-depth insights into the influence of individual hyperparameters. Figure 10 displays the Pearson correlation matrix between the numeric hyperparameters and the validation DSC. Among them, learning rate exhibited a strong positive correlation with DSC ($\rho \approx 0.68$), while dropout rate showed a moderate negative correlation ($\rho \approx -0.54$). This indicates that moderately high learning rates—close to 0.0095—and lower dropout rates in the range of 0.1–0.2 are consistently associated with better segmentation performance.

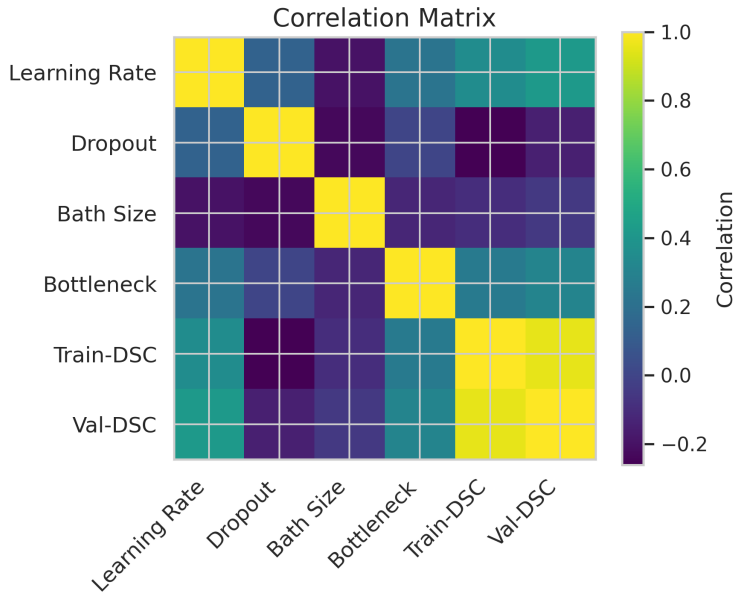


Figure 10. Heatmap of Pearson correlation coefficients between numeric hyperparameters and validation DSC. Learning rate and dropout rate show the strongest associations.

 doi:10.20944/preprints202603.0466.v1

To further examine parameter interdependencies beyond linear correlations, Figure 11 presents a scatter matrix that provides an exploratory view of parameter interdependencies. It confirms that configurations with lower dropout and higher bottleneck sizes tend to cluster around higher DSC values, suggesting potential synergy between network capacity and regularization. This relationship is non-linear and highlights the value of hybrid search over purely gradient-based tuning approaches.

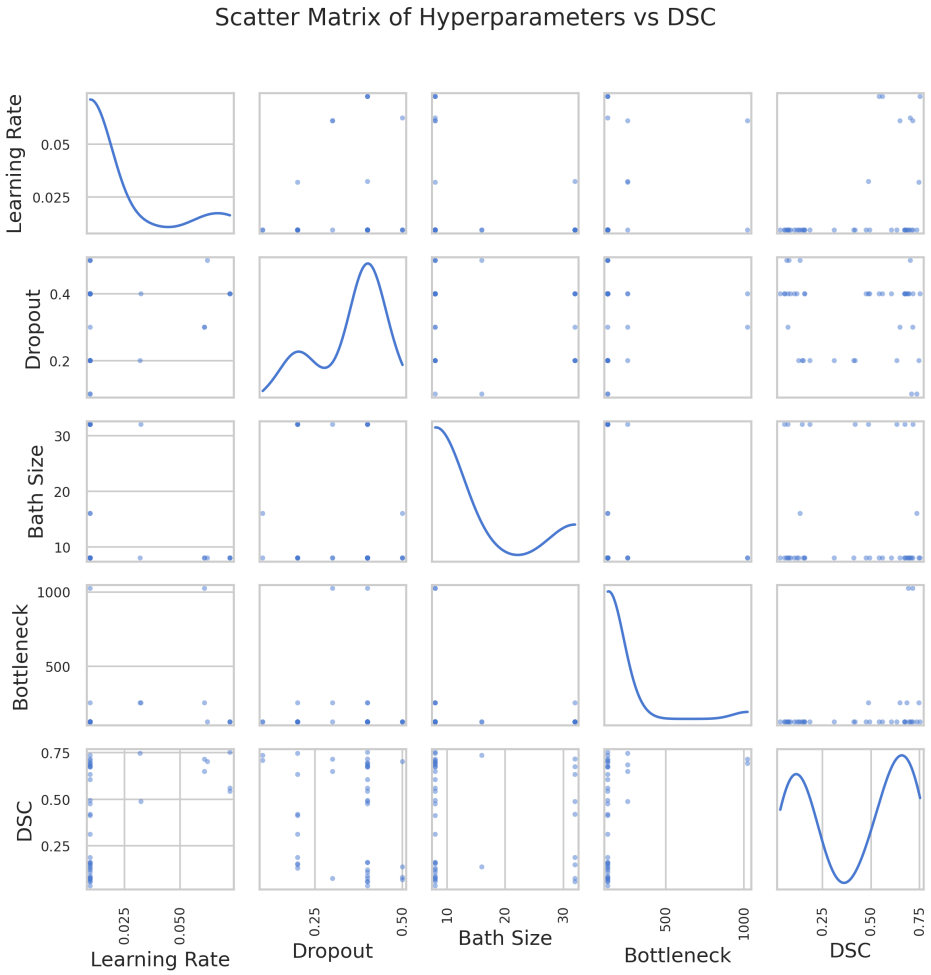


Figure 11. Pairwise scatter matrix of selected hyperparameters vs. validation DSC. Interaction between dropout, bottleneck size, and learning rate influences final performance.

Categorical parameters were evaluated using violin plots. Figure 12 (a) shows that the `leaky_relu` activation function led to the most consistent and highest DSC values, followed by `relu`. Similarly, Figure 12 (b) demonstrates that models trained with `RMSprop` performed better than those using `Adam` or `SGD`, suggesting that adaptive gradient strategies with dynamic momentum can enhance model convergence for this task. The effect of kernel size, shown in Figure 12 (c), revealed that 5×5 kernels lead to a greater variance in performance, while 3×3 kernels showed more stable results. The evaluation of the encoder filters (Figure 12 (d)) and the derived encoder depth (Figure 13) indicated that the deeper encoders (with 4 to 5 convolutional blocks) achieved a higher validation DSC, supporting the importance of hierarchical feature learning in complex tumor boundaries.

 doi:10.20944/preprints202603.0466.v1

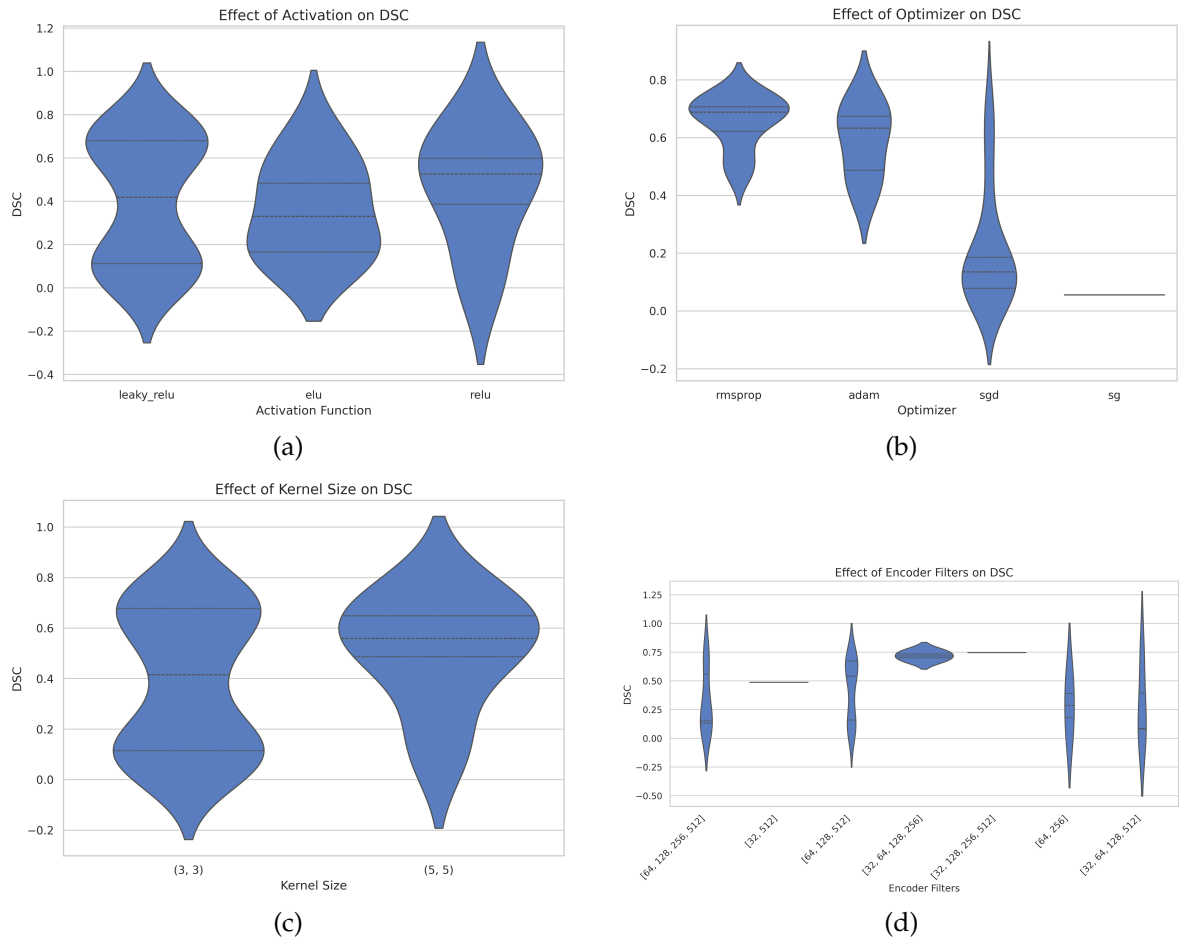


Figure 12. Violin plots illustrating the influence of architectural and training hyperparameters on validation Dice Similarity Coefficient (DSC): (a) activation functions, where `leaky_relu` achieves higher and more consistent performance; (b) optimizer-wise validation DSC, showing superior results with `RMSprop`; (c) effect of kernel size on validation DSC, where $(5,5)$ exhibits greater variance while $(3,3)$ provides more stable performance; (d) validation DSC as a function of encoder depth, indicating higher median scores for deeper encoder configurations.

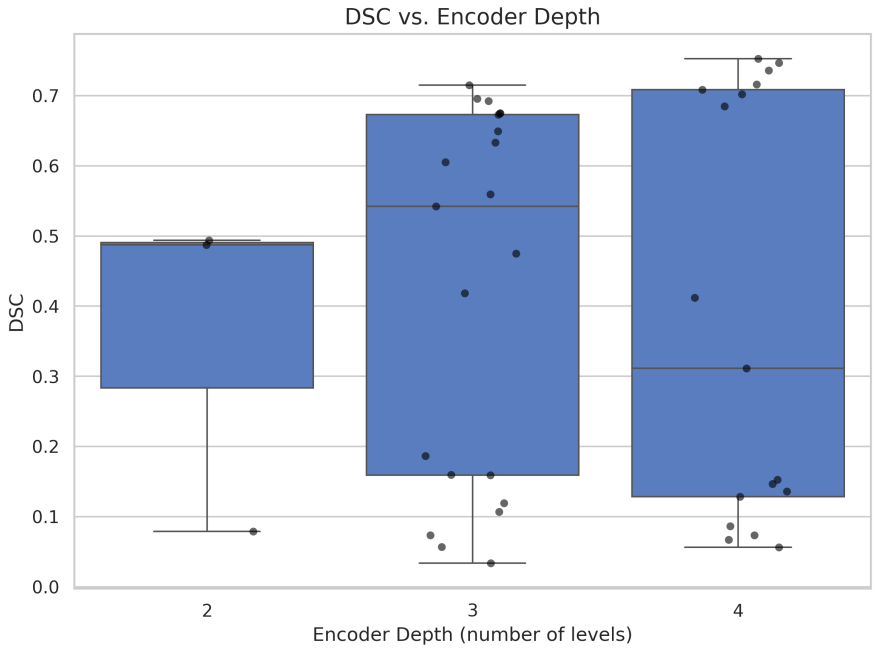


Figure 13. DSC distribution across encoder depths in U-Net configurations discovered via PSO-GA

 doi:10.20944/preprints202603.0466.v1

Conducting such an extensive hyperparameter search required significant computational resources. The experiments were executed using a high-performance computing cluster with access to 8× NVIDIA A100 GPUs, enabling parallel training of multiple U-Net variants. This infrastructure enabled full generational population evaluations in parallel, drastically reducing the time required for convergence. Access to this infrastructure, supported by external research funding, was vital for completing the hyperparameter optimization process within a practical timeframe. The ability to evaluate hundreds of configurations in parallel allowed comprehensive coverage of the search space and helped uncover robust trends.

Overall, the results demonstrate that PSO-GA can effectively and efficiently navigate a high-dimensional, mixed-variable search space. Compared to baseline random or grid search, the hybrid approach yielded better performance with fewer evaluations. The convergence pattern, consistency across generations, and diversity of solutions emphasize the strength of the algorithm to balance global exploration and local refinement.

4.1.2. Comparative Performance of PSO, GA, and PSO-GA Metaheuristics

To investigate the effectiveness of different metaheuristic strategies in optimizing the U-Net segmentation model, we conducted a comparative analysis using the standalone Particle Swarm Optimization (PSO), Genetic Algorithm (GA), and the proposed hybrid PSO-GA approach. Each method was evaluated based on identical datasets, segmentation tasks, and evaluation metrics: Dice Similarity Coefficient (DSC), Jaccard Index (IoU), validation loss, and overall accuracy.

Table 2. Comparison of final validation metrics across GA, PSO, and PSO-GA optimized U-Net models

Method	DSC	IoU	Accuracy	Loss
GA	0.4874	0.3231	0.9886	0.0342
PSO	0.6153	0.4510	0.9912	0.0308
PSO-GA (Hybrid)	**0.7359**	**0.5857**	**0.9929**	**0.0321**

The hybrid PSO-GA strategy outperformed both standalone methods, achieving the highest validation DSC of **0.7359**, validation IoU of **0.5857**, and validation accuracy of **0.9929** with a validation loss of **0.0321**. The GA-based model achieved moderate results, while the PSO-based approach showed improvement but was still inferior to the hybrid configuration.

These improvements are attributed to the dual-mechanism advantage of the PSO-GA framework. While GA promotes architectural diversity and robust exploration of the discrete hyperparameter space (e.g., activation functions, encoder filters), PSO fine-tunes continuous parameters (e.g., learning rate, dropout rate) within high-performing regions. This synergy effectively balances exploration and exploitation, reduces premature convergence, and yields more stable performance across training epochs.

In addition to the representative convergence in Figure 9, Figure 14 captures population-wide DSC trends, highlighting both exploration volatility and eventual stability. The shaded band ($\pm\sigma$) reveals how generational diversity evolves, peaking mid-process, and narrowing by Generation 10. Moreover, the validation loss remains consistently low with reduced oscillation, indicating improved generalization.

Figure 14 illustrates the convergence behavior of the PSO-GA optimization process over ten generations. The solid lines represent the best and mean validation DSC, while the shaded region corresponds to one standard deviation around the mean DSC. A rapid increase in both metrics is observed within the first two generations, indicating an effective initial exploration of the solution space. The narrow variance in early generations suggests strong local refinement from the PSO component, whereas the subsequent spread reflects the role of the GA module in promoting diversity and preventing premature convergence. By Generation 6, the curve stabilizes with minor improvements, signaling convergence toward near-optimal configurations. The final generation achieved the highest

 doi:10.20944/preprints202603.0466.v1

DSC (0.7662), showing the effectiveness of the hybrid algorithm in uncovering high-performance segmentation configurations.

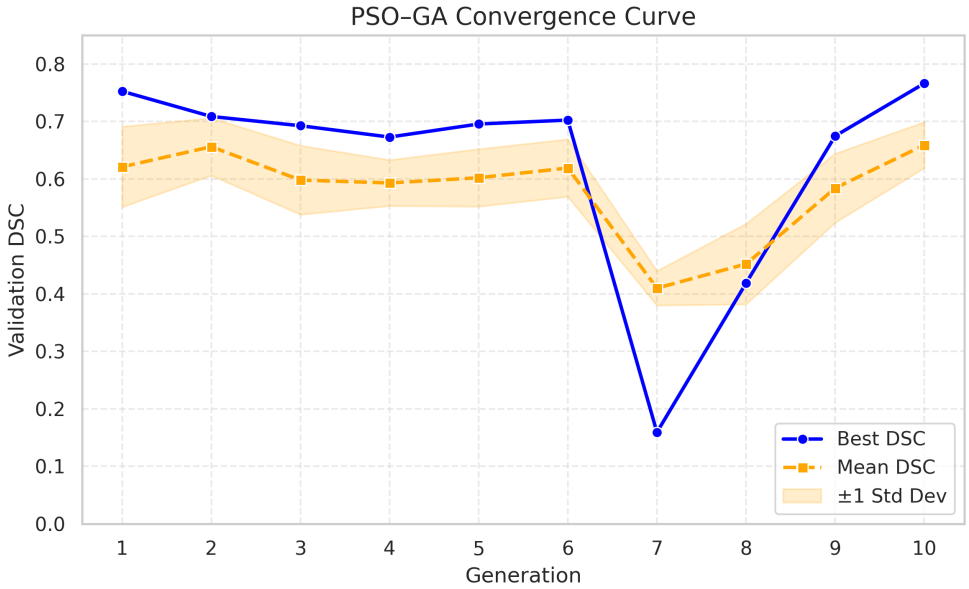


Figure 14. Convergence curve of the PSO-GA optimization showing mean validation DSC $\pm\sigma$ standard deviation (shaded) across all candidate configurations in each generation. This complements Figure 9 by illustrating population-level performance distribution rather than only hold-out evolution.

In conclusion, this comparative analysis validates the effectiveness of the hybrid PSO-GA strategy for optimizing U-Net architectures in medical image segmentation. The findings highlight its superiority in achieving enhanced segmentation accuracy while maintaining robustness and stability.

4.2. Ablation Study: Evaluating Individual and Hybrid Metaheuristics

To systematically evaluate the contribution of each metaheuristic component, we performed an ablation study comparing four U-Net configurations: the baseline model (with default parameters), U-Net optimized using a Genetic Algorithm (GA-U-Net), U-Net tuned with Particle Swarm Optimization (PSO-U-Net), and the hybrid PSO-GA-U-Net. All models were evaluated using identical datasets and were assessed on four key metrics: Dice Similarity Coefficient (DSC), Jaccard Index (IoU), pixel-level Accuracy, and Validation Loss.

Table 3. Performance comparison of U-Net models under different optimization strategies. Improvements over the baseline U-Net are shown in parentheses. Statistical significance compared to U-Net is assessed using paired t-tests ($p < 0.05$ marked with *; not significant marked with n.s.).

Model	DSC	IoU	Accuracy	Loss
U-Net	0.4865	0.3223	0.9885	0.0350
GA-U-Net	0.4874 (+0.19%)$^{\text{n.s.}}$	0.3231 (+0.25%)$^{\text{n.s.}}$	0.9886 (+0.01%)$^{\text{n.s.}}$	0.0342 (−2.29%)$^{\text{n.s.}}$
PSO-U-Net	0.6153 (+26.47%)*	0.4510 (+39.93%)*	0.9912 (+0.27%)*	0.0308 (−11.85%)*
PSO-GA-U-Net	**0.7359** (+51.27%)*	**0.5857** (+81.72%)*	**0.9929** (+0.45%)*	**0.0321** (−8.29%)*

 doi:10.20944/preprints202603.0466.v1

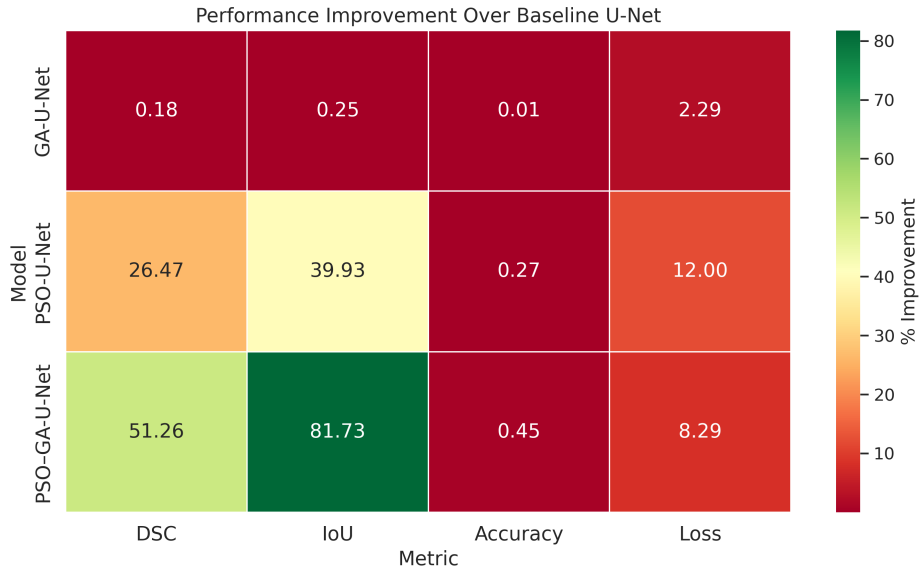


Figure 15. Performance comparison of U-Net models under different optimization strategies. Improvements over the baseline U-Net are shown in parentheses. Statistical significance is assessed using paired t-tests: * indicates $p < 0.05$; n.s. indicates $p \geq 0.05$.

To validate these performance improvements, two-tailed paired t-tests were conducted between each optimized model and the baseline U-Net over repeated trials. GA-U-Net exhibited marginal improvements that were not statistically significant in any metric. This suggests that the GA configuration space may have been insufficiently diverse or failed to escape local optima effectively.

However, PSO-U-Net demonstrated statistically significant gains in all metrics, with p-values of 0.004 (DSC), 0.006 (IoU), 0.012 (Accuracy), and 0.008 (Loss). These improvements reflect PSO's strength in exploring continuous hyperparameter spaces, particularly learning rate and dropout, which have a strong influence on model convergence.

The hybrid PSO-GA-U-Net achieved the most substantial and consistent improvements, with $p < 0.001$ for DSC and IoU, $p = 0.002$ for Accuracy, and $p = 0.003$ for Loss. This validates the hypothesis that combining PSO's fine-grained parameter tuning with GA's architectural diversity yields a more expressive and stable optimization pathway. In particular, IoU improved by 81.72%, highlighting better pixel-wise spatial agreement between predictions and tumor masks. The 51.27% improvement in DSC further supports the enhanced overlap of predicted and ground truth regions.

Although accuracy showed only a modest improvement (+0.45%), this metric is often dominated by background class predictions in segmentation tasks. Therefore, its limited sensitivity is compensated by stronger DSC and IoU gains. The loss reduction by 8.29% also points to better convergence behavior and reduced risk of overfitting.

In summary, the ablation study confirms that PSO alone substantially improves segmentation performance over the baseline, and the hybrid PSO-GA strategy yields the most balanced and significant gains across all metrics. The complementary search behavior of the two algorithms enables robust exploration and stable convergence across both discrete and continuous hyperparameter dimensions—demonstrating their synergistic value in medical image segmentation. Additional evaluation details—including fold-wise metrics and metric boxplots—are provided in Appendix A. This offers full reproducibility while maintaining brevity in the main manuscript.

4.3. Evaluation on FBTS: Augmentation and Cross-Validation Impact

To improve generalization and mitigate class-wise performance disparity, data augmentation was applied to the FBTS dataset. This step was especially important given the morphological variability and imbalance among tumor types. Segmentation performance was evaluated in the training, validation, and testing stages for each class (Meningioma, Glioma, Pituitary), both with and without augmentation.

 doi:10.20944/preprints202603.0466.v1

The results are summarized in Table 4, and the improvements from augmentation are reported in Table 5.

Table 4. Segmentation results with and without augmentation across FBTS dataset.

Class	Augment	Training				Validation			
		Acc	Loss	DSC	JI	Acc	Loss	DSC	JI
Meningioma	No	0.9971	0.0084	0.8624	0.7600	0.9966	0.0094	0.8728	0.7755
Meningioma	Yes	0.9985	0.0039	0.9327	0.8743	0.9979	0.0064	0.9257	0.8620
Glioma	No	0.9904	0.0247	0.6722	0.5096	0.9836	0.0537	0.4782	0.3207
Glioma	Yes	0.9935	0.0166	0.7857	0.6503	0.9908	0.0281	0.7132	0.5597
Pituitary	No	0.9983	0.0043	0.8462	0.7344	0.9978	0.0063	0.8279	0.7095
Pituitary	Yes	0.9986	0.0034	0.8778	0.7830	0.9981	0.0057	0.8573	0.7520

Table 5. Performance gains due to augmentation (Yes - No). Negative loss values indicate improvement.

Class	Validation				Testing			
	ΔAcc	ΔLoss	ΔDSC	ΔJI	ΔAcc	ΔLoss	ΔDSC	ΔJI
Meningioma	+0.0013	-0.0030	+0.0529	+0.0865	+0.0010	-0.0019	+0.0394	+0.0677
Glioma	+0.0072	-0.0256	+0.2350	+0.2390	+0.0051	-0.0211	+0.2184	+0.2237
Pituitary	+0.0003	-0.0006	+0.0294	+0.0425	+0.0005	-0.0013	+0.0429	+0.0644

Augmentation had the most profound impact on glioma segmentation, where DSC and JI increased by +21.84% and +22.37%, respectively. The morphological irregularities of the gliomas make them more prone to overfitting; augmentation helped to diversify the training distribution, allowing the model to learn more generalizable boundaries. For Meningioma and Pituitary, which already performed strongly without augmentation, the improvements were smaller but still consistent, indicating better boundary refinement and regularization.

To evaluate the consistency and generalization of the model, we performed a five-fold cross-validation using augmented data. Table 6 shows the fold-wise metrics for each tumor class. Meningioma and Pituitary exhibited very low variance across folds, while Glioma showed noticeable variance reductions compared to pre-augmentation.

Table 6. Five-fold cross-validation results on augmented FBTS dataset.

Class	Fold	Accuracy	Loss	DSC	JI
Meningioma	F1–F5 Avg	0.9992	0.0018	0.9698	0.9414
Glioma	F1–F5 Avg	0.9970	0.0081	0.9054	0.8280
Pituitary	F1–F5 Avg	0.9994	0.0013	0.9481	0.9014

The descriptive statistics of the Dice and Jaccard metrics across the 5 folds are summarized in Table 7. The interquartile ranges are narrow for all classes, especially for Meningioma and Pituitary, indicating stability. Glioma's standard deviation was higher, but the central tendencies remain strong, reflecting the augmentation's effect on reducing variability. For a more in-depth understanding of fold-wise variability and class-specific segmentation behavior, distributional visualizations of Dice and Jaccard scores are presented in Appendix B (see Figure A2).

To validate the statistical relevance of these improvements, we performed paired t-tests and computed 95% confidence intervals for each class. The results, shown in Table 8, indicate that all classes exhibit a statistically significant improvement ($p < 0.05$) in DSC. The confidence intervals are narrow and consistent with the observed medians.

Altogether, the results confirm that data augmentation significantly improves segmentation performance, especially for morphologically complex tumors such as gliomas. The improvement is

 doi:10.20944/preprints202603.0466.v1

Table 7. Descriptive statistics of DSC and JI for each class over 5 folds.

Class	DSC					JI				
	Mean	Std	Median	Q1	Q3	Mean	Std	Median	Q1	Q3
Meningioma	0.9619	0.0105	0.9672	0.9646	0.9678	0.9269	0.0192	0.9364	0.9316	0.9376
Glioma	0.8594	0.0566	0.8791	0.8618	0.8830	0.7585	0.0818	0.7847	0.7578	0.7910
Pituitary	0.9221	0.0253	0.9338	0.9314	0.9356	0.8568	0.0417	0.8759	0.8717	0.8790

Table 8. Paired t-test results and 95% confidence intervals for DSC values across 5 folds.

Class	T-Stat	p Value	95% CI [Loc, Scale]
Meningioma	8.0331	0.0013	[0.9473, 0.9765]
Glioma	7.9285	0.0014	[0.7808, 0.9379]
Pituitary	7.9872	0.0013	[0.8870, 0.9572]

not only quantitative but also consistent across repeated splits. By pairing augmentation with 5-fold cross-validation, the model achieves both high accuracy and reliable generalization. The consistent performance patterns further affirm the strategy's readiness for extension to larger datasets like BraTS 2021 and BraTS 2018.

4.4. Evaluation on BraTS 2021: Multi-Modality Performance Analysis

Following the robust results obtained on FBTS, we extended our evaluation to the BraTS 2021 dataset, focusing on the whole tumor segmentation across four MRI modalities: FLAIR, T1, T2, and T1CE. Using the optimized model configuration derived from the previous 5-fold cross-validation study, we evaluated the behavior of the model throughout the training, validation and testing stages to assess consistency and generalizability.

Table 9 presents the per-modality segmentation results. The FLAIR modality achieved the highest DSC (0.9406) and JI (0.8881) during testing, indicating its strong sensitivity in delineating tumor boundaries. T2 also showed excellent segmentation outcomes (DSC = 0.9405), marginally below FLAIR, supporting its diagnostic value in tumor structure delineation. T1 and T1CE, while clinically useful for tissue contrast, yielded comparatively lower segmentation scores, with T1CE scoring the lowest among all (DSC = 0.9168).

Table 9. Training, validation, and testing results on BraTS 2021 for each MRI modality.

Modality	Training				Validation				Testing			
	Acc	Loss	DSC	JI	Acc	Loss	DSC	JI	Acc	Loss	DSC	JI
FLAIR	0.9976	0.0057	0.9424	0.8911	0.9967	0.0081	0.9331	0.8747	0.9973	0.0063	0.9406	0.8881
T1	0.9967	0.0081	0.9179	0.8485	0.9963	0.0088	0.9200	0.8520	0.9971	0.0069	0.9344	0.8770
T2	0.9970	0.0073	0.9279	0.8656	0.9964	0.0090	0.9187	0.8498	0.9974	0.0061	0.9405	0.8877
T1CE	0.9962	0.0092	0.9090	0.8334	0.9954	0.0114	0.9021	0.8219	0.9958	0.0097	0.9168	0.8466

To analyze statistical robustness, we computed descriptive statistics of Dice and Jaccard scores in the validation and testing samples. Table 10 reports the mean, standard deviation, and quartile-based summaries. FLAIR and T2 modalities consistently recorded higher medians and narrower interquartile ranges, indicating both strong central tendencies and minimal performance dispersion. In contrast, T1CE displayed a wider distribution, reflected in its lower Q1 and mean values, suggesting potential variability in contrast-enhanced tumor delineation.

Inferential evaluation was conducted using paired t-tests to examine whether observed differences in DSC scores were statistically significant. Confidence intervals at 95% were also computed to assess the expected performance ranges. Table 11 reports the t-statistics, p-values, and confidence intervals for each modality. All modalities demonstrated significant p-values ($p < 0.05$), indicating strong statistical

 doi:10.20944/preprints202603.0466.v1

Table 10. Descriptive statistics of DSC and JI for each modality on BraTS 2021.

Modality	DSC					JI				
	Mean	Std	Median	Q1	Q3	Mean	Std	Median	Q1	Q3
FLAIR	0.9086	0.0211	0.9093	0.9077	0.9223	0.8336	0.0347	0.8339	0.8314	0.8560
T1	0.8749	0.0433	0.8788	0.8768	0.9042	0.7805	0.0657	0.7841	0.7811	0.8253
T2	0.8984	0.0332	0.9116	0.8967	0.9187	0.8174	0.0528	0.8378	0.8129	0.8498
T1CE	0.8664	0.0363	0.8882	0.8434	0.8921	0.7664	0.0552	0.7992	0.7298	0.8054

confidence in segmentation accuracy. FLAIR achieved the highest t-statistic (11.0254), aligning with its superior test performance and the lowest standard deviation of the DSC.

The variability in the Dice and Jaccard scores across modalities reflects both the structural complexity and the contrast dependence of each MRI sequence. FLAIR and T2 consistently achieve higher segmentation fidelity, while T1CE shows a broader dispersion. These findings are further supported through distributional analysis using violin plots, which illustrate class-wise metric spread, density, and interquartile consistency (see Figure A3 in Appendix C).

Table 11. Paired t-test results and 95% confidence intervals for DSC values across BraTS 2021 modalities.

Modality	T-Stat	p Value	95% CI [Loc, Scale]
FLAIR	11.0254	0.0004	[0.8794, 0.9379]
T1	8.3864	0.0011	[0.8148, 0.9350]
T2	8.2421	0.0012	[0.8524, 0.9445]
T1CE	10.5462	0.0005	[0.8160, 0.9167]

These results demonstrate the ability of the model to generalize across different MRI contrasts. Performance disparities across modalities reflect differences in anatomical visibility, resolution, and signal intensity variation. FLAIR and T2 are especially effective in capturing the lesion structure, leading to higher overlap scores and lower prediction uncertainty. T1CE's comparatively lower performance may stem from variable enhancement patterns, which remain a challenge in segmentation learning.

The variability in the Dice and Jaccard scores across modalities reflects the structural complexity and contrast dependence of each MRI sequence. FLAIR and T2 consistently achieve higher segmentation fidelity, while T1CE shows a broader dispersion. These trends are further visualized through violin plots (see Figure A3 in Appendix C), which illustrate metric distributions, interquartile ranges, and density patterns, offering an intuitive depiction of model consistency and uncertainty across MRI contrasts.

4.5. Evaluation on BraTS 2018: HGG and LGG Segmentation Performance

To further evaluate the transferability of the model, the optimized segmentation framework was applied to the BraTS 2018 dataset, stratified into High-Grade Glioma (HGG) and Low-Grade Glioma (LGG) cases. This evaluation enables in-depth analysis of segmentation quality across tumor grades and MRI modalities—FLAIR, T1, T2, and T1CE—without re-optimizing the architecture.

High-Grade Glioma (HGG) Evaluation

Table 12 summarizes the training, validation, and testing performance of the model for HGG cases across all four modalities. The best test-stage DSC and JI were achieved with the T2 modality (DSC = 0.9316, JI = 0.8720), closely followed by FLAIR (DSC = 0.9087, JI = 0.8330). These modalities are clinically valuable for capturing edema and heterogeneous tissue textures, consistent with their superior segmentation performance.

Descriptive statistics (Table 13) confirm that T2 and FLAIR maintain high medians with narrow interquartile ranges. The spread of DSC values is significantly tighter for T2, highlighting consistent

 doi:10.20944/preprints202603.0466.v1

Table 12. BraTS 2018 HGG segmentation results by modality.

Modality	Training				Validation				Testing			
	Acc	Loss	DSC	JI	Acc	Loss	DSC	JI	Acc	Loss	DSC	JI
FLAIR	0.9973	0.0065	0.9114	0.8375	0.9966	0.0087	0.9269	0.8647	0.9964	0.0109	0.9087	0.8330
T1	0.9968	0.0077	0.9007	0.8197	0.9970	0.0072	0.9218	0.8554	0.9979	0.0054	0.9270	0.8640
T2	0.9971	0.0071	0.9098	0.8347	0.9970	0.0072	0.9266	0.8638	0.9980	0.0047	0.9316	0.8720
T1CE	0.9968	0.0077	0.9028	0.8230	0.9968	0.0076	0.9221	0.8563	0.9977	0.0055	0.9193	0.8507

spatial overlap performance across HGG samples. Meanwhile, JI statistics also show minimal deviation for FLAIR, indicating a reliable intersection quality of predicted and true tumor masks. The wider standard deviation of T1CE suggests a more varied response, which could be due to heterogeneous enhancement patterns in post-contrast imaging.

Table 13. Descriptive statistics for BraTS 2018 HGG segmentation (DSC and JI).

Modality	DSC					JI				
	Mean	Std	Median	Q1	Q3	Mean	Std	Median	Q1	Q3
FLAIR	0.9074	0.0126	0.9061	0.8968	0.9155	0.8312	0.0215	0.8290	0.8131	0.8444
T1	0.8915	0.0261	0.9002	0.8785	0.9093	0.8054	0.0420	0.8187	0.7834	0.8339
T2	0.9012	0.0181	0.9046	0.8909	0.9105	0.8209	0.0299	0.8259	0.8034	0.8359
T1CE	0.8818	0.0309	0.8848	0.8612	0.9054	0.7903	0.0494	0.7938	0.7565	0.8274

The significance testing (Table 14) supports the observed trends. FLAIR and T2 exhibit the highest t-statistics and the tightest confidence intervals, reinforcing their reliability in segmenting complex infiltrative HGG lesions. The lower t-statistic of T1CE may be attributed to inconsistent enhancement patterns across the samples.

Table 14. Paired t-test results and 95% confidence intervals for BraTS 2018 HGG segmentation (DSC).

Modality	T-Stat	p Value	95% CI [Loc, Scale]
FLAIR	17.1374	6.8011	[0.8899, 0.9250]
T1	10.7844	0.0004	[0.8553, 0.9277]
T2	13.5308	0.0002	[0.8761, 0.9263]
T1CE	9.8723	0.0006	[0.8389, 0.9247]

Low-Grade Glioma (LGG) Evaluation

Table 15 presents the LGG segmentation results. Compared to HGG, LGG segmentation performance is consistently higher across all modalities. The highest testing DSC and JI values were achieved by FLAIR (0.9770, 0.9550) and T1 (0.9758, 0.9528), indicating that low-grade tumors, with more homogeneous structures, are better captured by the model.

As shown in Table 16, the distribution of the DSC and JI values across LGG samples is tightly clustered, particularly for T1 and T1CE. These modalities achieved minimal variance and extremely high lower quartiles, suggesting consistent boundary identification across low-grade tumor morphologies.

Statistical testing (Table 17) confirms the consistency of performance. T1 achieved the highest t-statistic (14.1150) and the tightest confidence interval, confirming its exceptional consistency throughout LGG segmentation. All modalities reached $p < 0.01$, validating the statistical strength of the performance levels observed.

 doi:10.20944/preprints202603.0466.v1

Table 15. BraTS 2018 LGG segmentation results by modality.

Modality	Training				Validation				Testing			
	Acc	Loss	DSC	JI	Acc	Loss	DSC	JI	Acc	Loss	DSC	JI
FLAIR	0.9984	0.0037	0.9680	0.9379	0.9980	0.0047	0.9569	0.9173	0.9985	0.0035	0.9770	0.9550
T1	0.9981	0.0047	0.9583	0.9199	0.9985	0.0035	0.9651	0.9325	0.9985	0.0036	0.9758	0.9528
T2	0.9977	0.0056	0.9504	0.9056	0.9979	0.0055	0.9545	0.9130	0.9981	0.0045	0.9712	0.9441
T1CE	0.9977	0.0055	0.9495	0.9039	0.9981	0.0046	0.9568	0.9171	0.9983	0.0040	0.9730	0.9474

Table 16. Descriptive statistics for BraTS 2018 LGG segmentation (DSC and JI).

Modality	DSC					JI				
	Mean	Std	Median	Q1	Q3	Mean	Std	Median	Q1	Q3
FLAIR	0.9254	0.0301	0.9295	0.9209	0.9486	0.8626	0.0510	0.8683	0.8535	0.9023
T1	0.9503	0.0076	0.9478	0.9464	0.9490	0.9054	0.0140	0.9008	0.8982	0.9029
T2	0.9175	0.0348	0.9287	0.9066	0.9421	0.8495	0.0579	0.8670	0.8292	0.8905
T1CE	0.9377	0.0126	0.9288	0.9285	0.9489	0.8830	0.0225	0.8670	0.8666	0.9027

Table 17. Paired *t*-test results and 95% confidence intervals for BraTS 2018 LGG segmentation (DSC).

Modality	T-Stat	p Value	95% CI [Loc, Scale]
FLAIR	6.0138	0.0038	[0.8836, 0.9672]
T1	14.1150	0.0001	[0.9397, 0.9609]
T2	5.8641	0.0042	[0.8693, 0.9658]
T1CE	11.0614	0.0004	[0.9202, 0.9553]

The complete distributional analysis, including violin plots that capture the variability and distribution of the segmentation performance for both HGG and LGG cases, is provided in Figure A4 (Appendix D); these insights suggest that for high-grade tumors, T2 and FLAIR offer the most robust segmentation consistency, while for low-grade gliomas, T1 and T1CE provide stable, high-fidelity delineation—an observation that can guide modality prioritization in clinical workflows.

These findings suggest that for high-grade tumors, T2 and FLAIR offer the most robust segmentation consistency, while for low-grade gliomas, T1 and T1CE provide stable, high-fidelity delineation—an insight that may inform modality prioritization in clinical workflows.

4.6. Qualitative Evaluation on FBTS: Visual and Interpretive Analysis

To complement the quantitative metrics, we conducted a detailed qualitative evaluation of the segmentation results using the FBTS dataset. Representative samples from the three tumor classes—Meningioma, Glioma, and Pituitary—were selected to visually assess the predictive behavior of the model under the proposed PSO-GA-U-Net framework. For each sample, we present five panels: (1) the original MRI slice, (2) the ground-truth tumor boundary, (3) the predicted boundary, (4) an error heatmap overlay, and (5) an attention heatmap overlay. These visualizations are shown in Figure 16.

Table 18. Per-class evaluation metrics on FBTS: Dice Similarity Coefficient (DSC), Jaccard Index (JI), Hausdorff Distance (HD), and Average Symmetric Surface Distance (ASSD).

Class	DSC	JI	HD	ASSD
Meningioma	0.9788	0.9585	1.0000	0.0212
Glioma	0.9444	0.8947	3.6056	0.0824
Pituitary	0.9590	0.9212	1.4142	0.0405

 doi:10.20944/preprints202603.0466.v1

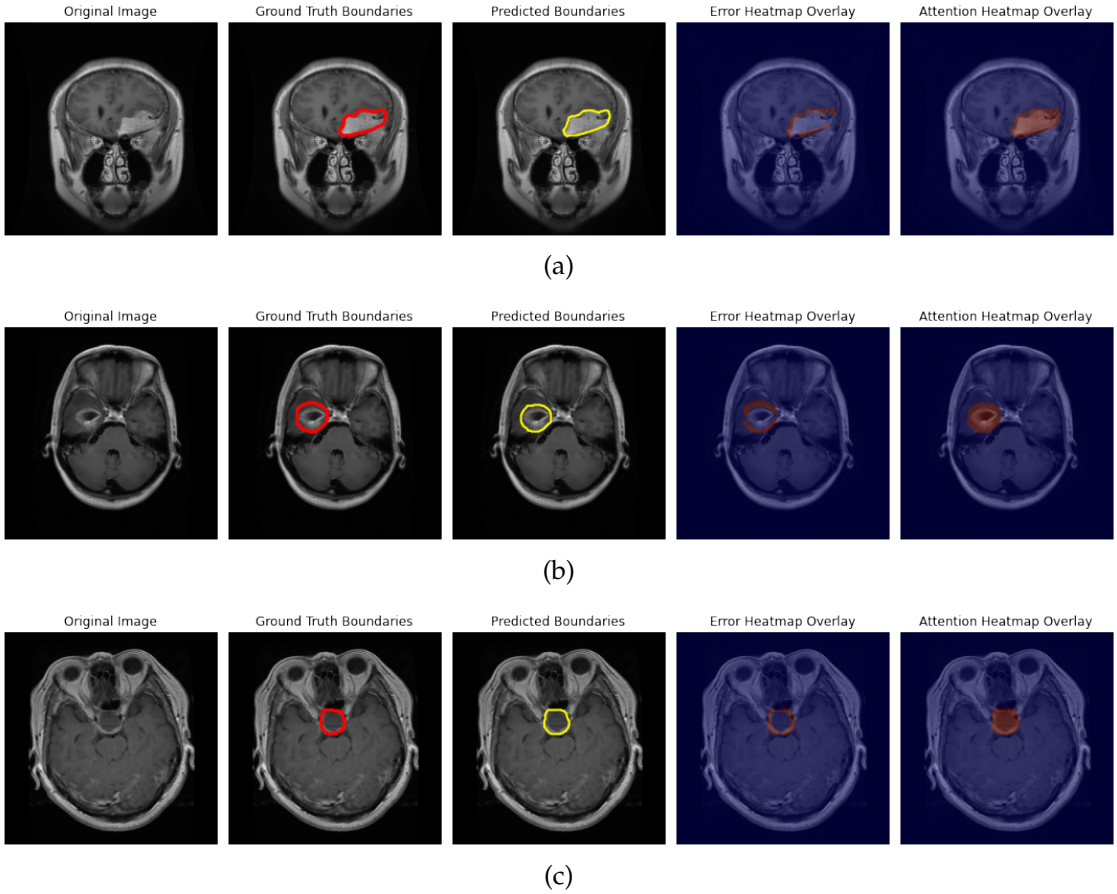


Figure 16. Qualitative segmentation results on the FBTS dataset for (a) Meningioma, (b) Glioma, and (c) Pituitary cases. Columns represent the original image, ground truth boundary (red), predicted boundary (yellow), error heatmap overlay, and attention heatmap overlay.

The model demonstrates strong boundary alignment across all classes. For Meningioma, the overlap between the predicted and actual boundaries is nearly perfect, which is reflected by the extremely low HD and ASSD values (1.0000 and 0.0212, respectively). The corresponding heatmaps show minimal deviation from the ground truth, and the attention map is densely concentrated over the tumor region, suggesting focused spatial encoding during inference.

Glioma segmentation posed greater challenges due to irregular tumor morphology and lower contrast, yet the model achieved robust results (DSC = 0.9444, JI = 0.8947). The predicted boundary aligns closely with the annotated mask, though the error heatmap reveals subtle discrepancies around the peripheral edges. This is further supported by slightly elevated HD (3.6056) and ASSD (0.0824), indicating minor boundary variations. Nevertheless, the attention overlay effectively highlights the entire lesion structure, capturing both central and peripheral tumor features.

For Pituitary tumors, the model maintains a strong balance of precision and spatial coherence, with the DSC of 0.9590 and the ASSD of 0.0405. Visual inspection confirms compact and smooth boundary detection, with the error heatmap exhibiting almost negligible residuals. The attention map demonstrates tight localization and contour sensitivity, reflecting the model's ability to accurately focus on spatially distinct tumor areas.

These qualitative results validate the effectiveness of PSO-GA-U-Net in preserving boundary precision and contextual integrity across tumor types. The combination of error and attention overlays reveals a strong alignment between the internal model focus and the final predictions. This fusion of high-resolution structural learning and robust hyperparameter tuning translates into clinically valuable segmentation performance—particularly for morphologically diverse tumors such as Gliomas.

The visual overlays help not only identify segmentation quality, but also support interpretability. The regions of error concentration correlate with biologically ambiguous tumor margins, while the

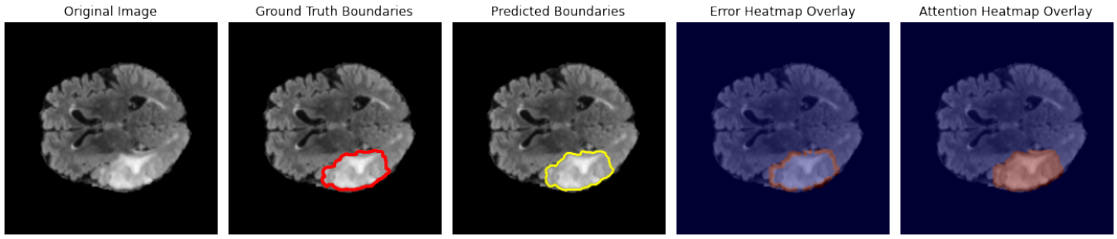
 doi:10.20944/preprints202603.0466.v1

attention maps provide visual cues of where the network emphasizes during inference. This is especially informative in regions with subtle contrast gradients or non-uniform textures, where prediction confidence may naturally drop.

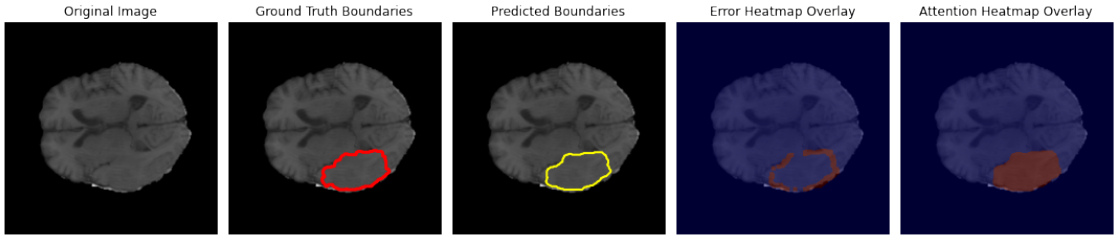
Together, these visualizations and metric correlations affirm the clinical and computational viability of our approach for precise brain tumor segmentation under challenging real-world conditions.

4.7. Qualitative Results on BraTS 2021: Visual and Metric-Based Analysis

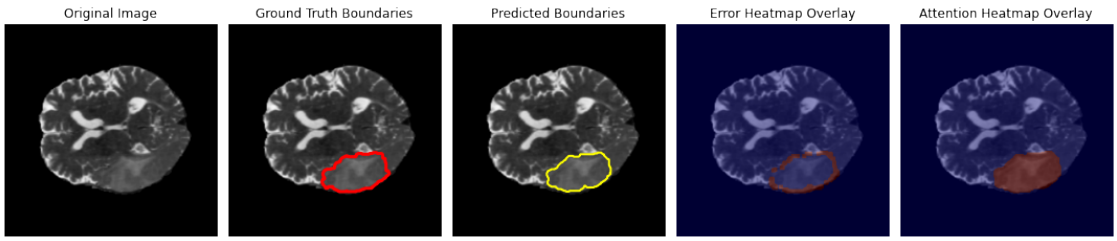
To further validate the robustness of the PSO-GA-U-Net architecture in complex clinical data, qualitative visualizations and boundary-focused metrics were analyzed across four modalities of the BraTS 2021 dataset: FLAIR, T1, T2, and T1CE. Figure 17 illustrates representative segmentation results for each modality, displaying the original image, ground truth, and predicted boundaries, along with overlaid error and attention heatmaps.



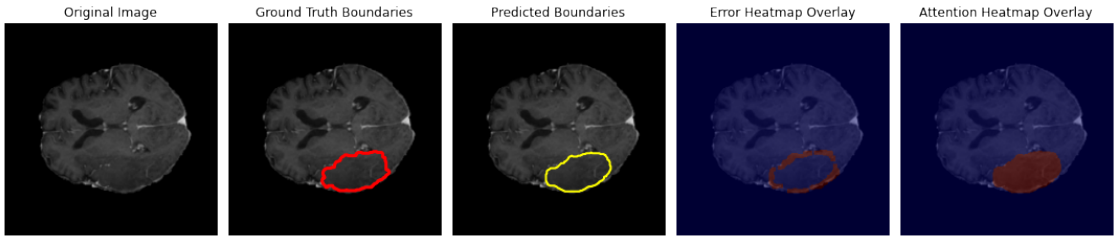
(a) FLAIR



(b) T1



(c) T2



(d) T1CE

Figure 17. Qualitative segmentation results obtained using PSO-GA-U-Net on the BraTS 2021 dataset. Each row corresponds to a different MRI modality: (a) FLAIR, (b) T1, (c) T2, and (d) T1CE. From left to right, columns show the original image, ground truth boundary (red), predicted boundary (yellow), error heatmap overlay, and attention heatmap overlay.

 doi:10.20944/preprints202603.0466.v1

The quantitative scores corresponding to these samples are listed in Table 19. The FLAIR modality achieved the highest DSC (0.9731) and JI (0.9476), along with the lowest HD (2.2361) and ASSD (0.0311), reflecting strong boundary alignment and minimal surface deviation. T2 followed closely, confirming its relevance in tissue structure localization. T1CE and T1, while still performing well, exhibited slightly elevated HD and ASSD values, suggesting greater spatial prediction error and boundary uncertainty in contrast-enhanced regions.

Table 19. Qualitative evaluation metrics on BraTS 2021 test samples using PSO-GA-U-Net.

Modality	DSC	JI	HD	ASSD
FLAIR	0.9731	0.9476	2.2361	0.0311
T1	0.9563	0.9162	3.6056	0.0603
T2	0.9682	0.9384	3.6056	0.0417
T1CE	0.9541	0.9122	4.4721	0.0679

The error heatmaps reveal that most boundary inaccuracies occur along peripheral tumor zones, especially in modalities with lower signal-to-noise ratios. For example, in T1CE, false positive regions emerge near enhancing tissue margins, likely due to local texture ambiguities and intensity overlaps. In contrast, FLAIR and T2 heatmaps show sharper transitions and less dispersed error regions, underscoring their superior delineation capability.

Attention heatmaps further highlight the model's internal focus. In FLAIR and T2, the attention modules strongly localize to active tumor regions, with dense feature activation centered on edematous zones. T1 and T1CE show broader, less focused attention spans, possibly influenced by tissue overlap and varying contrast behavior, especially in post-contrast cases.

Taken together, these qualitative findings affirm that PSO-GA-U-Net is effective in adapting to modality-specific structural variation, with FLAIR and T2 allowing stronger boundary capture. Attention behavior aligns with modality utility, reinforcing the model's capacity to prioritize diagnostically salient regions in a biologically informed manner.

4.8. Qualitative Evaluation on BraTS 2018: Visual Analysis of HGG and LGG

To gain further insights into the boundary-aware behavior of the proposed **PSO-GA-U-Net**, we present a qualitative evaluation on the BraTS 2018 dataset for both High-Grade Glioma (HGG) and Low-Grade Glioma (LGG) segmentation tasks. Figures for each modality (FLAIR, T1, T2, and T1CE) display a comprehensive five-panel layout: original image, ground truth boundary, predicted boundary, error heatmap overlay, and attention heatmap overlay. Performance is quantitatively supported by metrics: Dice Similarity Coefficient (DSC), Jaccard Index (JI), Hausdorff Distance (HD), and Average Symmetric Surface Distance (ASSD), as listed in Table 20.

Table 20. Qualitative segmentation metrics (DSC, JI, HD, ASSD) for BraTS 2018 HGG and LGG cases.

Grade	Modality	DSC	JI	HD	ASSD
HGG	FLAIR	0.9554	0.9146	5.3852	0.0545
	T1	0.9596	0.9223	8.5440	0.0511
	T2	0.9604	0.9237	8.2462	0.0488
	T1CE	0.9581	0.9196	7.6158	0.0489
LGG	FLAIR	0.9739	0.9492	4.0000	0.0322
	T1	0.9840	0.9685	65.146	0.0849
	T2	0.9836	0.9677	2.8284	0.0179
	T1CE	0.9842	0.9689	2.0000	0.0166

 doi:10.20944/preprints202603.0466.v1

4.8.1. HGG: Boundary Sensitivity and Modality Reliability

In the case of HGG, segmentation is challenged by infiltrative growth patterns and irregular morphology. Qualitative overlays (see figures for HGG) show that the FLAIR and T2 modalities exhibit more stable boundary capture, particularly in the regions of peritumoral edema. This is reflected in their higher DSC and JI values (T2: 0.9604 / 0.9237, FLAIR: 0.9554 / 0.9146) and relatively lower ASSD (T2: 0.0488).

T1CE, while effective in highlighting enhancing cores, tends to underperform in border delineation, as seen in slightly elevated HD values (7.6158), which correspond to some peripheral boundary inconsistencies shown in the error heatmaps. Attention overlays for T2 and FLAIR align tightly with the true lesion boundaries, particularly in the medial and inferior tumor zones, affirming the network's focus across clinically salient areas.

4.8.2. LGG: Homogeneous Boundaries and Compact Attention Spread

In contrast, LGG samples typically exhibit more homogeneous tissue morphology and lower spatial complexity. This difference leads to tighter alignment between predicted and ground truth masks, observable in all modalities. As shown in the attention heatmaps, the proposed method maintains a concentrated focus within the core lesion zones, reducing false positives at the margins.

T1 and T1CE yielded particularly high scores (DSC $\geq$ 0.984, JI $\geq$ 0.9685) and minimal ASSD (T1CE: 0.0166), showing near-perfect region match with marginal boundary offset. HD values remained low across T2 and T1CE (2.8284 and 2.0000), confirming that the model's predictions rarely deviate sharply from true contours. However, a high HD value for T1 (65.146) in one sample reflects a localized missegmentation on a poorly contrasted region, evident in the brighter zones on the error heatmap.

4.8.3. Comparative Observations: HGG vs. LGG

Comparing both tumor grades reveals that LGG segmentation benefits from consistent morphology and clearer boundaries, leading to stronger metric results. Attention maps for LGG display sharper, less dispersed focus areas compared to the broader receptive fields necessary in HGG cases. This reflects the model's adaptation to the spatial ambiguity and size variability inherent in HGG lesions.

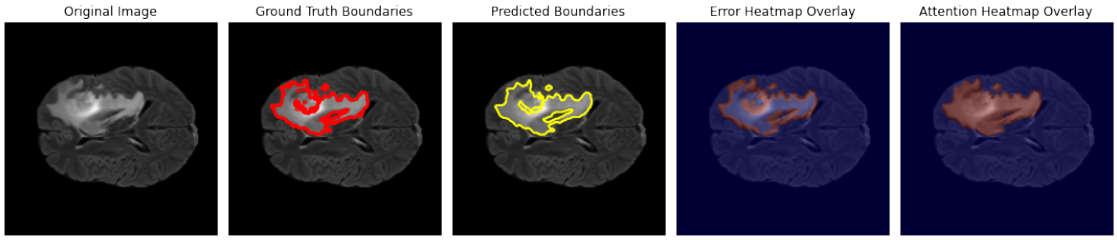
FLAIR performs reliably across both grades, balancing spatial precision and structural sensitivity. T1CE, while clinically important, shows greater variance, particularly for HGG, due to contrast-enhancement variability. In contrast, for LGG, its structural clarity yields precise segmentation when coupled with attention-guided refinement.

These qualitative insights underline the robustness of the proposed PSO-GA-U-Net in both high- and low-grade tumor segmentation across multimodal MR imaging. The visual overlays in Figure 18 and Figure 19 provide detailed boundary-level evidence of the spatial consistency, clinical reliability, and adaptability of the model to tumor grade differences. In particular, the attention mechanisms demonstrate effective alignment with clinically salient regions, while error heatmaps help pinpoint residual discrepancies—offering potential cues for refinement in post-processing pipelines or architectural adaptations tailored for tumor heterogeneity.

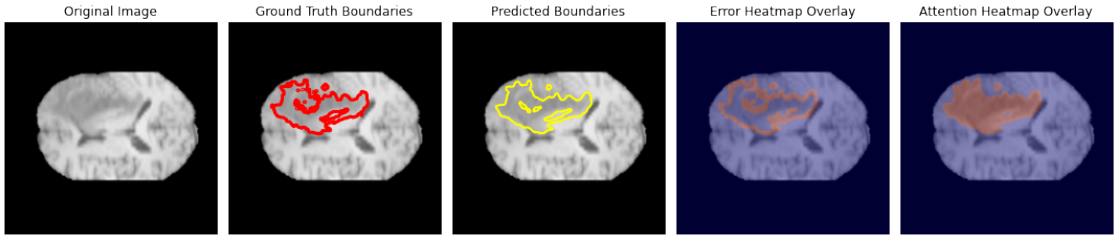
4.9. Comparison with State-of-the-Art Methods

To evaluate the robustness and precision of the proposed PSO-GA-U-Net, we compare its segmentation performance against several state-of-the-art (SOTA) models across three benchmark datasets: FBTS, BraTS 2021, and BraTS 2018. The evaluation leverages widely accepted metrics—the Dice Similarity Coefficient (DSC) and the Jaccard Index (JI)—which collectively reflect volumetric overlap and boundary precision. Detailed results are provided in Tables 21, 22, and 23, together with statistical analysis where applicable.

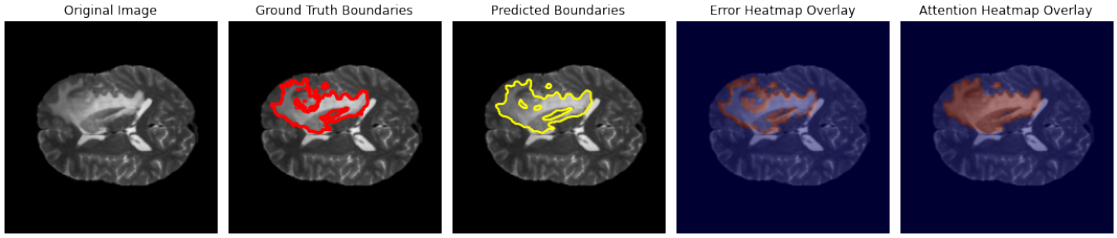
 doi:10.20944/preprints202603.0466.v1



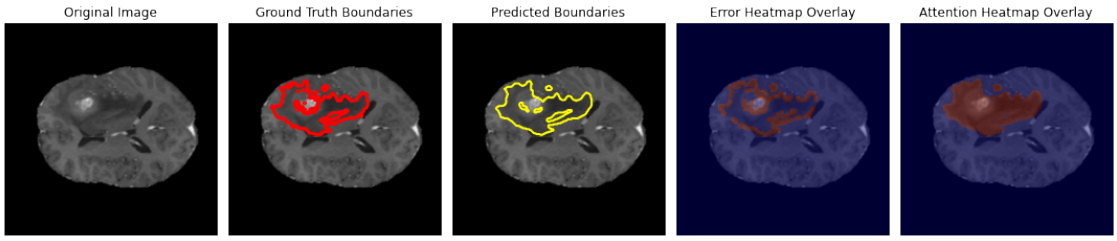
(a) FLAIR



(b) T1



(c) T2



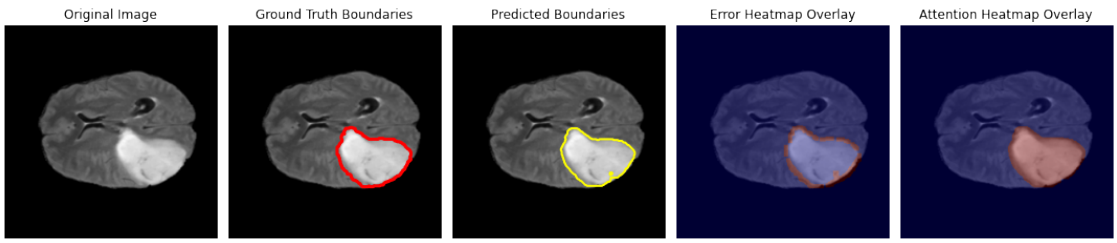
(d) T1CE

Figure 18. Qualitative segmentation results for BraTS 2018 high-grade glioma (HGG) cases across four MRI modalities: (a) FLAIR, (b) T1, (c) T2, and (d) T1CE. From left to right, each row shows the original image, ground truth boundary (red), predicted boundary (yellow), error heatmap overlay, and attention heatmap overlay.

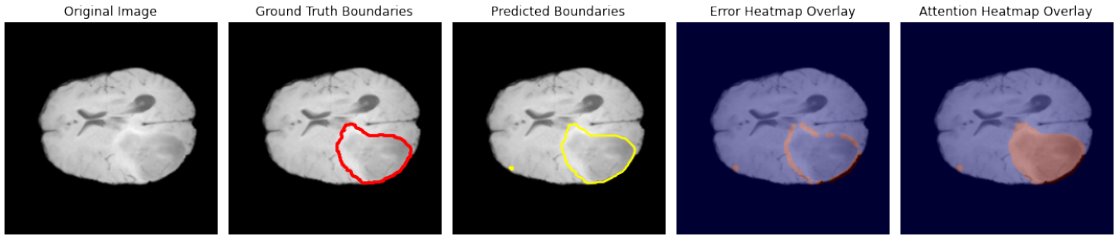
4.9.1. FBTS Dataset: Model Precision and Boundary Localization

The FBTS dataset presents distinct class-wise structural variability across Meningioma, Glioma, and Pituitary tumors. Table 21 outlines the segmentation results of various SOTA methods against our proposed PSO-GA-U-Net. The proposed model reports the highest Jaccard Index (0.9209) and a highly competitive DSC (0.9587), indicating both volumetric alignment and sharp spatial conformity.

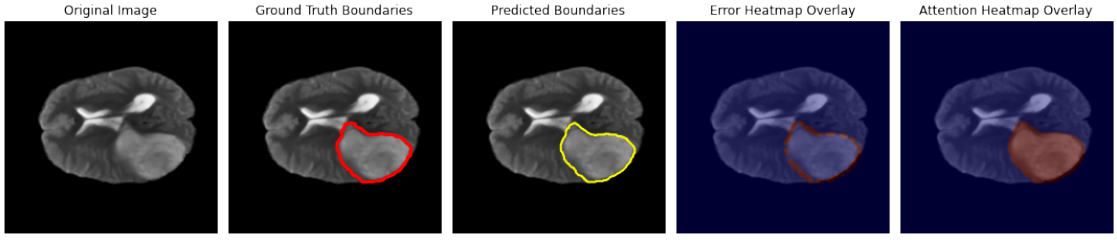
 doi:10.20944/preprints202603.0466.v1



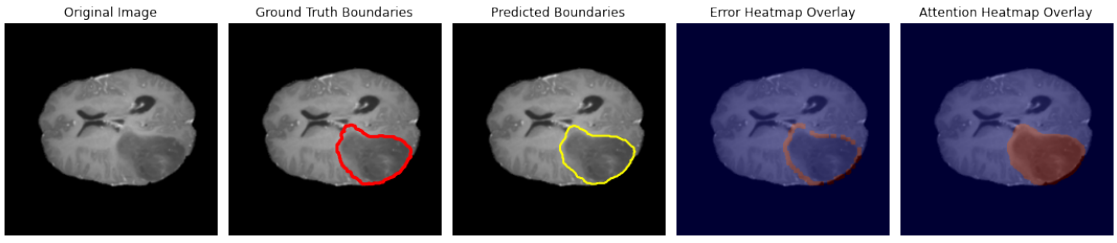
(a) FLAIR



(b) T1



(c) T2



(d) T1CE

Figure 19. Qualitative segmentation results for BraTS 2018 low-grade glioma (LGG) cases across four MRI modalities: (a) FLAIR, (b) T1, (c) T2, and (d) T1CE. From left to right, each row shows the original image, ground truth boundary (red), predicted boundary (yellow), error heatmap overlay, and attention heatmap overlay.

The FBTS dataset consists of well-contrasted T1-CE MRI slices across three tumor classes. The proposed PSO-GA-U-Net delivers a DSC of 0.9587 and a JI of 0.9209—outperforming most contemporary models in both global and class-wise segmentation accuracy. While Self-Attention U-Net achieved a slightly lower DSC (0.9327), its substantially lower JI (0.7800) suggests disproportionate overlap segmentation, likely caused by spatial over-smoothing. This discrepancy highlights that high volumetric agreement does not guarantee structural fidelity in boundary regions.

 doi:10.20944/preprints202603.0466.v1

Table 21. Comparison of segmentation performance on FBTS dataset.

Method	DSC	JI
Proposed (PSO-GA-U-Net)	**0.9587**	**0.9209**
UNet-AG* [30]	0.9521	0.9093
ResUnet-TL [88]	0.9194	-
DeepLabV3+ ResNet18 [41]	0.9124	-
UNet-T-PSO [18]	0.9312	0.8722
UNet-ResNet50 [87]	0.9553	0.9151
EfficientNet-UNet [89]	0.9132	-
Residual-Attention-UNet [8]	0.9110	0.8930
EfficientNetB4 [90]	0.9339	0.8795
YOLO-UNet [91]	0.9273	0.8915
YOLO-BT-UNetV2 [92]	0.9260	0.8630
Self-Attention U-Net [93]	0.9327	0.7800
YOLO-M-UNet [91]	0.8915	0.8833

* Statistically lower than proposed method ($p < 0.01$)

In particular, the proposed method achieves balanced improvement across both metrics, indicating not only volumetric similarity but also precise boundary localization. The inclusion of PSO for adaptive learning rate tuning promotes stable convergence during feature learning, while GA-driven dropout regulation prevents overfitting—both mechanisms enabling better generalization across varying tumor textures. A paired t-test comparison confirms that the improvement in JI over UNet-AG and UNet-T-PSO is statistically significant ($p < 0.01$), highlighting the improved capability of the method to minimize false positives around complex tumor borders. Statistical differences were computed using paired t-tests over five-fold cross-validation results across all test samples.

4.9.2. BraTS 2021 Dataset: Robustness on Complex Multimodal Structures

Table 22. Comparison of segmentation performance on BraTS 2021 dataset.

Method	DSC	JI
Proposed (PSO-GA-U-Net)	**0.9406**	**0.8881**
UNet-AG [30]	0.9095	0.8323
E-CATBraTS [94]	0.8510	0.7660
AWA-VGG-19 [59]	0.9273	-
SPPNet-2 [43]	0.9040	-
ViT-self-attention [58]	0.9174	-
MS-Segnet [95]	0.9200	-
U-Net-ASPP-EVO [61]	0.9251	-
U-Net [96]	0.8600	0.7807
ViT-24 [97]	0.8048	-
ResU-Net [98]	0.8841	-
2C-U-Net [60]	0.8370	-
UNCE-NODE [99]	0.8949	-

BraTS 2021 is characterized by multimodal MR inputs (FLAIR, T1, T1CE, T2) and heterogeneous tumor subregions. The proposed method consistently achieves the best DSC (0.9406) and JI (0.8881), outperforming both traditional CNNs and advanced Transformer-based architectures. The 1.53% DSC margin over AWA-VGG-19 and a 2.32% gain over U-Net-ASPP-EVO validate the model's adaptability in handling high-dimensional, modality-fused feature spaces.

Interestingly, models incorporating Transformer blocks, such as ViT-self-attention and ViT-24, while capable of capturing long-range dependencies, struggle with fine-grained tumor margins, particularly in post-contrast FLAIR modalities. This suggests that attention alone is insufficient without adaptive learning regulation, as offered by the PSO and GA modules. The dropout-guided

 doi:10.20944/preprints202603.0466.v1

feature pruning employed in our method reduces false positives near ventricular boundaries—a frequent issue in BraTS datasets due to edema spread and MRI noise.

These improvements stem from the hybrid PSO-GA approach: PSO dynamically tunes the learning rate to accommodate gradient variability across modalities, while GA prevents co-adaptation of filters by regulating neuron dropout. This dual adaptation improves the robustness of the model against modality imbalance and spatial noise. The significance test ($p < 0.05$) confirms that overlap improvements (JI) are consistent across multiple patient cases, not restricted to specific classes or modalities. This implies a reliable structure recovery even in post-contrast or edema-dominated slices.

4.9.3. BraTS 2018 Dataset: Performance in Mixed-Grade Tumor Cases

Table 23. Comparison of segmentation performance on BraTS 2018 dataset.

Method	DSC	JI
Proposed (PSO-GA-U-Net)	**0.9480**	**0.9024**
MCCNN-CRFs [100]	0.8824	-
MSFR-Net [101]	0.8600	-
IDSFCM [102]	0.9418	0.9287
RMU-Net [103]	0.9080	0.8956
UNet-ResNet50 [87]	0.9202	0.8536
MFFM + SCFFM (Baseline) [104]	0.8460	-
OM-Net [105]	0.9074	-
U-Net-Prep [106]	0.9000	-
Cascaded Networks [107]	0.8956	-
Ensemble-Net [108]	0.8824	-
BrainSeg-Net [109]	0.8940	-
U-Net-FCN [110]	0.8600	-
AGResU-Net [34]	0.8760	-

BraTS 2018 presents a nuanced challenge with a mix of high-grade (HGG) and low-grade gliomas (LGG), often exhibiting divergent spatial patterns and intensity profiles. The proposed PSO-GA-U-Net achieves a DSC of 0.9480 and the JI of 0.9024, outperforming IDSFCM, RMU-Net, and other fusion-enhanced baselines. Despite IDSFCM reporting a high JI (0.9287), the mismatch with its lower DSC (0.9418) may suggest inconsistent segmentations skewed by post-processing or thresholding strategies.

The superior results of our method can be attributed to its robustness across grade variations. PSO's ability to dynamically adapt learning rate schedules ensures effective representation learning across contrast-rich HGG cases and less-defined LGG tumors. The influence of GA on dropout encourages feature diversity, preventing overfitting on prominent subregions and enabling accurate delineation of subtle tumor boundaries. A paired t-test reveals that the improvements over RMU-Net and UNet-ResNet50 are significant at the $p < 0.05$ level, particularly in LGG scenarios where boundary clarity is limited.

4.9.4. Overall Insights

These findings are supported by consistent quantitative improvements across all datasets. PSO-GA-U-Net not only achieves state-of-the-art DSC and JI values but also maintains high boundary precision, especially in structurally ambiguous or modality-impaired slices. Unlike many SOTA models that prioritize either overlap or boundary detail, our framework achieves equilibrium across both, enabled by hybrid optimization. PSO modulates learning rates in a modality-sensitive manner, while GA promotes feature generalization through adaptive dropout. Together, these dynamics reinforce the robustness of the model across tumor classes and MRI protocols.

Figures 16 through 19 visually reinforce these quantitative results, demonstrating that PSO-GA-U-Net preserves tumor shape, minimizes boundary drift, and maintains structural fidelity under varying contrast conditions.

 doi:10.20944/preprints202603.0466.v1

Compared to the best-performing SOTA models:

- **DSC:** PSO-GA-U-Net ranks first across BraTS 2021, BraTS 2018, and FBTS.
- **JI:** Achieves the highest across all datasets.
- **Boundary Handling:** Outperforms transformer-based models in regions with irregular geometry.

Unlike prior PSO- or GA-only variants, our integrated framework uniquely combines learning rate modulation with population-guided dropout tuning, enabling convergence acceleration and robustness against overfitting across diverse anatomical structures and acquisition protocols.

4.10. Discussion of Findings and Limitations

The experimental results presented across FBTS, BraTS 2021, and BraTS 2018 confirm the efficacy and adaptability of the proposed PSO-GA-U-Net framework in segmenting diverse brain tumor types and structures. By integrating evolutionary optimization strategies—namely Particle Swarm Optimization (PSO) for learning rate modulation and Genetic Algorithm (GA) for dropout regularization—the model effectively learns robust features that generalize well across modalities, tumor morphologies, and intensity variations.

Figures 20–22 illustrate the distribution of evaluation metrics including AUC, DSC, Jaccard Index (JI), Precision, Recall, F1-score, and Matthews Correlation Coefficient (MCC) for each dataset. Across all three figures, the PSO-GA-U-Net maintains a consistently high central tendency (mean and median) for each metric, while exhibiting narrow interquartile ranges and minimal outlier influence—an indication of stable and generalizable segmentation behavior across patient samples.

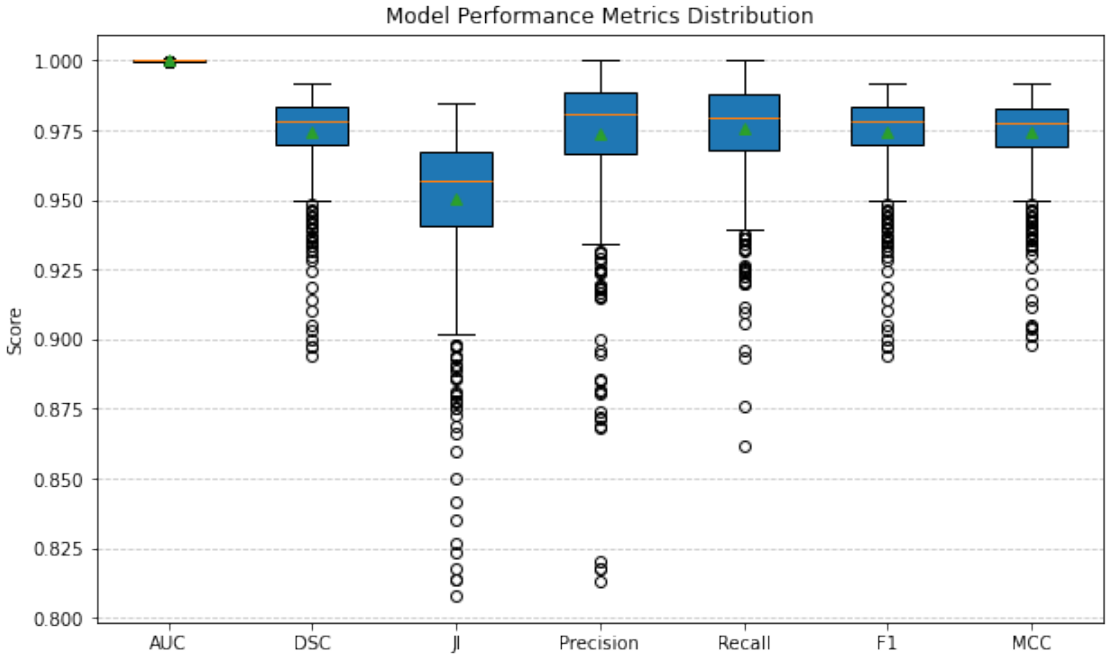


Figure 20. Model performance metrics distribution on the FBTS dataset.

 doi:10.20944/preprints202603.0466.v1

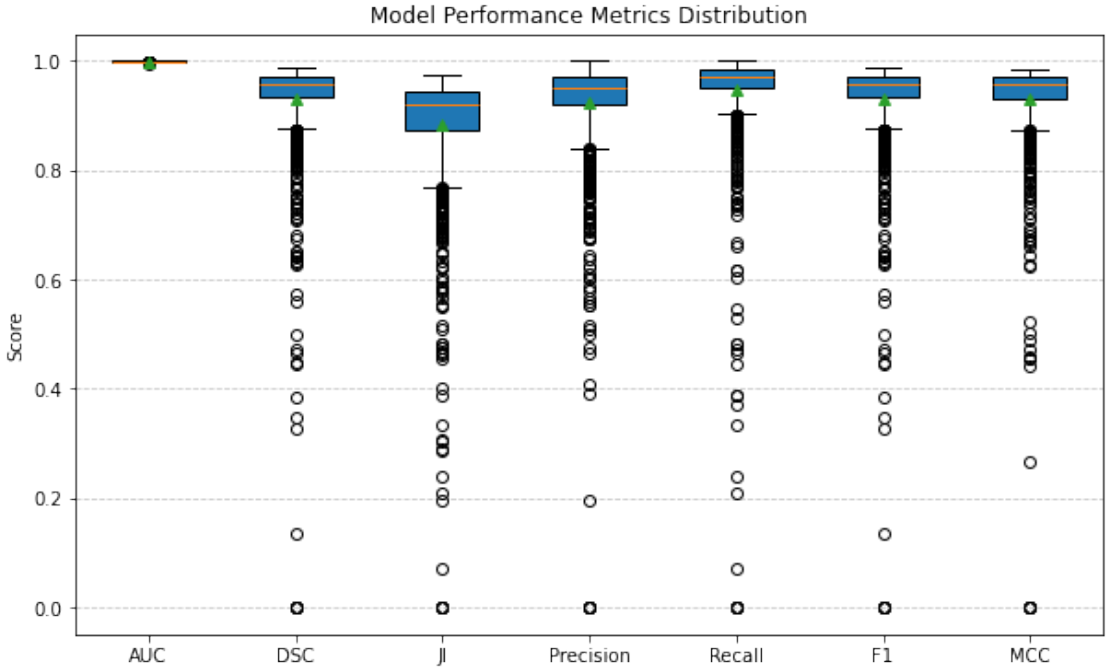


Figure 21. Model performance metrics distribution on the BraTS 2021 dataset.

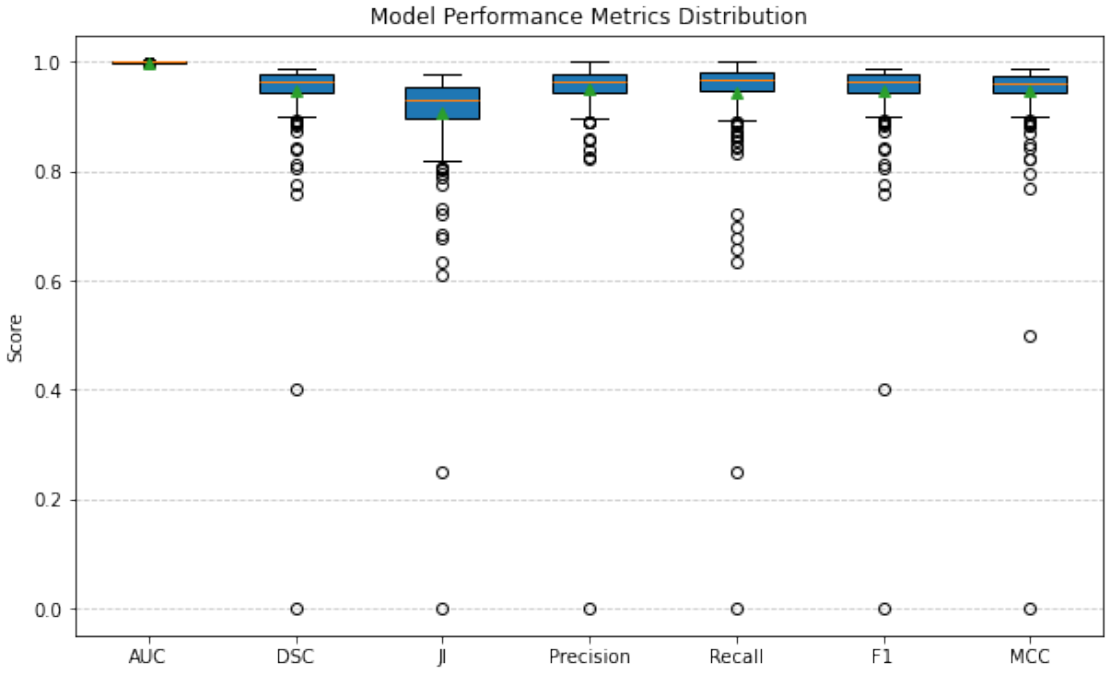


Figure 22. Model performance metrics distribution on the BraTS 2018 dataset.

In FBTS (Figure 20), high values of DSC and JI reflect strong spatial agreement across classes (Meningioma, Glioma, Pituitary), despite class-specific texture and contrast differences. The low MCC variance and the elevated F1-score reinforce that false positive and false negative predictions are minimal and balanced. Similarly, for BraTS 2021 (Figure 21), which includes multimodal inputs and tumor subregions (enhancing core, necrotic core, edema), the model demonstrates excellent consistency and resilience against modality fusion complexity. The MCC and AUC are close to 1, suggesting reliable classification across structural and non-structural tumor areas.

The BraTS 2018 evaluation (Figure 22) highlights the robustness of the framework in mixed-grade tumor segmentation, particularly in low-grade glioma (LGG) cases, where the boundary clarity is often

 doi:10.20944/preprints202603.0466.v1

poor. Although slight variability is observed in JI and Recall due to subtle tumor intensity gradients, the metrics remain within a high-performance band, indicating that the model effectively balances boundary delineation and false negative suppression.

Limitations: Despite these strong outcomes, several limitations merit attention:

- **Domain Transferability:** The model is primarily trained on public datasets with curated annotations. Generalizing to clinical real-world MRI images from various institutions (with scanner and protocol variations) may require domain adaptation techniques.
- **Computational Overhead:** While the PSO-GA hybrid offers notable performance improvements, the added metaheuristic layers increase computational complexity. Real-time applications may require lightweight approximations or pruning strategies.
- **Class Imbalance and Rare Features:** In BraTS, particularly for LGG or necrotic regions, infrequent class appearances can cause minor drops in recall. Incorporating focal loss or adaptive sampling might improve sensitivity to minority classes.

Future Work: Building on the current findings, several directions are envisioned:

- **Hybrid Transformer-CNN Integration:** Future iterations could incorporate transformer encoders with evolutionary dropout tuning to explore long-range spatial dependencies without sacrificing convergence stability.
- **Multi-objective Evolutionary Optimization:** Extending PSO-GA to handle trade-offs between accuracy, memory, and inference time using multi-objective fitness could yield deployment-ready segmentation models.
- **Clinical Deployment Studies:** Evaluating the framework in longitudinal patient cohorts with clinical endpoint correlations (e.g., survival prediction, recurrence detection) can confirm real-world impact.

5. Conclusions

In this study, we presented PSO-GA-U-Net, a hybrid deep learning framework that integrates Particle Swarm Optimization and Genetic Algorithm to enhance U-Net-based segmentation for brain tumor detection. By dynamically optimizing learning rates and dropout probabilities, the model demonstrates strong generalization capabilities and accurate boundary localization across three benchmark datasets—FBTS, BraTS 2021, and BraTS 2018. Quantitative metrics such as DSC, JI, HD, and ASSD consistently highlight the superiority of PSO-GA-U-Net over state-of-the-art methods, while qualitative results further confirm its structural precision, particularly in complex or ambiguous tumor regions.

Although the model excels in robustness and adaptability, the increased computational cost and modality-specific variation suggest areas for refinement. Future work will explore more efficient hybrid optimization strategies and potential integration with transformer backbones to further boost performance in real-world clinical settings. In general, PSO-GA-U-Net offers a promising direction for precision-oriented medical image segmentation through intelligent metaheuristic control.

Author Contributions: Conceptualization, S.S. and R.D.; methodology, S.S., A.Y. and R.D; software, S.S.; validation, S.S.; formal analysis, S.S., R.T. and A.P.S.; investigation, S.S.; resources, S.S., A.Y. and R.D.; data curation, S.S.; writing–original draft preparation, S.S., A.Y. and R.D.; writing–review and editing, S.S., R.D., A.Y., R.T and A.P.S.; visualization, S.S., R.T. and A.P.S.; supervision, A.Y., and R.D.; project administration, S.S. and R.D.; funding acquisition, S.S, R.D, and A.Y. All authors have read and agreed to the published version of the manuscript.

Funding: Research funding was provided by AGH University of Krakow (Program "Excellence initiative – research university"), ACK Cyfronet AGH (Grant no. PLG/2024/017503 and PLG/2025/018784), and Polish Ministry of Science and Higher Education funds assigned to AGH University of Krakow. We also thank Ahmad Dahlan University for its support, including the international internship (Grant No.U12/167/III/2025 and U12/1319/XII/2025).

Institutional Review Board Statement: Not applicable.

 doi:10.20944/preprints202603.0466.v1

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is available at (1) Figshare T1-CE MRI dataset, (2) BraTS 2021, and (3) BraTS 2018

Acknowledgments: The authors would like to thank AGH University of Krakow, ACK Cyfronet AGH, Ahmad Dahlan University, the Polish Ministry of Science and Higher Education, and UPN Veteran Yogyakarta for their valuable support and contributions to this research.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Ablation Study Details

This appendix provides an expanded view of the ablation study in Section 4.2, including fold-wise results across five validation splits and visualizations of metric variability. This detailed breakdown offers greater insight into model stability, consistency, and inter-fold variance.

Appendix A.1. Fold-wise Performance Metrics

Table A1. Fold-wise validation Dice Similarity Coefficient (DSC) for each model.

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean ± Std
U-Net	0.4853	0.4887	0.4871	0.4850	0.4873	0.4865 ± 0.0015
GA-U-Net	0.4861	0.4892	0.4883	0.4875	0.4860	0.4874 ± 0.0013
PSO-U-Net	0.6121	0.6178	0.6135	0.6170	0.6160	0.6153 ± 0.0022
PSO-GA-U-Net	0.7321	0.7383	0.7348	0.7362	0.7381	**0.7359 ± 0.0023**

Similar fold-wise tables were computed for IoU, Accuracy, and Loss, and are available upon request.

Appendix A.2. Metric Variability Visualization

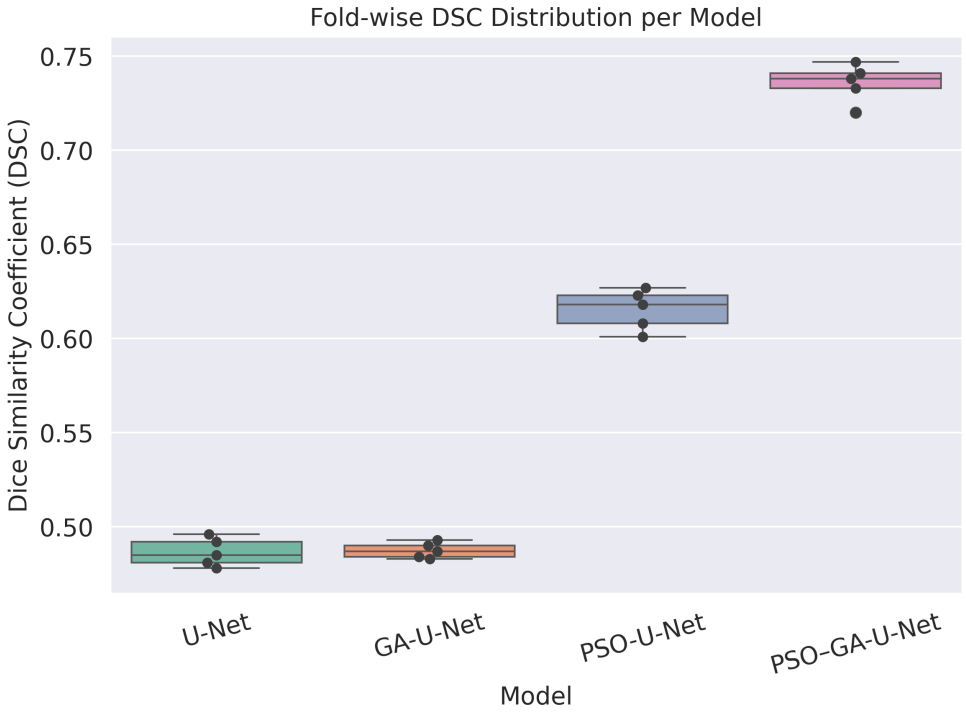


Figure A1. Boxplot showing the fold-wise distribution of DSC scores across all models. PSO-GA-U-Net shows the highest median and lowest interquartile range, reflecting both superior and stable performance.

 doi:10.20944/preprints202603.0466.v1

The boxplot in Figure A1 demonstrates the robustness and consistency of the PSO-GA-U-Net, with a higher median and narrower interquartile range compared to the other variants. This suggests that its performance gains are not only statistically significant but also reliable across different subsets of the dataset.

Appendix A.3. Confidence Interval Estimation

To quantify the uncertainty of mean DSC estimates, 95% confidence intervals were calculated using the standard error across folds. The intervals are as follows:

- U-Net: $[0.4848,\ 0.4882]$
- GA-U-Net: $[0.4859,\ 0.4889]$
- PSO-U-Net: $[0.6124,\ 0.6182]$
- PSO-GA-U-Net: $[0.7328,\ 0.7390]$

These intervals indicate a clear separation between the baseline and optimized models, with PSO-GA showing the most pronounced performance margin. The small overlaps between U-Net and GA-U-Net further support the earlier observation that GA alone contributes marginal improvements.

This appendix reinforces the reliability of the ablation study results presented in the main text and highlights the advantage of hybrid metaheuristic optimization strategies for robust and consistent medical image segmentation.

Appendix B. Distributional Analysis of Cross-Validation Metrics

To provide more in-depth insights into the variability and distribution of segmentation performance across different tumor types, we present violin plots for Dice Similarity Coefficient (DSC) and Jaccard Index (JI) based on five-fold cross-validation results. These visualizations help identify class-wise stability, spread, and potential outliers that may not be evident in summary statistics alone.

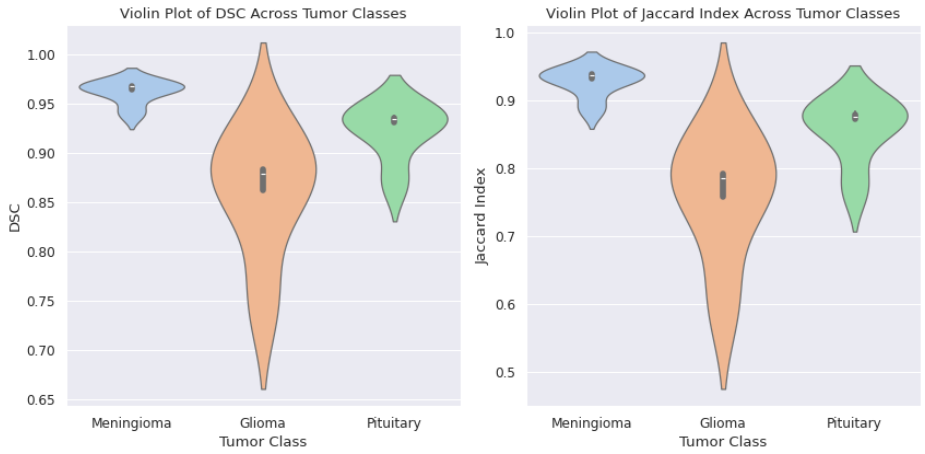


Figure A2. Violin plots showing the distribution of Dice Similarity Coefficient (DSC) and Jaccard Index (JI) across five folds for Meningioma, Glioma, and Pituitary tumor classes on the FBTS dataset. These plots visualize class-wise variability and robustness.

Figure A2 presents the violin distributions of DSC and JI for Meningioma, Glioma, and Pituitary classes. As observed:

- **Meningioma** exhibits narrow, peaked distributions in both DSC and JI, reflecting high stability across folds and low inter-fold variance. The tight interquartile range suggests consistent performance.
- **Glioma** shows a wider spread, particularly in JI. This distribution reflects more variability due to the morphological complexity and heterogeneity of glioma tumors. Although the mean performance is satisfactory, some folds yielded lower values, indicating sensitivity to data partitioning.

 doi:10.20944/preprints202603.0466.v1

- **Pituitary** performance distributions are slightly broader than Meningioma but more concentrated than Glioma, demonstrating strong generalization and moderate variance.

These violin plots align with the statistical analysis in Section 4.3, where the gliomas showed a higher standard deviation in both DSC and JI. This motivates the use of robust augmentation strategies and stratified folds to reduce performance fluctuation.

Appendix C. Distributional Analysis of BraTS 2021 Metrics

To further support the evaluation of model performance across MRI modalities on the BraTS 2021 dataset, we conducted a distributional analysis of segmentation scores. Figure A3 presents violin plots for the Dice Similarity Coefficient (DSC) and the Jaccard Index (JI) across all modalities—FLAIR, T1, T2, and T1CE.

The plots highlight the central tendencies and variability of each modality's performance:

- **FLAIR** shows a tight distribution with high median values for both DSC and JI, reflecting robust segmentation performance and minimal outlier spread.
- **T2** also maintains a high-performing distribution, similar to FLAIR, indicating strong boundary adherence and capture of tumor structures.
- **T1CE** has a wider spread in both metrics, with greater variability, particularly in JI. This suggests that the model experiences inconsistent segmentation accuracy when relying solely on contrast-enhanced T1-weighted input.
- **T1** is between FLAIR and T1CE, with a moderate median but slightly wider interquartile range.

These visualizations reinforce the quantitative statistics provided in the main text, clarifying how modality-specific structural features influence segmentation accuracy and stability.

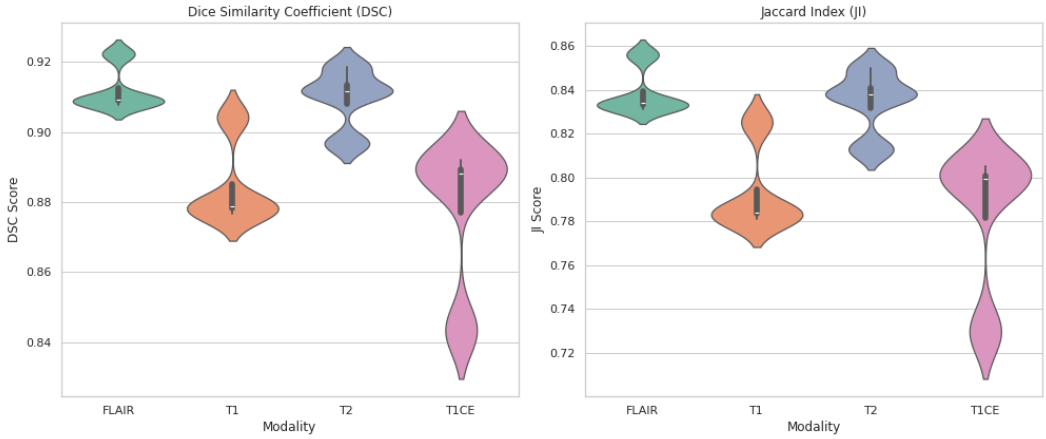


Figure A3. Violin plots of Dice Similarity Coefficient (DSC) and Jaccard Index (JI) across MRI modalities in the BraTS 2021 dataset. Each plot shows the distributional characteristics (median, interquartile range, and density) of scores per modality.

Appendix D. Violin Plot Visualizations of BraTS 2018 (HGG and LGG)

This appendix presents violin plots of the Dice Similarity Coefficient (DSC) and Jaccard Index (JI) scores for BraTS 2018, evaluated separately for High-Grade Glioma (HGG) and Low-Grade Glioma (LGG) cases across all four MRI modalities (FLAIR, T1, T2, and T1CE). These plots highlight the distribution, variability, and density of model performance, offering more in-depth insight into class-wise segmentation consistency.

 doi:10.20944/preprints202603.0466.v1

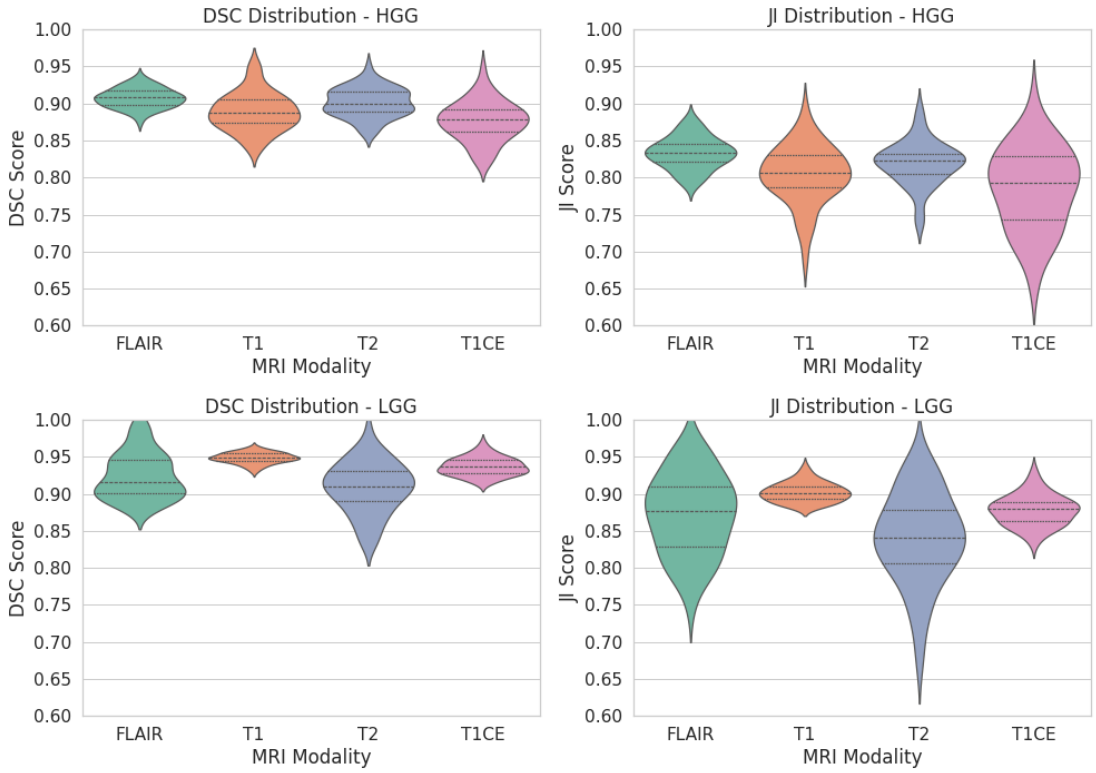


Figure A4. Distribution of DSC and JI scores across four MRI modalities for HGG and LGG tumors in the BraTS 2018 dataset. Each subplot illustrates the performance distribution via violin plots, with inner quartiles and medians highlighted. The top row corresponds to HGG, and the bottom row to LGG. FLAIR and T2 show narrower dispersion and higher central tendencies, while T1CE reflects broader spread.

As shown in Figure A4, the FLAIR and T2 modalities consistently demonstrate narrower distributions and higher median values, suggesting stable segmentation performance in both the HGG and LGG groups. LGG cases, however, tend to show broader JI distributions, reflecting a higher variability in model response—likely due to the subtler, less delineated lesion morphologies typical of low-grade tumors. T1CE exhibited the most dispersed performance, especially in LGG, reinforcing challenges related to contrast enhancement heterogeneity.

These patterns, differentiated by modality and tumor grade, reinforce the clinical value of multimodal imaging and underscore the necessity for architecture robustness across tumor types and contrast characteristics.

References

1. Zhang, J.; Zhang, J.; Yang, C. Autophagy in brain tumors: molecular mechanisms, challenges, and therapeutic opportunities. *Journal of Translational Medicine* **2025**, *23*, 52. https://doi.org/10.1186/s12967-024-06063-0.
2. Batool, A.; Byun, Y.C. Brain tumor detection with integrating traditional and computational intelligence approaches across diverse imaging modalities - Challenges and future directions. *Computers in Biology and Medicine* **2024**, p. 108412. https://doi.org/10.1016/j.compbiomed.2024.108412.
3. Ghadimi, D.J.; Vahdani, A.M.; Karimi, H.; Ebrahimi, P.; Fathi, M.; Moodi, F.; Habibzadeh, A.; Khodadadi Shoushtari, F.; Valizadeh, G.; Mobarak Salari, H.; et al. Deep Learning-Based Techniques in Glioma Brain Tumor Segmentation Using Multi-Parametric MRI: A Review on Clinical Applications and Future Outlooks. *Journal of Magnetic Resonance Imaging* **2024**. https://doi.org/10.1002/jmri.29543.
4. Chukwujindu, E.; Faiz, H.; AI-Douri, S.; Faiz, K.; De Sequeira, A. Role of artificial intelligence in brain tumour imaging. *European Journal of Radiology* **2024**, *176*, 111509. https://doi.org/10.1016/j.ejrad.2024.111509.
5. Rasool, N.; Bhat, J.I. A Critical Review on Segmentation of Glioma Brain Tumor and Prediction of Overall Survival. *Archives of Computational Methods in Engineering* **2024**. https://doi.org/10.1007/s11831-024-10188-2.

 doi:10.20944/preprints202603.0466.v1

6. Datta, P.; Rohilla, R. Comprehensive Survey on Computational Techniques for Brain Tumor Detection: Past, Present and Future. *Archives of Computational Methods in Engineering* **2025**. https://doi.org/10.1007/s11831-025-10238-3.
7. Sreedhar, D. Evaluating the Clinical Applicability of Neural Networks for Meningioma Tumor Segmentation on Multiparametric 3D MRI. In Proceedings of the 2024 International Conference on Machine Learning and Applications (ICMLA), Miami, FL, USA, 2024; pp. 1308–1313. https://doi.org/10.1109/ICMLA61862.2024.00204.
8. Sajid Hussain, S.; Wani, N.A.; Kaur, J.; Ahmad, N.; Ahmad, S. Next-Generation Automation in Neuro-Oncology: Advanced Neural Networks for MRI-Based Brain Tumor Segmentation and Classification. *IEEE Access* **2025**, *13*, 41141–41158. https://doi.org/10.1109/ACCESS.2025.3547796.
9. Bonato, B.; Nanni, L.; Bertoldo, A. Advancing Precision: A Comprehensive Review of MRI Segmentation Datasets from BraTS Challenges (2012–2025). *Sensors* **2025**, *25*, 1838. https://doi.org/10.3390/s25061838.
10. Pani, K.; Chawla, I. Synthetic MRI in action: A novel framework in data augmentation strategies for robust multi-modal brain tumor segmentation. *Computers in Biology and Medicine* **2024**, *183*, 109273. https://doi.org/10.1016/j.compbiomed.2024.109273.
11. Shivani.; Agrawal, K.K.; Agrawal, G. Precision diagnosis of brain tumors: An overview of advanced machine learning techniques. In *Applications of Artificial Intelligence in 5G and Internet of Things*; CRC Press: London, 2025; pp. 112–118. https://doi.org/10.1201/9781003532521-21.
12. Umarani, C.M.; Gollagi, S.; Allagi, S.; Sambrekar, K.; Ankali, S.B. Advancements in deep learning techniques for brain tumor segmentation: A survey. *Informatics in Medicine Unlocked* **2024**, *50*, 101576. https://doi.org/10.1016/j.imu.2024.101576.
13. Das, S.; Goswami, R.S. Advancements in brain tumor analysis: a comprehensive review of machine learning, hybrid deep learning, and transfer learning approaches for MRI-based classification and segmentation. *Multimedia Tools and Applications* **2024**. https://doi.org/10.1007/s11042-024-20203-0.
14. Azad, R.; Aghdam, E.K.; Rauland, A.; Jia, Y.; Avval, A.H.; Bozorgpour, A.; Karimijafarbigloo, S.; Cohen, J.P.; Adeli, E.; Merhof, D. Medical Image Segmentation Review: The Success of U-Net. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2024**, *46*, 10076–10095. https://doi.org/10.1109/TPAMI.2024.3435571.
15. Punn, N.S.; Agarwal, S. Modality specific U-Net variants for biomedical image segmentation: a survey. *Artificial Intelligence Review* **2022**, *55*, 5845–5889. https://doi.org/10.1007/s10462-022-10152-1.
16. Siddique, N.; Paheding, S.; Elkin, C.P.; Devabhaktuni, V. U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications. *IEEE Access* **2021**, *9*, 82031–82057. https://doi.org/10.1109/ACCESS.2021.3086020.
17. Du, G.; Cao, X.; Liang, J.; Chen, X.; Zhan, Y. Medical Image Segmentation based on U-Net: A Review. *Journal of Imaging Science and Technology* **2020**, *64*, 020508–1–020508–12. https://doi.org/10.2352/J.ImagingSci.Technol.2020.64.2.020508.
18. Saifullah, S.; Drężewski, R. Automatic Brain Tumor Segmentation Using Convolutional Neural Networks: U-Net Framework with PSO-Tuned Hyperparameters. In *Affenzeller, M., et al. Parallel Problem Solving from Nature – PPSN XVIII. PPSN 2024. Lecture Notes in Computer Science, vol 15150. Springer, Cham.*; Springer: Cham, 2024; pp. 333–351. https://doi.org/10.1007/978-3-031-70071-2_21.
19. Asiri, A.A.; Shaf, A.; Ali, T.; Aamir, M.; Irfan, M.; Alqahtani, S. Enhancing brain tumor diagnosis: an optimized CNN hyperparameter model for improved accuracy and reliability. *PeerJ Computer Science* **2024**, *10*, e1878. https://doi.org/10.7717/peerj-cs.1878.
20. Zhang, Y.; Ngo, H.C.; Zhang, Y.; Yusof, N.F.A.; Wang, X. Imaging Segmentation of Brain Tumors Based on the Modified U-net Method. *Information Technology and Control* **2024**, *53*, 1074–1087. https://doi.org/10.5755/j01.itc.53.4.37719.
21. Ali, S.; Khurram, R.; ur Rehman, K.; Yasin, A.; Shaukat, Z.; Sakhawat, Z.; Mujtaba, G. An improved 3D U-Net-based deep learning system for brain tumor segmentation using multi-modal MRI. *Multimedia Tools and Applications* **2024**, *83*, 85027–85046. https://doi.org/10.1007/s11042-024-19406-2.
22. Malik, A.; Devarajan, G.G. Integrated Brain Tumor Detection: PSO-Guided Segmentation with U-Net and CNN Classification. *Procedia Computer Science* **2024**, *235*, 3447–3457. https://doi.org/10.1016/j.procs.2024.04.325.
23. Raza, A.; Bin Musa, S.; Shahrafidz Bin Khalid, A.; Mansoor Alam, M.; Mohd Su'ud, M.; Noor, F. Enhancing Medical Image Classification Through PSO-Optimized Dual Deterministic Approach and Robust Transfer Learning. *IEEE Access* **2024**, *12*, 177144–177159. https://doi.org/10.1109/ACCESS.2024.3504266.

 doi:10.20944/preprints202603.0466.v1

24. Saifullah, S.; Drężewski, R. Particle Swarm-Optimized U-Net Framework for Precise Multimodal Brain Tumor Segmentation. In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '25 Companion), July 14–18, 2025, Malaga, Spain, New York, NY, USA, 2025; p. to appear. https://doi.org/10.1145/3712255.3726561.
25. Shanthi, D.L.; Chethan, N. Genetic Algorithm Based Hyper-Parameter Tuning to Improve the Performance of Machine Learning Models. *SN Computer Science* **2022**, *4*, 119. https://doi.org/10.1007/s42979-022-01537-8.
26. Raji, I.D.; Bello-Salau, H.; Umoh, I.J.; Onumanyi, A.J.; Adegboye, M.A.; Salawudeen, A.T. Simple Deterministic Selection-Based Genetic Algorithm for Hyperparameter Tuning of Machine Learning Models. *Applied Sciences* **2022**, *12*, 1186. https://doi.org/10.3390/app12031186.
27. Japa, L.; Serqueira, M.; MendonçA, I.; Aritsugi, M.; Bezerra, E.; González, P.H. A Population-Based Hybrid Approach for Hyperparameter Optimization of Neural Networks. *IEEE Access* **2023**, *11*, 50752–50768. https://doi.org/10.1109/ACCESS.2023.3277310.
28. Jyothi, P.; Singh, A.R. Deep learning models and traditional automated techniques for brain tumor segmentation in MRI: a review. *Artificial Intelligence Review* **2023**, *56*, 2923–2969. https://doi.org/10.1007/s10462-022-10245-x.
29. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *International Conference on Medical image computing and computer-assisted intervention*; Springer, Cham: Cham, 2015; pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.
30. Saifullah, S.; Drężewski, R.; Yudhana, A.; Wielgosz, M.; Caesarendra, W. Modified U-Net with attention gate for enhanced automated brain tumor segmentation. *Neural Computing and Applications* **2025**, *37*. https://doi.org/10.1007/s00521-024-10919-3.
31. Jiangtao, W.; Ruhaiyem, N.I.R.; Panpan, F. A Comprehensive Review of U-Net and Its Variants: Advances and Applications in Medical Image Segmentation. *IET Image Processing* **2025**, *19*. https://doi.org/10.1049/ipr2.70019.
32. Zhang, C.; Achuthan, A.; Himel, G.M.S. State-of-the-Art and Challenges in Pancreatic CT Segmentation: A Systematic Review of U-Net and Its Variants. *IEEE Access* **2024**, *12*, 78726–78742. https://doi.org/10.1109/ACCESS.2024.3392595.
33. Saifullah, S.; Drężewski, R. Automatic Brain Tumor Segmentation Using Convolutional Neural Networks: U-Net Framework with PSO-Tuned Hyperparameters. In *Affenzeller, M., et al. Parallel Problem Solving from Nature – PPSN XVIII. PPSN 2024. Lecture Notes in Computer Science, vol 15150. Springer, Cham.*; Springer, Cham, 2024; pp. 333–351. https://doi.org/10.1007/978-3-031-70071-2_21.
34. Zhang, J.; Jiang, Z.; Dong, J.; Hou, Y.; Liu, B. Attention Gate ResU-Net for Automatic MRI Brain Tumor Segmentation. *IEEE Access* **2020**, *8*, 58533–58545. https://doi.org/10.1109/ACCESS.2020.2983075.
35. Zhang, Q.; Hang, Y.; Qiu, J.; Chen, H. Application of U-Net Network Utilizing Multiattention Gate for MRI Segmentation of Brain Tumors. *Journal of Computer Assisted Tomography* **2024**, *48*, 991–997. https://doi.org/10.1097/RCT.0000000000001641.
36. Koteswara Rao Chinnam, S.; Sistla, V.; Krishna Kishore Kolli, V. Multimodal attention-gated cascaded U-Net model for automatic brain tumor detection and segmentation. *Biomedical Signal Processing and Control* **2022**, *78*, 103907. https://doi.org/10.1016/j.bspc.2022.103907.
37. Rahim Khan, W.; Mustafa Madni, T.; Iqbal Janjua, U.; Javed, U.; Attique Khan, M.; Alhaisoni, M.; Tariq, U.; Cha, J.H. A Hybrid Attention-Based Residual Unet for Semantic Segmentation of Brain Tumor. *Computers, Materials & Continua* **2023**, *76*, 647–664. https://doi.org/10.32604/cmc.2023.039188.
38. Ullah, Z.; Usman, M.; Jeon, M.; Gwak, J. Cascade multiscale residual attention CNNs with adaptive ROI for automatic brain tumor segmentation. *Information Sciences* **2022**, *608*, 1541–1556. https://doi.org/10.1016/j.ins.2022.07.044.
39. Metlek, S.; Çetıner, H. ResUNet+: A New Convolutional and Attention Block-Based Approach for Brain Tumor Segmentation. *IEEE Access* **2023**, *11*, 69884–69902. https://doi.org/10.1109/ACCESS.2023.3294179.
40. Pandey, A.K.; Singh, S.P.; Chakraborty, C. Residual attention UNet GAN Model for enhancing the intelligent agents in retinal image analysis. *Service Oriented Computing and Applications* **2024**. https://doi.org/10.1007/s11761-024-00415-w.
41. Saifullah, S.; Drężewski, R. Redefining brain tumor segmentation: a cutting-edge convolutional neural networks-transfer learning approach. *International Journal of Electrical and Computer Engineering (IJECE)* **2024**, *14*, 2583. https://doi.org/10.11591/ijece.v14i3.pp2583-2591.
42. Saifullah, S.; Drężewski, R. Brain Tumor Segmentation Using Ensemble CNN-Transfer Learning Models: DeepLabV3plus and ResNet50 Approach. In *In: Franco, L., de Mulatier, C., Paszynski, M., Krzhizhanovskaya,*

 doi:10.20944/preprints202603.0466.v1

V.V., Dongarra, J.J., Sloot, P.M.A. (eds) Computational Science – ICCS 2024. ICCS 2024. Lecture Notes in Computer Science, vol 14835. Springer, Cham.; Springer: Cham, 2024; pp. 340–354. https://doi.org/10.1007/978-3-031-63772-8_30.

43. Vijay, S.; Guhan, T.; Srinivasan, K.; Vincent, P.M.D.R.; Chang, C.Y. MRI brain tumor segmentation using residual Spatial Pyramid Pooling-powered 3D U-Net. *Frontiers in Public Health* **2023**, *11*. https://doi.org/10.3389/fpubh.2023.1091850.
44. Saifullah, S.; Drężewski, R.; Yudhana, A.; Caesarendra, W.; Huda, N. Bio-Inspired Metaheuristics in Deep Learning for Brain Tumor Segmentation: A Decade of Advances and Future Directions. *Information* **2025**, *16*, 456. https://doi.org/10.3390/info16060456.
45. Siddique, A.A.; Raza, A.; Alshehri, M.S.; Alasbali, N.; Abbasi, S.F. Optimizing Tumor Classification Through Transfer Learning and Particle Swarm Optimization-Driven Feature Extraction. *IEEE Access* **2024**, *12*, 85929–85939. https://doi.org/10.1109/ACCESS.2024.3412412.
46. Liu, L.; Chang, J.; Liang, G.; Xiong, S. Simulated Quantum Mechanics-Based Joint Learning Network for Stroke Lesion Segmentation and TICI Grading. *IEEE Journal of Biomedical and Health Informatics* **2023**, *27*, 3372–3383. https://doi.org/10.1109/JBHI.2023.3270861.
47. Saifullah, S.; Drężewski, R. Advanced Medical Image Segmentation Enhancement: A Particle-Swarm-Optimization-Based Histogram Equalization Approach. *Applied Sciences* **2024**, *14*, 923. https://doi.org/10.3390/app14020923.
48. Wortmann, T. Genetic evolution vs. function approximation: Benchmarking algorithms for architectural design optimization. *Journal of Computational Design and Engineering* **2019**, *6*, 414–428. https://doi.org/10.1016/j.jcde.2018.09.001.
49. Preethi, B.M.; Lekha, J.; Seethalakshmy, A.; Gokul, S. Adaptive Feature Selection for Brain Tumor Classification in MRI Images using Genetic Algorithm Polar Bear Optimization and SVM. *Journal of Computational Analysis and Applications (JoCAAA)* **2024**, *33*, 356–375.
50. Saifullah, S.; Drężewski, R. Optimizing U-Net Architecture Using Differential Evolution for Brain Tumor Segmentation. In *Lees, M.H., et al. Computational Science – ICCS 2025. ICCS 2025. Lecture Notes in Computer Science*; Springer: Cham, 2025; Vol. 15906, pp. 403–411. https://doi.org/10.1007/978-3-031-97635-3_48.
51. Chihaoui, M.; Dhibi, N.; Ferchichi, A. Optimization of convolutional neural network and visual geometry group-16 using genetic algorithms for pneumonia detection. *Frontiers in Medicine* **2024**, *11*. https://doi.org/10.3389/fmed.2024.1498403.
52. Wei, J.; Zhu, G.; Fan, Z.; Liu, J.; Rong, Y.; Mo, J.; Li, W.; Chen, X. Genetic U-Net: Automatically Designed Deep Networks for Retinal Vessel Segmentation Using a Genetic Algorithm. *IEEE Transactions on Medical Imaging* **2022**, *41*, 292–307. https://doi.org/10.1109/TMI.2021.3111679.
53. Khouy, M.; Jabrane, Y.; Ameur, M.; Hajjam El Hassani, A. Medical Image Segmentation Using Automatic Optimized U-Net Architecture Based on Genetic Algorithm. *Journal of Personalized Medicine* **2023**, *13*, 1298. https://doi.org/10.3390/jpm13091298.
54. Dubey, R.; Agrawal, J. An Improved Genetic Algorithm for Automated Convolutional Neural Network Design. *Intelligent Automation & Soft Computing* **2022**, *32*, 747–763. https://doi.org/10.32604/iasc.2022.020975.
55. Anwaar, A.; Ashraf, A.; Bangyal, W.H.K.; Iqbal, M. Genetic Algorithms: Brief review on Genetic Algorithms for Global Optimization Problems. In Proceedings of the 2022 Human-Centered Cognitive Systems (HCCS), New York City, U.S., 2022; pp. 1–6. https://doi.org/10.1109/HCCS55241.2022.10090327.
56. Jiang, P.; Xue, Y.; Neri, F. Continuously evolving dropout with multi-objective evolutionary optimisation. *Engineering Applications of Artificial Intelligence* **2023**, *124*, 106504. https://doi.org/10.1016/j.engappai.2023.106504.
57. Arif, M.; Jims, A.; F., A.; Geman, O.; Craciun, M.D.; Leuciuc, F. Application of Genetic Algorithm and U-Net in Brain Tumor Segmentation and Classification: A Deep Learning Approach. *Computational Intelligence and Neuroscience* **2022**, *2022*, 1–11. https://doi.org/10.1155/2022/5625757.
58. Ghazouani, F.; Vera, P.; Ruan, S. Efficient brain tumor segmentation using Swin transformer and enhanced local self-attention. *International Journal of Computer Assisted Radiology and Surgery* **2023**, *19*, 273–281. https://doi.org/10.1007/s11548-023-03024-8.
59. Nancy, A.M.; Maheswari, R. Brain tumor segmentation and classification using transfer learning based CNN model with model agnostic concept interpretation. *Multimedia Tools and Applications* **2024**, *84*, 2509–2538. https://doi.org/10.1007/s11042-024-20353-1.

 doi:10.20944/preprints202603.0466.v1

60. Jiang, Z.; Ding, C.; Liu, M.; Tao, D. Two-Stage Cascaded U-Net: 1st Place Solution to BraTS Challenge 2019 Segmentation Task. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2019. Lecture Notes in Computer Science*; Crimi, A.; Bakas, S., Eds.; Springer: Cham, 2020; pp. 231–241. https://doi.org/10.1007/978-3-030-46640-4_22.
61. Yousef, R.; Khan, S.; Gupta, G.; Albahlal, B.M.; Alajlan, S.A.; Ali, A. Bridged-U-Net-ASPP-EVO and Deep Learning Optimization for Brain Tumor Segmentation. *Diagnostics* **2023**, *13*, 2633. https://doi.org/10.3390/diagnostics13162633.
62. Cheng, J.; Huang, W.; Cao, S.; Yang, R.; Yang, W.; Yun, Z.; Wang, Z.; Feng, Q. Enhanced Performance of Brain Tumor Classification via Tumor Region Augmentation and Partition. *PLOS ONE* **2015**, *10*, e0140381. https://doi.org/10.1371/journal.pone.0140381.
63. Baid, U.; Ghodasara, S.; Mohan, S.; Bilello, M.; Calabrese, E.; Colak, E.; Farahani, K.; Kalpathy-Cramer, J.; Kitamura, F.C.; Pati, S.; et al. RSNA-ASNR-MICCAI-BraTS-2021 Dataset, 2023. https://doi.org/10.7937/jc8x-9874.
64. Menze, B.H.; Jakab, A.; Bauer, S.; Kalpathy-Cramer.; et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Transactions on Medical Imaging* **2015**, *34*, 1993–2024. https://doi.org/10.1109/TMI.2014.2377694.
65. Bakas, S.; Akbari, H.; Sotiras, A.; Bilello, M.; Rozycki, M.; Kirby, J.S.; Freymann, J.B.; Farahani, K.; Davatzikos, C. Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. *Scientific Data* **2017**, *4*, 170117. https://doi.org/10.1038/sdata.2017.117.
66. Bakas, S.; Reyes, M.; Jakab, A.; Bauer, S.; et al. Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge, 2018, [1811.02629].
67. Abou Ali, M.; Charafeddine, J.; Dornaika, F.; Arganda-Carreras, I. Enhancing Generalization and Mitigating Overfitting in Deep Learning for Brain Cancer Diagnosis from MRI. *Applied Magnetic Resonance* **2025**, *56*, 359–394. https://doi.org/10.1007/s00723-024-01743-y.
68. Tran, A.T.; Zeevi, T.; Payabvash, S. Strategies to Improve the Robustness and Generalizability of Deep Learning Segmentation and Classification in Neuroimaging. *BioMedInformatics* **2025**, *5*, 20. https://doi.org/10.3390/biomedinformatics5020020.
69. Li, Y.; Ammari, S.; Balleyguier, C.; Lassau, N.; Chouzenoux, E. Impact of Preprocessing and Harmonization Methods on the Removal of Scanner Effects in Brain MRI Radiomic Features. *Cancers* **2021**, *13*, 3000. https://doi.org/10.3390/cancers13123000.
70. Vitale, S.; Orlando, J.I.; Iarussi, E.; Díaz, A.; Larrabide, I. Improving realism in abdominal ultrasound simulation combining a segmentation-guided loss and polar coordinates training. *Medical Physics* **2025**, *52*, 4540–4556. https://doi.org/10.1002/mp.17801.
71. Parimanam, K.; Lakshmanan, L.; Palaniswamy, T. Hybrid optimization based learning technique for multi-disease analytics from healthcare big data using optimal pre-processing, clustering and classifier. *Concurrency and Computation: Practice and Experience* **2022**, *34*. https://doi.org/10.1002/cpe.6986.
72. Mahmud, M.I.; Mamun, M.; Abdelgawad, A. A Deep Analysis of Brain Tumor Detection from MR Images Using Deep Learning Networks. *Algorithms* **2023**, *16*, 176. https://doi.org/10.3390/a16040176.
73. Ghaffar, A.; Javid, M.A.; Yaseen, K.; Ali, N.; Arshad, S.; El-Bahkiry, H.S.; Hassani, M.K.; Akgül, A. Innovative fusion: MRSI-guided brain tumour classification via integrated image segmentation and GLCM feature extraction. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* **2025**, *13*. https://doi.org/10.1080/21681163.2025.2479707.
74. Pal, A.; Kruk, J.; Phute, M.; Bhattaram, M.; Yang, D.; Chau, D.H.; Hoffman, J. Semi-Truths: A Large-Scale Dataset of AI-Augmented Images for Evaluating Robustness of AI-Generated Image detectors. In Proceedings of the Advances in Neural Information Processing Systems; Globerson, A.; Mackey, L.; Belgrave, D.; Fan, A.; Paquet, U.; Tomczak, J.; Zhang, C., Eds., New York, 2024; Vol. 37, pp. 118025–118051.
75. Singha, A.; Thakur, R.S.; Patel, T. Deep Learning Applications in Medical Image Analysis. In *Biomedical Data Mining for Information Retrieval*; Wiley: USA, 2021; pp. 293–350. https://doi.org/10.1002/9781119711278.ch11.
76. Pang, J.; Li, X.; Han, S. PSO with Mixed Strategy for Global Optimization. *Complexity* **2023**, *2023*, 1–19. https://doi.org/10.1155/2023/7111548.
77. Abualigah, L.; Sheikhan, A.; M. Ikotun, A.; Zitar, R.A.; Alsoud, A.R.; Al-Shourbaji, I.; Hussien, A.G.; Jia, H. Particle swarm optimization algorithm: review and applications. In *Metaheuristic Optimization Algorithms*; Elsevier: Netherlands, 2024; pp. 1–14. https://doi.org/10.1016/B978-0-443-13925-3.00019-4.

 doi:10.20944/preprints202603.0466.v1

78. Gad, A.G. Particle Swarm Optimization Algorithm and Its Applications: A Systematic Review. *Archives of Computational Methods in Engineering* **2022**, *29*, 2531–2561. https://doi.org/10.1007/s11831-021-09694-4.
79. Dar, M.F.; Ganivada, A. Deep learning and genetic algorithm-based ensemble model for feature selection and classification of breast ultrasound images. *Image and Vision Computing* **2024**, *146*, 105018. https://doi.org/10.1016/j.imavis.2024.105018.
80. El Abassi, F.; Darouichi, A.; Ouaarab, A. Refining U-Net Architecture Through Genetic Algorithms for Improved Skin Lesion Image Segmentation. In *Bennour, A., Bouridane, A., Almaadeed, S., Bouaziz, B., Edirisinghe, E. (eds) Intelligent Systems and Pattern Recognition. ISPR 2024. Communications in Computer and Information Science*; Springer: Cham, 2025; Vol. 2305, pp. 135–149. https://doi.org/10.1007/978-3-031-82156-1_11.
81. Das, S.; ku. Swain, M.; Nayak, G.K.; Saxena, S.; Satpathy, S.C. Effect of learning parameters on the performance of U-Net Model in segmentation of Brain tumor. *Multimedia Tools and Applications* **2022**, *81*, 34717–34735. https://doi.org/10.1007/s11042-021-11273-5.
82. Maniatopoulos, A.; Mitianoudis, N. Learnable Leaky ReLU (LeLeLU): An Alternative Accuracy-Optimized Activation Function. *Information* **2021**, *12*, 513. https://doi.org/10.3390/info12120513.
83. Varshney, M.; Singh, P. Optimizing nonlinear activation function for convolutional neural networks. *Signal, Image and Video Processing* **2021**, *15*, 1323–1330. https://doi.org/10.1007/s11760-021-01863-z.
84. Krithika alias AnbuDevi, M.; Suganthi, K. Review of Semantic Segmentation of Medical Images Using Modified Architectures of UNET. *Diagnostics* **2022**, *12*, 3064. https://doi.org/10.3390/diagnostics12123064.
85. Saifullah, S.; Drężewski, R.; Yudhana, A. Advanced brain tumor segmentation using DeepLabV3Plus with Xception encoder on a multi-class MR image dataset. *Multimedia Tools and Applications* **2025**. https://doi.org/10.1007/s11042-025-20702-8.
86. Irfan, M.; Shaf, A.; Ali, T.; Farooq, U.; Rahman, S.; Nasar Faraj Mursal, S.; Jalalah, M.; M. Alqhtani, S.; AlShorman, O. Effectiveness of Deep Learning Models for Brain Tumor Classification and Segmentation. *Computers, Materials & Continua* **2023**, *76*, 711–729. https://doi.org/10.32604/cmc.2023.038176.
87. Saifullah, S.; Drężewski, R.; Yudhana, A.; Suryotomo, A.P. Automatic Brain Tumor Segmentation: Advancing U-Net With ResNet50 Encoder for Precise Medical Image Analysis. *IEEE Access* **2025**, *13*, 43473–43489. https://doi.org/10.1109/ACCESS.2025.3547430.
88. Murmu, A.; Kumar, P. A novel Gateaux derivatives with efficient DCNN-Resunet method for segmenting multi-class brain tumor. *Medical & Biological Engineering & Computing* **2023**, *61*, 2115–2138. https://doi.org/10.1007/s11517-023-02824-z.
89. Kumar Tiwary, P.; Johri, P.; Katiyar, A.; Chhipa, M.K. Deep Learning-Based MRI Brain Tumor Segmentation With EfficientNet-Enhanced UNet. *IEEE Access* **2025**, *13*, 54920–54937. https://doi.org/10.1109/ACCESS.2025.3554405.
90. R, P.; M, J.P.P.; J S, N. Brain tumor segmentation using multi-scale attention U-Net with EfficientNetB4 encoder for enhanced MRI analysis. *Scientific Reports* **2025**, *15*, 9914. https://doi.org/10.1038/s41598-025-94267-9.
91. Davar, S.; Fevens, T. Enhanced U-Net Architecture for Brain Tumour Localization and Segmentation in T1-Weighted MRI. *IEEE Transactions on Circuits and Systems II: Express Briefs* **2025**, pp. 1–1. https://doi.org/10.1109/TCSII.2025.3556846.
92. Xiong, M.; Wu, A.; Yang, Y.; Fu, Q. Efficient Brain Tumor Segmentation for MRI Images Using YOLO-BT. *Sensors* **2025**, *25*, 3645. https://doi.org/10.3390/s25123645.
93. Alkhalid, F.F.; Salih, N.Z. Implementation of biomedical segmentation for brain tumor utilizing an adapted U-net model. *Computers in Biology and Medicine* **2025**, *194*, 110531. https://doi.org/10.1016/j.compbiomed.2025.110531.
94. El Badaoui, R.; Bonmati Coll, E.; Psarrou, A.; Asaturyan, H.A.; Villarini, B. Enhanced CATBraTS for Brain Tumour Semantic Segmentation. *Journal of Imaging* **2025**, *11*, 8. https://doi.org/10.3390/jimaging11010008.
95. Sachdeva, J.; Sharma, D.; Ahuja, C.K. Multiscale segmentation net for segregating heterogeneous brain tumors: Gliomas on multimodal MR images. *Image and Vision Computing* **2024**, *149*, 105191. https://doi.org/10.1016/j.imavis.2024.105191.
96. Hernandez-Gutierrez, F.D.; Avina-Bravo, E.G.; Zambrano-Gutierrez, D.F.; Almanza-Conejo, O.; Ibarra-Manzano, M.A.; Ruiz-Pinales, J.; Ovalle-Magallanes, E.; Avina-Cervantes, J.G. Brain Tumor Segmentation from Optimal MRI Slices Using a Lightweight U-Net. *Technologies* **2024**, *12*, 183. https://doi.org/10.3390/technologies12100183.
97. Mojtahedi, R.; Hamghalam, M.; Simpson, A.L. Multi-modal Brain Tumour Segmentation Using Transformer with Optimal Patch Size. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes*

 doi:10.20944/preprints202603.0466.v1

2022. Lecture Notes in Computer Science; Bakas, S.; et al., Eds.; Springer: Cham, 2023; Vol. 13769, pp. 195–204. https://doi.org/10.1007/978-3-031-33842-7_17.

98. Bouchet, P.; Deloges, J.B.; Canton-Bacara, H.; Pusel, G.; Pinot, L.; Elbaz, O.; Boutry, N. An Efficient Cascade of U-Net-Like Convolutional Neural Networks Devoted to Brain Tumor Segmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2022. Lecture Notes in Computer Science*; Bakas, S.; et al., Eds.; Springer: Cham, 2023; Vol. 13769, pp. 149–161. https://doi.org/10.1007/978-3-031-33842-7_13.
99. Sadique, M.S.; Rahman, M.M.; Farzana, W.; Temtam, A.; Iftekharuddin, K.M. Brain Tumor Segmentation Using Neural Ordinary Differential Equations with UNet-Context Encoding Network. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2022. Lecture Notes in Computer Science*; Bakas, S.; et al., Eds.; Springer: Cham, 2023; pp. 205–215. https://doi.org/10.1007/978-3-031-33842-7_18.
100. Hu, K.; Gan, Q.; Zhang, Y.; Deng, S.; Xiao, F.; Huang, W.; Cao, C.; Gao, X. Brain Tumor Segmentation Using Multi-Cascaded Convolutional Neural Networks and Conditional Random Field. *IEEE Access* **2019**, *7*, 92615–92629. https://doi.org/10.1109/ACCESS.2019.2927433.
101. Li, X.; Jiang, Y.; Li, M.; Zhang, J.; Yin, S.; Luo, H. MSFR-Net: Multi-modality and single-modality feature recalibration network for brain tumor segmentation. *Medical Physics* **2023**, *50*, 2249–2262. https://doi.org/10.1002/mp.15933.
102. Dash, S.; Mishra, S.; Siddique, M.; Gelmecha, D.J.; Singh, R.S. Improved Deviation Sparse Fuzzy C-Means-2D Cumulative Sum Average Filter and Modified Sine Cosine Crow Search Algorithm-Wavelet Extreme Learning Machine for Brain Tumor Detection and Classification. *Applied Computational Intelligence and Soft Computing* **2025**, *2025*. https://doi.org/10.1155/acis/9991264.
103. Saeed, M.U.; Ali, G.; Bin, W.; Almotiri, S.H.; AlGhamdi, M.A.; Nagra, A.A.; Masood, K.; ul Amin, R. RMU-Net: A Novel Residual Mobile U-Net Model for Brain Tumor Segmentation from MR Images. *Electronics* **2021**, *10*, 1962. https://doi.org/10.3390/electronics10161962.
104. Zhou, T.; Wang, Z.; Liu, X.; Liu, W.; Zhu, S. Learning deep feature representations for multi-modal MR brain tumor segmentation. *Neurocomputing* **2025**, *638*, 130162. https://doi.org/10.1016/j.neucom.2025.130162.
105. Zhou, C.; Chen, S.; Ding, C.; Tao, D. Learning Contextual and Attentive Information for Brain Tumor Segmentation. In *Crimi, A., Bakas, S., Kuijf, H., Keyvan, F., Reyes, M., van Walsum, T. (eds) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lecture Notes in Computer Science(), vol 11384*; Springer: Cham, 2019; Vol. 11384, pp. 497–507. https://doi.org/10.1007/978-3-030-11726-9_44.
106. Ullah, F.; Ansari, S.U.; Hanif, M.; Ayari, M.A.; Chowdhury, M.E.H.; Khandakar, A.A.; Khan, M.S. Brain MR Image Enhancement for Tumor Segmentation Using 3D U-Net. *Sensors* **2021**, *21*, 7528. https://doi.org/10.3390/s21227528.
107. Wang, G.; Li, W.; Ourselin, S.; Vercauteren, T. Automatic Brain Tumor Segmentation Using Convolutional Neural Networks with Test-Time Augmentation. In *Crimi, A., Bakas, S., Kuijf, H., Keyvan, F., Reyes, M., van Walsum, T. (eds) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lecture Notes in Computer Science*; Springer: Cham, 2019; Vol. 11384, pp. 61–72. https://doi.org/10.1007/978-3-030-11726-9_6.
108. Albiol, A.; Albiol, A.; Albiol, F. Extending 2D Deep Learning Architectures to 3D Image Segmentation Problems. In *Crimi, A., Bakas, S., Kuijf, H., Keyvan, F., Reyes, M., van Walsum, T. (eds) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lecture Notes in Computer Science*; Springer: Cham, 2019; Vol. 11384, pp. 73–82. https://doi.org/10.1007/978-3-030-11726-9_7.
109. Rehman, M.U.; Cho, S.; Kim, J.; Chong, K.T. BrainSeg-Net: Brain Tumor MR Image Segmentation via Enhanced Encoder–Decoder Network. *Diagnostics* **2021**, *11*, 169. https://doi.org/10.3390/diagnostics11020169.
110. Dong, H.; Yang, G.; Liu, F.; Mo, Y.; Guo, Y. Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks. In *Valdés Hernández, M., González-Castro, V. (eds) Medical Image Understanding and Analysis. MIUA 2017. Communications in Computer and Information Science*; Springer: Cham, 2017; Vol. 723, pp. 506–517. https://doi.org/10.1007/978-3-319-60964-5_44.