Lightweight Grey Wolf Optimization–Driven Hyperparameter Tuning of U-Net for Robust Brain Tumor Segmentation

Shoffan Saifullah[1,2][0000−0001−6799−3834], Rafał Dreżewski[1][0000−0001−8607−3478], and Anton Yudhana[3][0000−0001−5451−454X]

[1] Faculty of Computer Science, AGH University of Krakow, Krakow 30-059, Poland
{saifulla,drezew}@agh.edu.pl
[2] Department of Informatics, Universitas Pembangunan Nasional Veteran Yogyakarta, Yogyakarta 55281, Indonesia shoffans@upnyk.ac.id
[3] Department of Electrical Engineering, Universitas Ahmad Dahlan, Yogyakarta 55191, Indonesia eyudhana@ee.uad.ac.id

Abstract. Accurate brain tumor segmentation from magnetic resonance imaging (MRI) is essential for clinical diagnosis and treatment planning, yet convolutional neural network performance is highly sensitive to hyperparameter selection. This paper proposes a lightweight Grey Wolf Optimization (GWO)–driven framework for automated hyperparameter tuning of a parameterized U-Net. The approach jointly optimizes architectural and training parameters under a constrained evaluation budget using reduced-resolution training, enabling efficient search without increasing model complexity. Experimental results on the Figshare Brain Tumor Segmentation dataset demonstrate strong performance, achieving Dice scores of up to 0.9820 under five-fold cross-validation. The optimized model generalizes well to BraTS 2021, reaching whole-tumor Dice scores up to 0.9778. These results demonstrate that lightweight and interpretable metaheuristic optimization can effectively improve segmentation performance while maintaining computational efficiency.

Keywords: Brain tumor segmentation · U-Net · Grey Wolf Optimization · Hyperparameter tuning · MRI.

1 Introduction

Accurate brain tumor segmentation from MRI is essential for diagnosis and treatment planning [11]. Manual annotation is time-consuming and variable, motivating automated deep learning approaches [17]. Encoder–decoder CNNs, particularly U-Net, are widely used due to their ability to capture multiscale context while preserving spatial detail [9].

However, U-Net performance is highly sensitive to hyperparameters such as depth, filter size, kernel size, learning rate, and regularization [8]. In medical imaging, empirical tuning often leads to suboptimal performance. Conventional

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_38

search methods (grid, random, Bayesian) are inefficient or unstable for high-dimensional non-convex problems [12].

Grey Wolf Optimization (GWO) is a lightweight metaheuristic with strong exploration–exploitation balance [10]. Although metaheuristic tuning is effective, most studies focus only on the final performance without analyzing the optimization behavior. This work proposes a lightweight GWO-based hyperparameter tuning framework for U-Net with analysis of convergence, sensitivity, and Top-K robustness.

The main contributions are:

- A lightweight GWO-based hyperparameter optimization for U-Net.
- Systematic analysis of optimization behavior, including convergence, top-performing configurations, and hyperparameter sensitivity.
- Statistical validation of hyperparameter effects using non-parametric tests.
- Improved segmentation performance without increasing model complexity.

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 presents the proposed framework and the experimental setup, Section 4 reports the results, and Section 5 concludes the paper.

2 Related Work

U-Net and its variants are widely used for brain tumor segmentation due to their encoder–decoder design with skip connections [20]. Extensions using residual connections, attention mechanisms, and multiscale convolutions improve feature representation [1], but performance is often influenced by empirically selected hyperparameters.

Hyperparameter optimization is therefore critical. Conventional methods (grid, random, Bayesian) are inefficient for high-dimensional problems [5]. Metaheuristics have shown effectiveness, but most works report only final performance without analyzing convergence, sensitivity, or robustness [14]. GWO provides a lightweight alternative [10], but is typically used as a black-box optimizer. This work addresses these gaps with a lightweight GWO framework and systematic optimization analysis.

3 Methodology

3.1 Parameterized U-Net and Hyperparameter Representation

The segmentation backbone is based on the classical U-Net encoder–decoder architecture with skip connections [20]. Instead of modifying the topology, U-Net is expressed in a parametric form where both the architectural and the training components are optimized (Fig. 1).

The model capacity and the receptive field are jointly influenced by the network depth D and the kernel size K, where increasing D expands the hierarchical

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_38

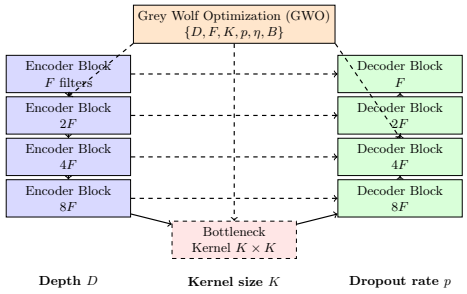


Fig. 1: Parameterized U-Net architecture optimized by GWO.

context and larger K improves local spatial information. Regularization is controlled by dropout p, while training dynamics depends on learning rate η and batch size B. Each candidate configuration is encoded as a vector (Eq. 1).

$$\mathbf{x} = [\eta,\, p,\, F,\, K,\, B,\, D]\,, \tag{1}$$

where each dimension corresponds to a specific hyperparameter, with discrete values obtained via quantization.

Candidate solutions are evaluated using the Dice similarity coefficient (DSC):

$$\mathrm{DSC} = \frac{2|A \cap B|}{|A| + |B|}, \tag{2}$$

and the optimization objective is defined in Eq. 3, where transforms the maximization of DSC into a minimization problem.

$$f(\mathbf{x}) = 1 - \mathrm{DSC}. \tag{3}$$

3.2 GWO Optimization and Training Procedure

Grey Wolf Optimization (GWO) models [10] a leadership hierarchy of α, β, δ, and ω wolves. Each wolf represents a candidate solution, where its position vector $\mathbf{X}_i$ corresponds to the hyperparameter vector defined in Eq. 1. The best solutions are denoted as $\mathbf{X}_\alpha$, $\mathbf{X}_\beta$, and $\mathbf{X}_\delta$. Position updates (Eq. 5) are computed on the basis of the Distance (Eq. 4).

$$\mathbf{D}_i = |\mathbf{C}_i \cdot \mathbf{X}_{\text{leader}} - \mathbf{X}_i|\,, \tag{4}$$

$$\mathbf{X}_i = \mathbf{X}_{\text{leader}} - \mathbf{A}_i \cdot \mathbf{D}_i, \tag{5}$$

where $\mathbf{X}_{\text{leader}} \in \{\mathbf{X}_\alpha, \mathbf{X}_\beta, \mathbf{X}_\delta\}$.

The coefficient vectors are defined in Eq. 6.

$$\mathbf{A}_i = 2a \cdot \mathbf{r}_1 - a, \quad \mathbf{C}_i = 2 \cdot \mathbf{r}_2, \tag{6}$$

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_38

where $\mathbf{r}_1, \mathbf{r}_2 \sim U(0,1)$ and a decreases linearly from 2 to 0 to balance exploration and exploitation.

The final position update is obtained by averaging the leading wolves (Eq. 7).

$$\mathbf{X}(t+1) = \frac{\mathbf{X}_\alpha + \mathbf{X}_\beta + \mathbf{X}_\delta}{3}. \tag{7}$$

To ensure computational efficiency, the optimization phase uses reduced-resolution inputs, short training schedules, and a compact wolf population. Candidate configurations are initialized by uniform sampling within the predefined search bounds, decoded according to Eq. 1, and evaluated using the fitness in Eq. 3. At each iteration, the leading wolves $\mathbf{X}_\alpha$, $\mathbf{X}_\beta$, and $\mathbf{X}_\delta$ guide the population update through Eq. 4–Eq. 7. After optimization, the best configuration $\mathbf{x}_\alpha$ is retrained using full-resolution data. All evaluations are logged to enable analysis of convergence behavior, hyperparameter sensitivity, and robustness.

3.3 Experimental Setup

Experiments are conducted on the Figshare Brain Tumor Segmentation (FBTS) dataset and BraTS 2021. FBTS consists of contrast-enhanced T1-weighted MRI slices for meningioma, glioma, and pituitary tumors with binary masks [7], while BraTS 2021 provides multimodal MRI (FLAIR, T1, T1ce, T2) with annotations for whole tumor (WT), tumor core (TC), and enhancing tumor (ET) [4]. Fixed 80:20 training–validation splits are used.

All images are converted to grayscale, resized, and min–max normalized to $[0,1]$ as $I_{\text{norm}} = \frac{I - \min(I)}{\max(I) - \min(I)}$. During optimization, the inputs are downscaled, while full-resolution images are used for the final training.

U-Net is trained with Adam and binary cross-entropy. Hyperparameters are optimized via GWO under a constrained budget using short training cycles, followed by full-resolution retraining of the best configuration. Random seeds are fixed. The optimized hyperparameters and search ranges are listed in Table 1.

Table 1: Optimized hyperparameters and search ranges.

Hyperparameter	**Symbol**	**Range**	**Hyperparameter**	**Symbol**	**Range**
Learning rate	η	$[10^{-5}, 10^{-2}]$	Kernel size	K	$\{3, 5\}$
Dropout rate	p	$[0.0, 0.5]$	Batch size	B	$\{4, 8, 16\}$
Base filters	F	$\{16, 32, 48, 64\}$	Network depth	D	$\{2, 3, 4\}$

Performance is evaluated using overlap and boundary metrics: Dice (Eq. 2), Jaccard (JI), Hausdorff Distance (HD), and ASSD (Eq. 8).

$$\begin{aligned} \text{JI}(A,B) &= \frac{|A \cap B|}{|A \cup B|}, \quad \text{HD}(A,B) = \max\left\{ \sup_{a \in A} \inf_{b \in B} \|a - b\|, \sup_{b \in B} \inf_{a \in A} \|b - a\| \right\}, \\ \text{ASSD}(A,B) &= \frac{1}{|A| + |B|} \left(\sum_{a \in A} \inf_{b \in B} \|a - b\| + \sum_{b \in B} \inf_{a \in A} \|b - a\| \right), \end{aligned} \tag{8}$$

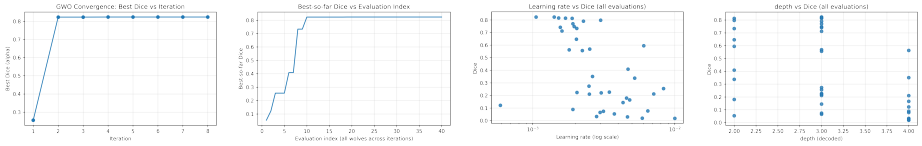
ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_38

where A and B denote boundary points of predicted and ground-truth segmentations.

4 Results

4.1 Optimization Behavior and Hyperparameter Analysis

Fig. 2 summarizes both convergence behavior and hyperparameter sensitivity. The best Dice improves rapidly from ≈ 0.26 to > 0.82 in two iterations and stabilizes thereafter, indicating efficient early convergence. The best-so-far curve exceeds 0.80 within ≈ 10 evaluations, confirming that the lightweight search budget is sufficient to identify a high-quality configuration.



(a) Best Dice (b) Best-so-far (c) Learning rate (d) Depth

Fig. 2: Optimization convergence and hyperparameter sensitivity of the proposed GWO framework.

Sensitivity analysis shows that the learning rate is the dominant factor, with optimal values around 10^{-3}. Moderate depth ($D = 3$) provides stable performance, while extreme configurations lead to degraded accuracy. These trends are consistent with the convergence dynamics. Statistical analysis across all evaluated configurations (N = 40) confirms that the learning rate has a significant effect ($p < 0.01$, $\varepsilon^2 > 0.14$), while depth and filter size exhibit moderate influence. Top-performing configurations (Dice > 0.81) are concentrated around learning rates of $(1\text{–}2) \times 10^{-3}$, moderate dropout ($\approx 0.1$–$0.2$), and depth $D = 3$, indicating a stable high-performing region rather than a single isolated optimum.

4.2 Quantitative Results and Generalization

Table 2 summarizes performance on FBTS and BraTS 2021. On FBTS, the model achieves high accuracy across tumor types, with Dice up to 0.9820 for meningioma and 0.9568 for pituitary, while glioma improves significantly after cross-validation (0.9189). On BraTS 2021, strong generalization is observed, with a whole tumor Dice up to 0.9778 and consistent performance across modalities.

4.3 Comparison with State-of-the-Art

Table 3 compares the proposed method with representative approaches. On FBTS, the proposed method achieves DSC = 0.9820, significantly exceeding typical U-Net variants. On BraTS 2021, it achieves DSC = 0.9778, outperforming

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_38

Table 2: Segmentation performance on FBTS and BraTS 2021.

FBTS					BraTS 2021						
Class	DSC_{val}	JI_{val}	DSC_{cv}	JI_{cv}	Modality	DSC_{WT}	JI_{WT}	DSC_{TC}	JI_{TC}	DSC_{ET}	JI_{ET}
Meningioma	0.9338	0.8765	0.9820	0.9652	FLAIR	0.9778	0.9566	0.9087	0.8341	0.9395	0.8862
Glioma	0.7532	0.6110	0.9189	0.8663	T1	0.8904	0.8044	0.9608	0.9245	0.9454	0.8966
Pituitary	0.8779	0.7849	0.9568	0.9207	T2	0.9764	0.9539	0.9520	0.9081	0.9476	0.9011
					T1CE	0.8983	0.8196	0.9729	0.9498	0.9658	0.9340

several recent methods. Compared to NAS/AutoML approaches, which require extensive computational resources, the proposed lightweight GWO framework achieves competitive performance at significantly lower cost.

Table 3: Comparison with state-of-the-art methods on FBTS and BraTS 2021 using DSC.

FBTS		BraTS 2021 (WT)	
Method	DSC	Method	DSC
Ensemble-CNN. [3]	0.748	DenseUNet+ [6]	0.9500
FCNN (U-Net) [2]	0.753	UNet-AG [16]	0.9483
U-Attention Net [19]	0.704	ResUNet50 [15]	0.9553
MS-UNet [18]	0.800	PSO-UNet [13]	0.9578
Proposed GWO-UNet	**0.9820**	**Proposed GWO-UNet**	**0.9778**

4.4 Qualitative Results

Fig. 3 shows representative segmentation results, demonstrating strong boundary alignment and high overlap with ground truth, consistent with the quantitative Dice and Jaccard scores. Boundary discrepancies are minimal and primarily localized at ambiguous tumor regions, aligning with low HD and ASSD values.

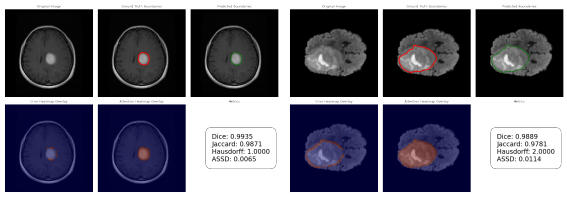


Fig. 3: Representative qualitative segmentation results on FBTS (Meningioma) and BraTS 2021 (FLAIR).

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_38

5 Conclusion

This paper presented a lightweight Grey Wolf Optimization (GWO)–based framework for hyperparameter tuning of U-Net for brain tumor segmentation. By optimizing key architectural and training parameters under a constrained budget, the method efficiently identified stable high-performing configurations without increasing model complexity. The optimized model achieved strong performance on FBTS (up to DSC = 0.9820) and demonstrated robust generalization on BraTS 2021 (up to DSC = 0.9778). The convergence, sensitivity, and statistical analyses further confirmed the effectiveness and interpretability of the optimization process. Overall, the results show that lightweight metaheuristic optimization can provide accurate and computationally efficient segmentation.

Acknowledgement. This work was supported by ACK Cyfronet AGH (Grant no. PLG/2025/018784), and Polish Ministry of Science and Higher Education funds assigned to AGH University of Krakow.

References

1. Abueed, O., Wang, Y., Khasawneh, M.: A Systematic Review of U-Net Optimizations: Advancing Tumour Segmentation in Medical Imaging. IET Image Processing **19**(1) (jan 2025). https://doi.org/10.1049/ipr2.70203
2. Agravat, R.R., Raval, M.S.: 3D Semantic Segmentation of Brain Tumor for Overall Survival Prediction. In: Crimi, A., Bakas, S. (eds.) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2020. Lecture Notes in Computer Science(), vol. 12659, pp. 215–227. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-72087-2_19
3. Ali, M.J., et al.: Glioma Segmentation Using Ensemble of 2D/3D U-Nets and Survival Prediction Using Multiple Features Fusion. In: Crimi, A., Bakas, S. (eds.) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2020. Lecture Notes in Computer Science(), vol. 12659, pp. 189–199. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-72087-2_17
4. Baid, U., Ghodasara, S., Mohan, S., others.: RSNA-ASNR-MICCAI-BraTS-2021 Dataset (2023). https://doi.org/10.7937/jc8x-9874
5. Bian, K., Priyadarshi, R.: Machine Learning Optimization Techniques: A Survey, Classification, Challenges, and Future Research Issues. Archives of Computational Methods in Engineering **31**, 4209–4233 (mar 2024). https://doi.org/10.1007/s11831-024-10110-w
6. Çetiner, H., Metlek, S.: DenseUNet+: A novel hybrid segmentation approach based on multi-modality images for brain tumor segmentation. Journal of King Saud University - Computer and Information Sciences **35**(8), 101663 (2023). https://doi.org/10.1016/j.jksuci.2023.101663
7. 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 **10**(10), e0140381 (2015). https://doi.org/10.1371/journal.pone.0140381

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_38

8. 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 **81**(24), 34717–34735 (oct 2022). https://doi.org/10.1007/s11042-021-11273-5
9. 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 **19**(1) (jan 2025). https://doi.org/10.1049/ipr2.70019
10. Liu, Y., As'arry, A., Hassan, M.K., Hairuddin, A.A., Mohamad, H.: Review of the grey wolf optimization algorithm: variants and applications. Neural Computing and Applications **36**(6), 2713–2735 (feb 2024). https://doi.org/10.1007/s00521-023-09202-8
11. 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 **32**(3), 1525–1569 (apr 2025). https://doi.org/10.1007/s11831-024-10188-2
12. Saifullah, S., Drężewski, R.: Bayesian Optimization-Driven U-Net Architecture Tuning for Brain Tumor Segmentation. In: Proceedings of the 6th International Electronic Conference on Applied Sciences. MDPI, Basel, Switzerland (2025). https://doi.org/10.3390/engproc2026124022
13. 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 Companion, pp. 323–326. ACM, New York, NY, USA (2025). https://doi.org/10.1145/3712255.3726561
14. 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 **16**(6), 456 (may 2025). https://doi.org/10.3390/info16060456
15. 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 **13**, 43473–43489 (2025). https://doi.org/10.1109/ACCESS.2025.3547430
16. 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 **37**, 5521–5558 (2025). https://doi.org/10.1007/s00521-024-10919-3
17. Shaheema, B., Muppalaneni, N.B.: Multimodal brain image segmentation: a recent review, challenges and future perspectives. Multimedia Tools and Applications **84**(39), 48023–48070 (aug 2025). https://doi.org/10.1007/s11042-025-21068-7
18. Soltaninejad, M., Pridmore, T., Pound, M.: Efficient MRI Brain Tumor Segmentation Using Multi-resolution Encoder-Decoder Networks. In: Crimi, A., Bakas, S. (eds.) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2020. Lecture Notes in Computer Science(), vol. 12659, pp. 30–39. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-72087-2_3
19. Xu, J.H., Teng, W.P.K., Wang, X.J., Nürnberger, A.: A Deep Supervised U-Attention Net for Pixel-Wise Brain Tumor Segmentation. In: Crimi, A., Bakas, S. (eds.) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2020. Lecture Notes in Computer Science(), vol. 12659, pp. 278–289. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-72087-2_24
20. Yang, W., Zhang, R., Chow, S.K.K., Li, Z.: A survey of U-Net variant network for MRI brain tumor segmentation. Discover Artificial Intelligence (dec 2025). https://doi.org/10.1007/s44163-025-00525-0

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_38