

Optimizing U-Net Architecture Using Differential Evolution for Brain Tumor Segmentation

Shoffan Saifullah^{1,2}[0000–0001–6799–3834] and Rafal
Dreżewski¹[0000–0001–8607–3478]

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

Abstract. Accurate brain tumor segmentation is essential for effective diagnosis and treatment planning. This study proposes DE-UNet, an enhanced U-Net architecture optimized using Differential Evolution (DE) to improve segmentation of multimodal MRI scans. The model was evaluated on two benchmark datasets: Figshare Brain Tumor Segmentation (FBTS) and BraTS 2021 datasets, focusing on whole tumor segmentation across four MRI modalities: FLAIR, T1, T1-CE, and T2. DE-UNet outperformed state-of-the-art methods, achieving Dice Similarity Coefficient (DSC) and Jaccard Index (JI) scores of 0.9160/0.8472 on FBTS and 0.9094/0.8371 on BraTS 2021. DE effectively optimized key hyperparameters—learning rate, dropout, batch size, and filter sizes—enhancing the model generalization across tumor types and imaging conditions. Visual analysis confirmed accurate tumor boundary delineation. These results highlight the potential of DE-UNet as a robust and precise tool for clinical brain tumor segmentation.

Keywords: Brain Tumor Segmentation · Differential Evolution · MRI Modalities · U-Net Optimization · Medical Image Analysis.

1 Introduction

Accurate brain tumor segmentation is crucial for diagnosis, treatment planning, and prognosis [15]. MRI is widely used due to its superior soft tissue contrast and non-invasive nature [3], but manual segmentation is time-consuming, subjective, and inconsistent [9]. Tumor heterogeneity further complicates the task, emphasizing the need for reliable automated methods [7].

U-Net and its variants have shown strong performance in medical image segmentation [19, 21], but limitations persist in hyperparameter tuning, handling class imbalance, and adapting to multimodal MRI. Recent models like DeepLabV3+[23] and U-Net extensions[22] improve performance but remain suboptimal for multi-class and multi-modal cases. Metaheuristics such as PSO [20] and GA [10] aid tuning but often converge prematurely [6], whereas Differential Evolution (DE) offers more robust and adaptive optimization [12, 16].

This study introduces DE-UNet, a U-Net architecture optimized via DE to improve segmentation of multimodal MRI brain tumor images. Evaluated on the FBTS and BraTS 2021 datasets, it targets whole tumor segmentation across classes and modalities, achieving superior Dice Similarity Coefficient (DSC) and Jaccard Index (JI) compared to state-of-the-art methods. Section 2 details the proposed framework and datasets, Section 3 presents results and comparisons, and Section 4 concludes the study.

2 Methods

The proposed DE-UNet was evaluated on two benchmark datasets (Fig. 1): the Figshare Brain Tumor Segmentation (FBTS) [5] and BraTS 2021[2], selected for their diversity in tumor types and MRI modalities. FBTS includes 3064 slices labeled as Meningioma (708), Glioma (1426), and Pituitary (930), each with expert-annotated binary masks. BraTS 2021 comprises 1251 multimodal slices across T1, T1-CE, T2, and FLAIR, with corresponding whole tumor masks.

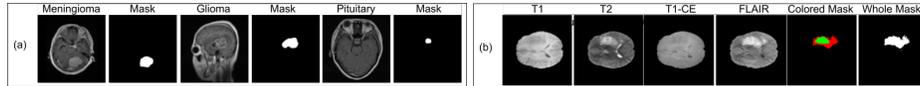


Fig. 1: Sample images from the datasets: (a) FBTS and (b) BraTS 2021.

All images were resized to 256×256 pixels using bicubic interpolation [21], and intensities were normalized to $[0, 1]$ to reduce scanner-induced variability. Segmentation masks were binarized to separate tumor from background. This standardized preprocessing enhanced contrast, reduced noise, and ensured consistent inputs for training and evaluation.

DE-UNet integrates a U-Net backbone with Differential Evolution (DE) to optimize four hyperparameters: learning rate, dropout rate, batch size, and number of filters. The architecture follows an encoder-decoder structure with skip connections and bottleneck dropout [20], as shown in Fig. 2.

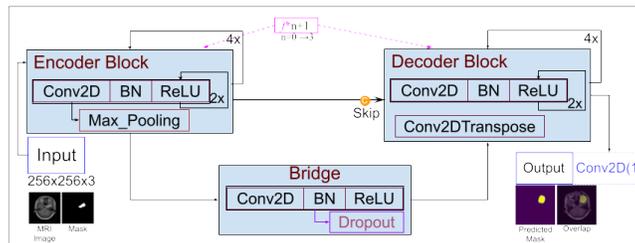


Fig. 2: DE-UNet: Encoder-decoder architecture with skip connections.

Algorithm 1 DE Algorithm for U-Net hyperparameter optimization.**Require:** NP, F, CR, G **Ensure:** Optimized hyperparameter configuration x^* 1: Initialize a population of NP candidate solutions within predefined bounds.2: **for** generation $g = 1$ to G **do**3: **for** each candidate solution x_i in the population **do**4: **Mutation:** Randomly select 3 distinct solutions (x_a, x_b, x_c) from the population.5: Generate mutant vector: $v_i = x_a + F \cdot (x_b - x_c)$ 6: **Crossover:** Generate trial vector u_i as:

$$u_{ij} = \begin{cases} v_{ij} & \text{if } \text{rand}(0, 1) < CR \text{ or } j = j_{\text{rand}} \\ x_{ij} & \text{otherwise} \end{cases}$$

7: **Selection:** Replace x_i with u_i if $f(u_i) < f(x_i)$ 8: **end for**9: **end for**10: Return the best-performing solution x^*

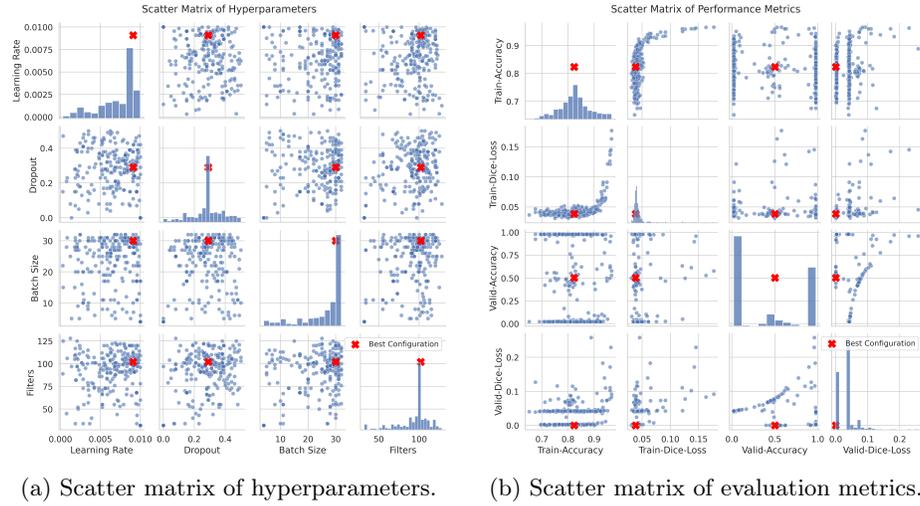
DE was configured with a population size of $NP = 5$, mutation factor $F = 0.8$, crossover probability $CR = 0.9$, and a maximum of $G = 10$ generations. Continuous parameters (learning rate: $[10^{-5}, 10^{-2}]$, dropout: $[0.1, 0.5]$) were mutated within bounds, while discrete parameters (batch size: 8, 16, 32, 64; filters: 16, 32, 64, 128) were rounded post-mutation. The hyperparameter ranges were selected based on prior studies [4, 20] and validated empirically. Optimization was performed separately on the FBTS and BraTS 2021 datasets.

The search minimized a composite loss: $\mathcal{L}_{\text{total}} = 1 - \alpha \cdot \text{DSC} - (1 - \alpha) \cdot \text{JI}$, where $\alpha = 0.5$ balances overlap accuracy (Dice Similarity Coefficient, DSC) and boundary agreement (Jaccard Index, JI). This objective encourages robust segmentation across tumor types and modalities. The DE operations—mutation, crossover, and selection—are detailed in Algorithm 1, which iteratively returns the best-performing configuration x^* for final training and evaluation.

DE-UNet was evaluated on the FBTS and BraTS 2021 datasets using a server with 8 NVIDIA A100-SXM4-40GB GPUs. All images and masks were resized to 256×256 pixels, converted to three-channel format, and normalized to $[0, 1]$. An 80/20 training-validation split was applied. During the DE search, each candidate configuration was trained for 10 epochs to evaluate validation performance.

The model was trained using the Adam optimizer and binary cross-entropy loss, with DE optimizing the learning rate, dropout rate, batch size, and filter size. After selecting the best configuration, final training and evaluation were performed on both datasets.

Model performance was assessed using Accuracy, Dice Similarity Coefficient (DSC), and Jaccard Index (JI) [19, 20, 18]. Higher DSC and JI indicate stronger overlap between predicted and ground truth masks. Evaluation was performed per tumor class in FBTS (Meningioma, Glioma, Pituitary) and per modality in BraTS 2021 (FLAIR, T1, T1-CE, T2) to assess robustness and generalization.



(a) Scatter matrix of hyperparameters. (b) Scatter matrix of evaluation metrics.

Fig. 3: Effect of DE-tuned hyperparameters on model performance (red: best).

3 Results and Discussion

Differential Evolution (DE) was applied to optimize four U-Net hyperparameters: learning rate, dropout rate, batch size, and initial filter count. Performance was assessed using training and validation accuracy and Dice-based loss.

Fig. 3 visualizes the influence of these hyperparameters on the performance. The hyperparameter matrix (Fig. 3a) shows that learning rate and dropout were the most sensitive, directly impacting the convergence. Batch size and filter count affected the training stability and efficiency. The performance matrix (Fig. 3b) highlights how small hyperparameter shifts can yield notable gains in validation accuracy and Dice scores.

The optimal configuration—learning rate 0.009094, dropout 0.286, batch size 30, and 102 filters—achieved the training accuracy of 0.9302, the validation accuracy of 0.9788, and the validation Dice loss of 0.0024. Among all parameters, the learning rate and the dropout had the strongest impact on generalization, while batch size and filter count mainly influenced the training dynamics.

The DE-optimized U-Net was evaluated on the FBTS and BraTS 2021 datasets to assess segmentation performance across diverse tumor types and MRI modalities. Table 1 presents the results on the FBTS dataset. The model achieved consistently high accuracy, Dice Similarity Coefficient (DSC), and Jaccard Index (JI) across all tumor classes. Meningioma achieved the best DSC (0.9348), while Glioma—despite its irregular morphology—maintained strong performance. Pituitary tumors exhibited stable, near-perfect accuracy and overlap.

Table 2 summarizes the performance on BraTS 2021 before and after extended training. Improvements are shown as percentage point (pp) gains. T1 and T2 exhibited the largest relative gains, particularly in DSC and JI. T1-CE

Table 1: Performance metrics on the FBTS dataset.

Tumor Type	Training				Validation			
	Accuracy	Loss	DSC	JI	Accuracy	Loss	DSC	JI
Meningioma	0.9983	0.0042	0.9286	0.8677	0.9984	0.0038	0.9348	0.8784
Glioma	0.9971	0.0070	0.9023	0.8231	0.9968	0.0080	0.8943	0.8103
Pituitary	0.9991	0.0022	0.9183	0.8509	0.9991	0.0021	0.9200	0.8539

Table 2: Performance on BraTS 2021 before and after extended training.

Modality	Training				Validation			
	Accuracy (pp)	Loss (pp)	DSC (pp)	JI (pp)	Accuracy (pp)	Loss (pp)	DSC (pp)	JI (pp)
FLAIR	0.9956 (+0.18)	0.0110 (-0.47)	0.8941 (+4.31)	0.8103 (+7.21)	0.9961 (+0.15)	0.0095 (-0.37)	0.9068 (+3.31)	0.8304 (+5.69)
T1	0.9935 (+0.32)	0.0157 (-0.79)	0.8464 (+7.58)	0.7353 (+12.12)	0.9930 (+0.39)	0.0168 (-0.96)	0.8327 (+9.07)	0.7154 (+14.3)
T1-CE	0.9950 (+0.34)	0.0119 (-0.82)	0.8823 (+8.05)	0.7900 (+13.86)	0.9940 (+0.43)	0.0147 (-1.06)	0.8576 (+10.37)	0.7524 (+17.34)
T2	0.9946 (+0.28)	0.0135 (-0.72)	0.8707 (+6.65)	0.7714 (+11.1)	0.9942 (+0.34)	0.0147 (-0.89)	0.8602 (+7.97)	0.7550 (+13.23)

achieved the highest overall segmentation accuracy (DSC: 0.9613, JI: 0.9258), highlighting the benefit of contrast-enhanced imaging. These results validate DE-UNet’s effectiveness across MRI modalities and tumor structures. These findings are consistent with prior studies showing that modality contrast significantly influences segmentation quality [20].

The DE-optimized U-Net was evaluated on the FBTS and BraTS 2021 test sets using Accuracy, Loss, Dice Similarity Coefficient (DSC), and Jaccard Index (JI). Table 3 summarizes the model’s final test performance. On the FBTS dataset, Meningioma achieved the highest DSC (0.9410) and JI (0.8895), while Pituitary showed the best accuracy (0.9991). Glioma remained more challenging due to its irregular morphology, resulting in the lower DSC (0.8922).

Table 3: Test performance of the DE-optimized U-Net model.

Class	Accuracy	Loss	DSC	JI	Modality	Accuracy	Loss	DSC	JI
Meningioma	0.9985	0.0035	0.9410	0.8895	FLAIR	0.9976	0.0058	0.9447	0.8956
Glioma	0.9966	0.0083	0.8922	0.8069	T1	0.9969	0.0073	0.9277	0.8657
Pituitary	0.9991	0.0020	0.9148	0.8451	T1-CE	0.9983	0.0041	0.9634	0.9297
					T2	0.9976	0.0058	0.9399	0.8873

For BraTS 2021, the T1-CE modality achieved the highest segmentation performance (DSC: 0.9634, JI: 0.9297), benefiting from enhanced tumor contrast. FLAIR and T2 also performed well in delineating edema regions. T1 showed comparatively lower overlap metrics but remained effective in less complex cases.

Fig. 4 presents qualitative segmentation results, comparing ground truth (red contours) with predicted masks (green contours). The model produced high-quality segmentations across all tumor types and modalities, with T1-CE and FLAIR showing the closest alignment. Despite Glioma complexity, prediction masks demonstrated strong overlap with expert annotations. These results confirm the model’s robustness in handling modality-specific and tumor-type variations. DE-UNet adapts well to anatomical complexity, achieving reliable segmentations suitable for clinical diagnostic support.

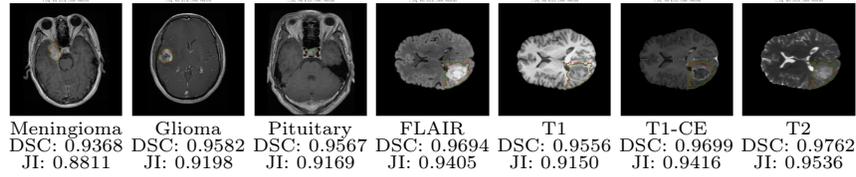


Fig. 4: Qualitative results of DE-UNet (red: ground truth, green: prediction).

Table 4: Comparison of DE-UNet with state-of-the-art models.

Method	FBTS Dataset		Method	BraTS 2021	
	DSC	JI		DSC	JI
Proposed DE-UNet	0.9160	0.8472	Proposed DE-UNet	0.9094	0.8371
DeepLabV3+Xception [23]	0.8115	0.8018	UNet [8]	0.8600	0.7807
KFCM-CNN [14]	0.8884	0.8204	U-Net base [25]	0.9080	-
U-Net based [1]	0.8900	0.8100	SPPNet-2 [25]	0.9040	-
MST-based [13]	0.8469	0.7443	UNCE-NODE [17]	0.8949	-
U-Net with ResNet [11]	0.9011	-	nnU-Net [24]	0.8900	-

Table 4 compares the proposed DE-UNet against state-of-the-art (SOTA) methods on the FBTS and BraTS 2021 datasets. DE-UNet outperformed all baselines across Dice Similarity Coefficient (DSC) and Jaccard Index (JI), demonstrating superior segmentation accuracy.

On the FBTS dataset, DE-UNet achieved the DSC of 0.9160 and the JI of 0.8472, outperforming the previous best model, U-Net with ResNet [11] (DSC: 0.9011). For BraTS 2021, DE-UNet scored 0.9094 in DSC and 0.8371 in JI, exceeding prior models such as UNet [8] and modular approaches such as UNCE-NODE [17]. These improvements reflect the impact of differential evolution (DE) in dynamically tuning learning rate, dropout, batch size, and filter count, unlike traditional static or manually configured models.

DE-UNet’s modality-agnostic architecture and deeper encoder-decoder design contribute to its robust performance across MRI modalities (FLAIR, T1, T1-CE, T2). The model captures fine tumor boundaries without requiring modality-specific preprocessing or architectural modifications. These strengths enhance generalization to diverse tumor characteristics and support clinical applicability, positioning DE-UNet as a competitive and practical solution for brain tumor segmentation.

4 Conclusions

This study proposed DE-UNet, a U-Net architecture enhanced with Differential Evolution (DE) for optimized brain tumor segmentation. By automatically tuning key hyperparameters—learning rate, dropout, batch size, and filters—DE-UNet achieved superior performance on the FBTS and BraTS 2021 datasets, with DSC of 0.9160 and 0.9094, respectively. The model demonstrated strong generalization across MRI modalities (FLAIR, T1, T1-CE, T2) and tumor types,

outperforming several state-of-the-art methods. Visual results confirmed accurate tumor boundary alignment with expert annotations. These findings highlight the effectiveness of DE-driven optimization in medical image segmentation. Future work will explore hybrid metaheuristics and validation on larger, multi-modal datasets to enhance clinical applicability.

Acknowledgement. Research funding was provided by AGH University of Krakow (Program “Excellence initiative – research university”), ACK Cyfronet AGH (Grant no. PLG/2024/017503), and Polish Ministry of Science and Higher Education funds assigned to AGH University of Krakow.

References

1. Akter, A., et al.: Robust clinical applicable CNN and U-Net based algorithm for MRI classification and segmentation for brain tumor. *Expert Systems with Applications* **238**, 122347 (2024). <https://doi.org/10.1016/j.eswa.2023.122347>
2. Baid, U., et al.: RSNA-ASNR-MICCAI-BraTS-2021 Dataset (2023). <https://doi.org/10.7937/jc8x-9874>
3. 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* p. 108412 (2024). <https://doi.org/10.1016/j.compbimed.2024.108412>
4. Biswas, S., et al.: Improving differential evolution through bayesian hyperparameter optimization. In: 2021 IEEE Congress on Evolutionary Computation (CEC). pp. 832–840. IEEE (2021). <https://doi.org/10.1109/CEC45853.2021.9504792>
5. Cheng, J.: brain tumor dataset. Figshare (2017). <https://doi.org/10.6084/m9.figshare.1512427.v5>
6. Gad, A.G.: Particle swarm optimization algorithm and its applications: A systematic review. *Archives of Computational Methods in Engineering* **29**(5), 2531–2561 (2022). <https://doi.org/10.1007/s11831-021-09694-4>
7. Ghadimi, D.J., 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>
8. Hernandez-Gutierrez, F.D., et al.: Brain tumor segmentation from optimal MRI slices using a lightweight U-Net. *Technologies* **12**(10), 183 (2024). <https://doi.org/10.3390/technologies12100183>
9. Kaifi, R.: A review of recent advances in brain tumor diagnosis based on AI-based classification. *Diagnostics* **13**(18), 3007 (2023). <https://doi.org/10.3390/diagnostics13183007>
10. Khouy, M., Jabrane, Y., Ameer, M., Hajjam El Hassani, A.: Medical image segmentation using automatic optimized U-Net architecture based on genetic algorithm. *Journal of Personalized Medicine* **13**(9), 1298 (aug 2023). <https://doi.org/10.3390/jpm13091298>
11. Kumar Sahoo, A., Parida, P., Muralibabu, K., Dash, S.: Efficient simultaneous segmentation and classification of brain tumors from MRI scans using deep learning. *Biocybernetics and Biomedical Engineering* **43**(3), 616–633 (2023). <https://doi.org/10.1016/j.bbe.2023.08.003>

12. Kuş, Z., et al.: Differential evolution-based neural architecture search for brain vessel segmentation. *Engineering Science and Technology, an International Journal* **46**, 101502 (2023). <https://doi.org/10.1016/j.jestch.2023.101502>
13. Mayala, S., et al.: Brain tumor segmentation based on minimum spanning tree. *Frontiers in Signal Processing* **2** (2022). <https://doi.org/10.3389/frsip.2022.816186>
14. Rao, S.K.V., Lingappa, B.: Image analysis for MRI based brain tumour detection using hybrid segmentation and deep learning classification technique. *International Journal of Intelligent Engineering and Systems* **12**(5), 53–62 (2019). <https://doi.org/10.22266/ijies2019.1031.06>
15. 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* (oct 2024). <https://doi.org/10.1007/s11831-024-10188-2>
16. Ren, L., et al.: Multi-level thresholding segmentation for pathological images: Optimal performance design of a new modified differential evolution. *Computers in Biology and Medicine* **148**, 105910 (2022). <https://doi.org/10.1016/j.compbimed.2022.105910>
17. Sadique, M.S., et al.: Brain tumor segmentation using neural ordinary differential equations with UNet-Context encoding network. In: Bakas, S., et al. (eds.) *Brain-lesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2022*, pp. 205–215. LNCS, Springer, Cham (2023). https://doi.org/10.1007/978-3-031-33842-7_18
18. Saifullah, S., Dreżewski, R.: Brain tumor segmentation using ensemble CNN-transfer learning models: DeepLabV3plus and ResNet50 approach. In: Franco, L., et al. (eds.) *Computational Science – ICCS 2024*, LNCS, vol. 14835, pp. 340–354. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-63772-8_30
19. Saifullah, S., Dreżewski, R.: Improved brain tumor segmentation using modified U-Net based on Particle Swarm Optimization image enhancement. In: *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. pp. 611–614. GECCO '24 Companion, Association for Computing Machinery, New York, NY, USA (2024). <https://doi.org/10.1145/3638530.3654339>
20. Saifullah, S., et al.: Automatic brain tumor segmentation using convolutional neural networks: U-Net framework with PSO-tuned hyperparameters. In: Affenzeller, M., et al. (eds.) *Parallel Problem Solving from Nature – PPSN XVIII*, LNCS, vol. 15150, pp. 333–351. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-70071-2_21
21. Saifullah, S., et al.: Modified U-Net with attention gate for enhanced automated brain tumor segmentation. *Neural Computing and Applications* (2024). <https://doi.org/10.1007/s00521-024-10919-3>
22. Saifullah, S., et al.: Optimizing brain tumor segmentation through CNN U-Net with CLAHE-HE image enhancement. *Proceedings of the 2023 1st International Conference on Advanced Informatics and Intelligent Information Systems (ICAI3S 2023)* pp. 90–101 (2024). https://doi.org/10.2991/978-94-6463-366-5_9
23. Saifullah, S., et al.: Advanced brain tumor segmentation using DeepLabV3Plus with Xception encoder on a multi-class MR image dataset. *Multimedia Tools and Applications* (feb 2025). <https://doi.org/10.1007/s11042-025-20702-8>
24. Sørensen, P.J., et al.: Repurposing the public BraTS dataset for postoperative brain tumour treatment response monitoring. *Tomography* **10**(9), 1397–1410 (sep 2024). <https://doi.org/10.3390/tomography10090105>
25. Vijay, S., et al.: MRI brain tumor segmentation using residual spatial pyramid pooling-powered 3D U-Net. *Frontiers in Public Health* **11** (2023). <https://doi.org/10.3389/fpubh.2023.1091850>