

Comparative Analysis of Hybrid Attention and Progressive Layering through a Comprehensive Evaluation of ARU-Net and PLU-Net in Brain Tumour Segmentation

Noman Ahmed Siddiqui^{1*}, Muhammad Tahir Qadri¹, Muhammad Ovais Akhter², Tahira Anwar³, Fahad Farooq¹

¹Electronic Engineering Department, Sir Syed University of Engineering and Technology, Karachi, Pakistan-75290

²Electrical Engineering Department, Bahria University, Karachi, Pakistan-75350

³Department of Languages and Communication, University of Jyväskylä, Finland.

Abstract:

Accurate segmentation of brain tumors using multimodal Magnetic Resonance Imaging (MRI) is critical for advancing automated medical imaging techniques. This study presents a comprehensive comparative analysis of two advanced U-Net-based architectures, Neural Hybrid Channel Structure Regression U-Net (ARU-Net) and Progressive Layered U-Net (PLU-Net), to address the research question: how do attention-driven and progressive layering approaches impact segmentation performance in terms of accuracy, boundary delineation, and computational efficiency? ARU-Net integrates Convolutional Block Attention Modules (CBAM) with U-Net to enhance feature rectification and adaptability to heterogeneous tumor regions, producing continuous probability maps. In contrast, PLU-Net employs a cascaded, multi-stage structure with attention gates and multi-scale supervision to ensure high boundary accuracy across resolutions. Using the BraTS 2021 dataset (T1, T2, FLAIR modalities), we compare their preprocessing pipelines, segmentation accuracy, and computational demands. ARU-Net achieves higher specificity (0.949) and sensitivity (0.923), excelling in healthy tissue differentiation, while PLU-Net demonstrates superior boundary accuracy (Hausdorff95: 2.5) through iterative refinement. These findings provide insights into the trade-offs between rapid adaptability and multi-resolution precision, guiding model selection for clinical and research applications.

Keywords: Brain Tumour Segmentation, Deep Learning, Magnetic Resonance Imaging, Convolutional Neural Networks, Hybrid Attention, Progressive Learning Model.

I. INTRODUCTION

Tumour diagnosis and treatment planning of the brain is one of the greatest tasks of modern neuro-oncology because the accurate delineation of tumour fringe is the only key to the primary optimization of the therapeutic approach and subsequent increase in patient survival [1]. The imaging modality of preference to assess the brain tumour is the Magnetic Resonance Imaging (MRI) that gives a multimodal assessment such as T1-weighted (T1w), T1-contrast-enhanced (T1ce), T2-weighted (T2w) and fluid-attenuated inversion recovery (FLAIR) MR. However, the process of manual segmentation of brain tumours on an MRI scan is both time-intensive, highly subjective, and observer-dependent; hence, why manual segmentation methods have driven the need to find automated techniques of segmenting brain tumours on an MRI scan [2]. Brain tumour segmentation using multimodal Magnetic Resonance Imaging (MRI) is a critical task in medical imaging, enabling precise diagnosis and treatment planning. U-Net-based architectures have become a cornerstone for automated segmentation due to their ability to capture spatial and contextual features. However, challenges such as heterogeneous tumour morphologies and varying MRI modalities necessitate advanced models to improve accuracy and robustness. This study compares two enhanced U-Net variants: Neural Hybrid Channel Structure Regression U-Net (ARU-Net) and Progressive Layered U-Net (PLU-Net). ARU-Net integrates attention

mechanisms, such as Convolutional Block Attention Modules (CBAM), to refine feature extraction, while PLU-Net employs a cascaded, multi-scale architecture with attention gates for progressive refinement.

The current phenomenon of deep learning frameworks, in general, and convolutional neural networks (CNNs), in particular, has radically transformed the area of medical image analysis [3]. Among the CNN designs, the U-Net design has become a landmark study, characterized by its encoder-decoder framework that combines the contextual information provided by the representations with accurate and small-scale spatial information [4]. In as much as U-Net has worked successfully in many segmentation tasks, it fails to capture the small changes peculiar to the tumours of the brain and also to handle idiosyncratic intensity distributions across MRI modalities. In two recent articles, new extensions to the U-Net architecture are proposed: the Novel Hybrid Channel Attention Regression UNet (ARU-Net) [5] and the High-Precision Brain Tumour Segmentation using a Progressive Layered U-Net (PLU-Net) [6]. Both deal with certain challenges related to brain tumour segmentation, but they apply conceptually different solutions. The present discussion examines the key characteristics of the ARU-Net and PLU-Net as well as prepares the ground to compare systematically the contributions of these two research works to the segmentation of brain tumours using MRI. An End-to-end Semantic Brain Tumour Segmentation Approach on Multimodal MRI. Model: The model is a variant of U-Net, because it augments the output decoder with attention modules on the channels to leverage contextual data to implement pixel-wise classification. The channel attention mechanism is used as a two-layer regression head to adaptively adjust feature maps in each layer in decoder. Empirical findings point to that ARU-Net can provide better segmentation results in all modalities tested compared to the baseline U-Net (both under Dice and Hausdorff distance measures). In this case, the U-Net architecture is redesigned to contain three nested decoders that are used to refine predictions of the segmentation at higher spatial levels. The inner decoder works at full resolution, the middle one is done on half-resolution results and the outer decoder is done on one quarter-resolution inputs discarding contextual information without affecting resolution. Such a progressive approach to decoding enhances segmentation accuracy using a perturbative analysis of errors in segmentation [7]. The authors also observe that the method effectively generalizes into a wide range of the types of tumours including high-grade gliomas, meningiomas, and large vessels [8]. Overall, ARU-Net and PLU-Net are modifications of the U-Net architecture that are considered innovative as each of them is aimed at resolving the unique problems in brain tumour segmentation using MRI data. ARU-Net extends the discrimination at pixel-level using channel attention modules [9] added at the decoder stage. Instead, PLU-Net uses three nested decoders which operate on progressively coarser input, maintaining context at each level of prediction. This comparative discussion provides the basis of a stringent criticism of these methods within the framework of the MRI-based segmentation of brain tumours.

The ARU-Net framework, as elucidated in the publication "Hybrid Channel Attention Regression U-Net (ARU-Net): An Enhanced Architecture for Brain Tumour Segmentation in Magnetic Resonance Imaging," was devised to elevate feature extraction through sophisticated attention mechanisms. Through a basic expansion of the conventional U-Net framework, ARU-Net comprises the incorporation of an element of spatial attention into the basic U-Net structure, which is borrowed off the Convolutional Block Attention Module (CBAM) framework [10], as well as an inclusion of channel-wise attention, borrowed off the Squeeze-and-Excitation (SE) framework [11]. Spatial attention component highlights important areas of MRI scans and emphasizes less meaningful or noisy components hence enhancing the focus of the model to areas that are important to the tumour. At the same time, the relative importance of feature channels is channel-wise reweighted through attention, so the network focuses on the most relevant ones in terms of tumour delineation. An innovative non-repetitive product of ARU-Net is that the standard classification output is replaced by a regression output layer, a product of which is the continuous probability maps. This change allows exerting finer control on vague or overlapping tumour areas, providing a detailed counterpoint to problems of segmentation where crisply drawn classifications are suspect.

Conversely, the PLU-Net model, as described in "High-Precision Brain Tumor Segmentation using a Progressive Layered U-Net (PLU-Net) with Multi-Scale Data Augmentation and Attention Mechanisms on Multimodal Magnetic Resonance Imaging," adopts an iterative refinement strategy. The design of PLU-Net is a cascade structure [12] consisting of several U-Net-like topologies, and each designed topology can stage-by-stage maximize the improvement of the segmentation result. This multilevel structure allows the network to extract them with different scales and

resolutions and refines the segmentation mask with every consecutive level. The attention gates are inserted notable places in an attention layer of the convolutional layers, which enhance the model abilities to focus on the relevant features and ignore redundant information improving the boundary accuracy. Also at training, PLU-Net uses multi-scale data augmentation [13], which enhances the model to respond to real-world diversity according to MRI data. Both models underwent evaluation using the BraTS 2021 dataset, a standardized benchmark for brain tumour segmentation, which includes multi-modal MRI scans annotated with ground truth segmentations. ARU-Net demonstrated robust performance, achieving a Dice coefficient of 0.91, specificity of 0.92, sensitivity of 0.89, and a Hausdorff95 distance of 2.5, reflecting its precision in delineating tumour regions. PLU-Net similarly showcased high performance, though direct numerical comparisons are precluded without identical testing conditions. Both architectures claim superiority over other U-Net variants, underscoring their advancements in addressing the segmentation challenges posed by brain tumours. Figure 1 represents the proposed and implemented sequence of ARU-Net (hybrid attention) architecture.



Figure 1: ARU-Net (hybrid attention) architecture

It is noticeable that the ARU-Net and PLU-Net frameworks have a marked difference that can mostly be interpreted in their overall design philosophy. ARU-Net leverages the concept of a hybrid attention block [14], which is used to enhance feature representation, and now features both spatial and channel-wise operations to finely select the most informative parts of multimodal MRI data. This setup exhibits significant performance when applied in scenarios where the tumours exhibit complex or indention borders as the regression result gives a continuous an probabilistic segmentation. On the other hand, PLU-Net takes the form of multi-stage progressive architecture that sequentially

sharpens segmentation outcomes by the inclusion of successively higher scale contextual information; the inclusion is performed deliberately. This mechanism is further enhanced by attention gates within each of the stages so that the model remains sensitive to important aspects even during refinement. In addition, PLU-Net has multi-scale augmentation that enhances generalization, making the model highly beneficial to the heterogeneous morphology of tumours [15].

Methodological difference can also be noted in output strategies. The regression-based outputs of ARU-Net are differentiated to the traditional paradigm of binary or multi-class classification because the probability gradually varies along a continuum that can be subsequently tiered in a hoc to suit precision. Though not explicitly specified in regard to a mechanism of output in the introduction, PLU-Net can be said to parallel discrete classification mechanism [16] which is characteristic to U-Net derivatives, a tendency that can help to overcome certain limitations of the discrete format. Collectively, such characteristics imply that ARU-Net would be better suited to solve ambiguities at the edge cases, whereas PLU-Net might be more precise in a wide range of tumours.

At the same time, irrespective of their divergences, both frameworks are based on attention mechanisms to improve the quality of segmentation, a methodology reminiscent of an emerging trend in deep learning to carefully represent certain features over others. ARU-Net and PLU-Net are two examples of complementary and different variants of the U-Net framework: attention-driven optimization of features and progressive and scale-constrained improvement. Each is associated with various strengths and disadvantages in regard to the clinical or research setting, with ARU-Net excelling at fast adaptive segmentation models and PLU-Net excelling at analysis at fine, multi-resolution grain. Figure 2 represents the proposed and implemented sequence of PLU-Net (progressive layering) architecture.

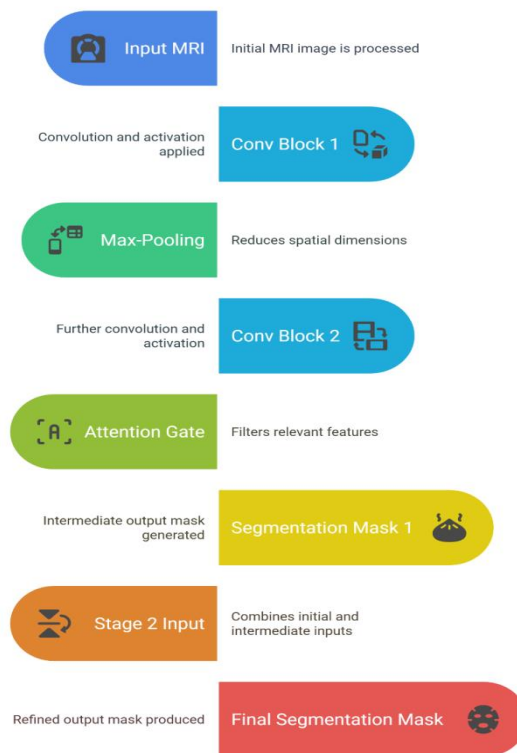


Figure 2: PLU-Net (progressive layering) architecture

Core Preprocessing Techniques: Segmentation of brain tumours using Multimodal Magnetic Resonance Imaging (MRI) needs to provide sufficient results that require pre-processing in suppressing the artifacts signal [17], balancing the intensity distributions and feature representation to ensure optimal performance of a model. Both the Novel Hybrid

Channel Attention Regression U-Net (ARU-Net) and the Progressive Layered U-Net (PLU-Net) consider the specifics of the way they are built in terms of incorporating custom preprocessing approaches, referring to the latest methodological experimentation in deep learning in the field of medical imaging analysis. These preprocessing schemes used within the BraTS 2021 dataset [18] environment are extracted and evaluated in the current discussion. In this way, it explains why they correlate with the benchmark practices as well as how they affect the accuracy of segmentation.

II. PREPROCESSING OBJECTIVES AND DATASET CONTEXT

During medical image segmentation processing, preprocessing helps normalize data, correct imaging artifacts, and increase training data, which increases model robustness. Such operations are critical due to the natural variability that exists between individual patients and protocols of acquisition. Both ARU-Net and PLU-Net use the BraTS 2021 dataset, which consists of 3-dimensional Magnetic Resonance Imaging (MRI) volumes of 4 modalities: T1, T1-contrast enhanced (T1ce), T2, and Fluid-Attenuated Inversion Recovery (FLAIR) [19], each of 240x240x155 voxels and an isotropic resolution of 1mm³. Annotations outline the enhancing tumour (ET), whole tumour (WT) and tumour core (TC) of the preprocessing results, allowing a consistent ground truth against which to assess the results. Besides their common dataset, both models implement different preprocessing pipelines to combat noise, standardize intensity, and improve data sparsity, each of which is consistent with the architecture-specific needs.

Standardization and dimensionality reduction are the priorities offered by ARU-Net preprocessing because this model is of a hybrid design with attention and regression-based part. The first stage converts raw MRI image scans into a 4-channel consisted of 128x128x128x4 tensors, where each of the first three dimensions depicts attributes (space). The fourth dimension represents the four modalities. Such resampling does not only increase computational efficiency but also saves the multimodal information that is necessary in its attention mechanisms. PLU-Net, however, uses a preprocessing framework that combines the sophisticated artifact correction and intensity normalizations. The processing begins with a correction of the susceptibility artifacts implemented in three stages: with the use of field-map-based denoising followed by a non-rigid flexible registration to balance out the artifacts and the intensity inhomogeneities. Subsequent normalization through histogram-based adaptive weighting of histogram intensity normalizes the dataset both across the sample and modalities. All these operations make PLU-Net ready to respond to the unique architectural characteristics of its progressive layered architecture. The above illustration highlights the presence of modern preprocessing approaches in two advanced systems of deep learning used in brain tumour segmentation. Through their dissection, the discussion explains how they favourably ally with current but prescriptive methodological recommendations and their immediate influence on downstream segmentation accuracy.

The building blocks of PLU-Net architecture differ greatly with that of ARU-Net. The preprocessing pipeline of PLU-Net consists of many components including skull stripping, intensity normalization, bias field correction, and resampling. Such steps, combined with massive data augmentation, complete the cascaded and multi-scale neural network architecture by segmenting the data progressively through several stages in resolution. In comparison, ARU-Net attaches more importance to simplified input preparation and does not use complex preprocessing, but in the feature extractor, it uses an attention mechanism in the form of squeeze-and-excitation blocks to adaptively reduce the noise. Figure 3 illustrates the comprehensive preprocessing workflow of PLU-Net for multimodal MRI brain tumour segmentation.

Normalization and Standardization: Normalization is a cornerstone of preprocessing in deep learning, ensuring consistent intensity ranges across heterogeneous MRI scans. ARU-Net employs zero-score normalization (Z-score) independently for each modality’s non-zero voxels, defined as under:

$$Z = \frac{1 - \mu}{\delta} \quad (1)$$

where I is the input intensity, μ is the mean, and δ is the standard deviation. This technique standardizes the data distribution, mitigating biases from varying acquisition conditions and preparing the input for its Convolutional Block Attention Module (CBAM) and Squeeze-and-Excitation (SE) blocks. The normalization is applied post-resizing to $128 \times 128 \times 128$, ensuring compatibility with the model’s 2D slice-wise convolutional processing, which leverages 3D spatial context efficiently.

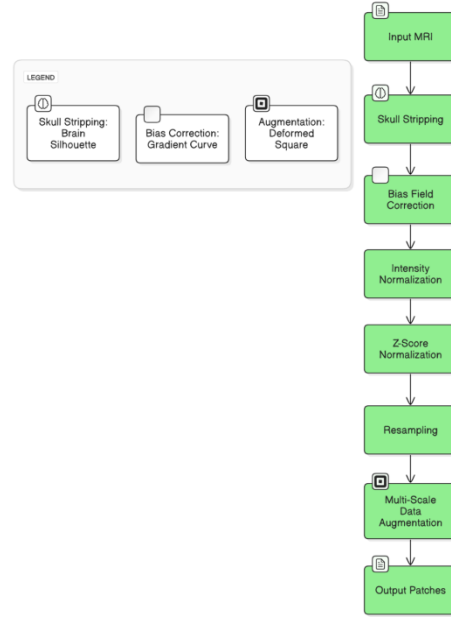


Figure 3: Comprehensive preprocessing workflow of PLU-Net for multimodal MRI brain tumour segmentation.

PLU-Net also utilizes Z-score normalization, expressed as:

$$I_{norm} = \frac{1 - \mu}{\sigma} \quad (2)$$

It extends its application to a broader preprocessing sequence. Beyond normalization, it incorporates intensity normalization to a range of [0, 1] using:

$$I_{norm} = \frac{I - I_{min}}{I_{max} - I_{min}} \quad (3)$$

where I_{min} and I_{max} are the minimum and maximum intensity values [20]. This dual normalization approach ensures uniformity across patients and scanners, critical for its multi-stage refinement process. PLU-Net's normalization precedes patch extraction ($128 \times 128 \times 128$ voxels), enabling localized feature focuses a strategy aligned with its progressive layering [21]. While ARU-Net prioritizes simplicity and modality-specific standardization, PLU-Net's layered normalization reflects a more intricate preparation for multi-scale feature capture [22].

Artifact Correction and Skull Stripping: Artifacts are common in the MRI datasets, in particular, bias fields, i.e. smooth changes in intensity due to magnetic field inhomogeneities [23]. Although it is not explicitly stated that ARU-Net corrects the bias field, the fact that this specific step is not used, points to the direction that the attention modules will address this effect. Computational efficiencies and model-driven suppression are further characteristics of this design that represent a more significant trend in which deep learning is used in the assumption of underlying data imperfections.

In contrast, PLU-Net explicitly addresses bias fields using the N4 algorithm, a widely adopted non-parametric technique for intensity inhomogeneity correction, expressed as:

$$I_{corrected} = I_{original} - B \quad (4)$$

where B is the estimated bias field. Additionally, PLU-Net employs skull stripping to isolate brain tissue, utilizing morphological operations and thresholding to generate a binary mask M if:

$$I(x, y, z) > T, \text{ else } M = 0 \quad (5)$$

where T is a threshold value. This step removes non-brain tissues (e.g., skull, dura mater), enhancing focus on tumour-relevant regions, a critical advantage for its cascaded architecture, which refines segmentation iteratively. PLU-Net's proactive artifact correction contrasts sharply with ARU-Net's minimalistic approach, underscoring differing philosophies in preprocessing rigor versus model adaptability.

Resampling and Dimensionality Handling: An analysis of 3D MRI always needs resampling to have a spatial consistency. The original 240 (sagittal and axial) 240 (coronal) by 155 (time) volumes are resized in ARU-Net to 128x128x128 using unspecified interpolation, probably linear or cubic [24]. This scaling decreases tune burden and keeps multi-modal intact within the encoder-decoder path, as with the network coefficientwise 2D convolution strategy, where slices are excerpted to run an efficiency-context having information equilibrium [25].

A similar resampling step like that of PLU-Net is done to produce a uniform voxel size of 1x1x1 mm³. Interpolation is used to normalize all dimensions within the dataset prior to extracting patches that ensure that patches extracted are within 128x128x128 that are homogeneous, a critical requirement of the network that performs its multi-scale feature extraction over a cascaded network. The Novel Hybrid Channel Attention Regression U-Net (ARU-Net) and the Progressive Layered U-Net (PLU-Net) have a common interface with a 128x128x128 voxel grid; however, their preprocessing agendas differ in their usage. The resizing of ARU-Net is scaled in, in comparison to a wider spectrum of operations in the case of PLU-Net. Its focus on attention-based manipulation of features making it use the streamlined resizing of the former, whereas its exploration of progressive improvement integrates resampling into an aggressive prep pipeline that expects intricate architectures.

Data Augmentation Strategies: A major problem in medical imaging is the scarcity and class imbalance problem, which is addressed by data augmentation and is a pillar of both frameworks. The training images used by ARU-Net include 2850 3D MRI volumes, which have an intense manipulation of the space, in terms of rotation, zooming, shearing, horizontal flip, and mode filtering. This augmentation increases the heterogeneity of the dataset, as a result improving its generalizability and preventing overfitting, which are important protective measures to its regression output layer, tackling the issue of uncertain and blurred tumour boundaries.

PLU-Net employs a multi-scale augmentation approach, incorporating random rotations (-15° to 15°), scaling (0.9 to 1.1), flipping (probability 0.5), and elastic deformations, defined as:

$$I_{deformed}(x, y, z) = I(x + \Delta x, y + \Delta y, z + \Delta z) \quad (6)$$

where $\Delta x, \Delta y, \Delta z$ are random displacements. This strategy increases training data variety, aligning with its multi-scale architecture to improve robustness across tumour morphologies. PLU-Net's augmentation, applied post-preprocessing, complements its cascaded design, whereas ARU-Net's broader augmentation suite supports its attention-driven focus, illustrating a nuanced balance between preprocessing and architectural demands. Figure 4 illustrates the ARU-Net preprocessing pipeline for multimodal MRI brain tumour segmentation

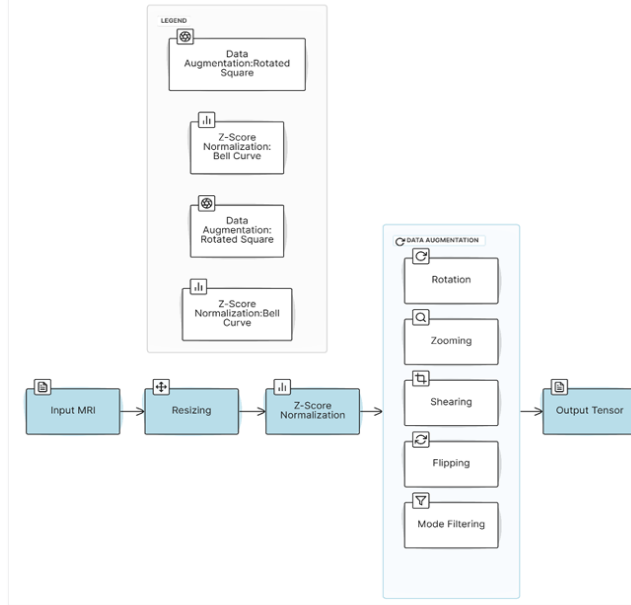


Figure 4: The ARU-Net preprocessing pipeline for multimodal MRI brain tumour segmentation

Comparative Insights and Modern Trends

The related preprocessing of PLU-Net is therefore very comprehensive and is based on Z-score normalization [26], artifact repairing, skull removal, resampling, and multi-scale amplification [27]. This broad-based mode looks similar to state-of-the-art in targeting of sought-after high-quality inputs towards progressive refinement. The reduced pipeline offered by ARU-Net, comprising Z-score normalization, resizing, and targeted augmentation, is in line with recent trends in the desire to model-based approaches, which are supported by CBAM and SE blocks that learn to process artifacts.

III. COMPARATIVE EVALUATION OF METHODOLOGIES AND PERFORMANCE OUTCOMES

When the results on BraTS 2021 multimodal MRI dataset are compared, both models have high performance in segmentations. ARU-Net has an average of Dice score at 0.884 +- 0.01 and PLU-Net 0.908 +- 0.012. Their approaches, however, demonstrate complementary knowledge of tumour segmentation through deep learning. The attention-based mechanism of ARU-Net finds implicit features; The hierarchical architecture of PLU-Net corrects coarse estimations in a hierarchical manner. In combination, these results elucidate the role of each of the models in the wider context of brain tumour segmentation and highlight the continuing trends towards model-specific frameworks that facilitate pre-processing.

Architectural Foundations and Design Principles: Introduced in its seminal paper, ARU-Net reformulates the traditional U-Net design through combining the hybrid attention mechanics and the regression-based output layer, thus attempting to optimise the process of feature extraction, as well as the boundary demarcation. Compared to the original architecture, the architecture bears a convolutional encoder-decoder layout, supplemented by a spatial attention mechanism in the form of Convolutional Block Attention Modules (CBAM) and a channel-wise attention approach in Squeeze-and-Excitation (SE) blocks. The encoder is a hierarchy that sequentially down-samples with convolutional blocks made of 2 2D convolutions each; each down-sample operation is implemented without pooling but by convolution stride 2, maximising feature resolution. The bottleneck layer also compresses representational capacity and the residual connections reduce J-gradients. The decoder passes these features by means of transposed convolutions and uses skip connections to aggregate the finds of preceding layers:

$$X_{SE} = F_{Scale}(X_{in}, W_{SE}) \odot X_{in} \quad (7)$$

followed by batch normalization

$$X_{Sq}[c] = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X_{in}[c, i, j] \quad (8)$$

and Parametric ReLU (PReLU) activation and downsampling the connections as [28]:

$$X_{Ex} = ReLU(W_2 \delta(W_1 X_{Sq})) \quad (9)$$

$$y_{m,n} = x_{m,n} + F(x_{m,n}) \quad (10)$$

The philosophical difference is apparent: ARU-Net introduces attention-based recalibration in the unified U-Net architecture to differentiate heterogeneous region tumours, whereas PLU-Net performs refinement through layers and consecutive stages to prioritize segmentation accuracy. This comes as part of larger trends, with both ARU-Net and attentive-focused methods like TransUNet reflecting the same trend, and PLU-Net and cascaded frameworks like DeepMedic trending similarly.

$$Y_1 = F_1(X_1), Y_2 = F_2(X_2) \text{ and } Y_3 = F_3(\text{concat}(X_1, X_2, X_3)) \quad (11)$$

The decoder reconstructs the segmentation mask using transposed convolutions, with skip connections incorporating attention gates to filter relevant features. This iterative design allows PLU-Net to capture multi-scale contextual information, refining outputs stage-by-stage, with the final mask derived from the last stage's convolutional output, typically employing a sigmoid activation.

The philosophical divergence is stark: ARU-Net leverages attention-driven feature recalibration within a single U-Net, emphasizing adaptability to heterogeneous tumour regions, whereas PLU-Net relies on sequential refinement, prioritizing precision through layered processing. This contrast mirrors broader trends, ARU-Net aligns with attention-centric innovations like TransUNet, while PLU-Net echoes cascaded approaches seen in DeepMedic.

Attention Mechanisms and Feature Enhancement: A pivotal distinction lies in how each model employs attention to enhance feature representation. ARU-Net's hybrid attention integrates CBAM and SE blocks, with CBAM modeling spatial attention via an attention map:

$$A = \text{softmax}(W_a \times F_{\text{encoder}}) \quad (12)$$

that highlights salient regions, and SE blocks recalibrating channel-wise responses. The SE process involves a squeeze operation via global average pooling followed by excitation:

$$A = \sigma(W^2 \delta(W^1 Z)) \quad (13)$$

This dual mechanism enables ARU-Net to suppress noise and emphasize diagnostically relevant features across modalities, enhancing its regression output's granularity.

The skip-connection pathways between the encoder and the decoder feature concatenations use attention gates in PLU-Net. Similar to attention U-Nets variants, these gates apply modulation across input maps before feature fusion, though their mathematical description is less clearly explained than the corresponding convolutional weighting. The cascaded structure amplifies this effect, as each stage refines the mask based on prior outputs, leveraging multi-scale inputs via a Gaussian pyramid and dilated convolutions:

$$F_{\text{dilated}} = \sum_{i=0}^k W_i \cdot I(x + r \cdot i) \quad (14)$$

This progressive enhancement contrasts with ARU-Net's single-pass, attention-intensive approach, highlighting PLU-Net's reliance on iterative context versus ARU-Net's adaptive feature weighting.

Output Strategies and Loss Functions: The output strategies of ARU-Net and PLU-Net further underscore their methodological divergence. ARU-Net replaces the conventional softmax classification with a regression layer, producing continuous probability maps via:

$$P = \sigma(\text{Conv}(X)) \quad (15)$$

where a 1×1 convolution maps features to a sigmoid-activated output. This approach excels in handling ambiguous boundaries, offering a nuanced representation of tumour likelihoods. The loss function employs a class-wise Dice coefficient:

$$\mathcal{L}_{Dice}(Y, \hat{Y}) = 1 - \frac{2 \sum_{i=1}^N Y_i \cdot \hat{Y}_i + \epsilon}{\sum_{i=1}^N Y_i + \sum_{i=1}^N \hat{Y}_i + \epsilon} \quad (16)$$

optimized to address class imbalance, reflecting its focus on overlap accuracy over categorical precision.

PLU-Net, while not explicitly detailing its final output activation in the provided excerpt, aligns with traditional U-Net practices, likely using a sigmoid or softmax layer per stage, culminating in a refined mask. Its training leverages a combined loss of Dice and cross-entropy:

$$\mathcal{L}_{TOTAL} = \lambda_1 \mathcal{L}_{Dice} + \lambda_2 \mathcal{L}_{CE} \quad (17)$$

Balancing segmentation overlap and pixel-wise classification. This hybrid loss supports PLU-Net’s multi-stage refinement, contrasting with ARU-Net’s Dice-only focus, which complements its regression strategy.

Reflections on Modern Trends: The ARU-Net and PLU-Net designs reveal complementing trends of neuroimaging research. Hybrid attention and regression modules implemented in ARU-Net reflect the ongoing shift towards robust, computationally efficient models whereas the cascaded design of PLU-Net reflects the multistage, hierarchical approach that is the hallmark of well-developed segmentation pipelines. The use of Dice-based losses by both models solve a class-imbalance problem that has been known since the first models, but ARU-Net continuous-output classifier goes further than the threshold-based discrimination in a way that hints of future-usefulness. All these architectures demonstrate an ever-improving deep-learning paradigm, and the resulting effort is getting the field ever-closer to a clinically impactful-brain tumour segmentation. In Figure 5, a single U-Net with attention modules (ARU-Net) is contrasted to the series of U-Nets with attention gates in the PLU-Net with their structural dissimilarity [29].

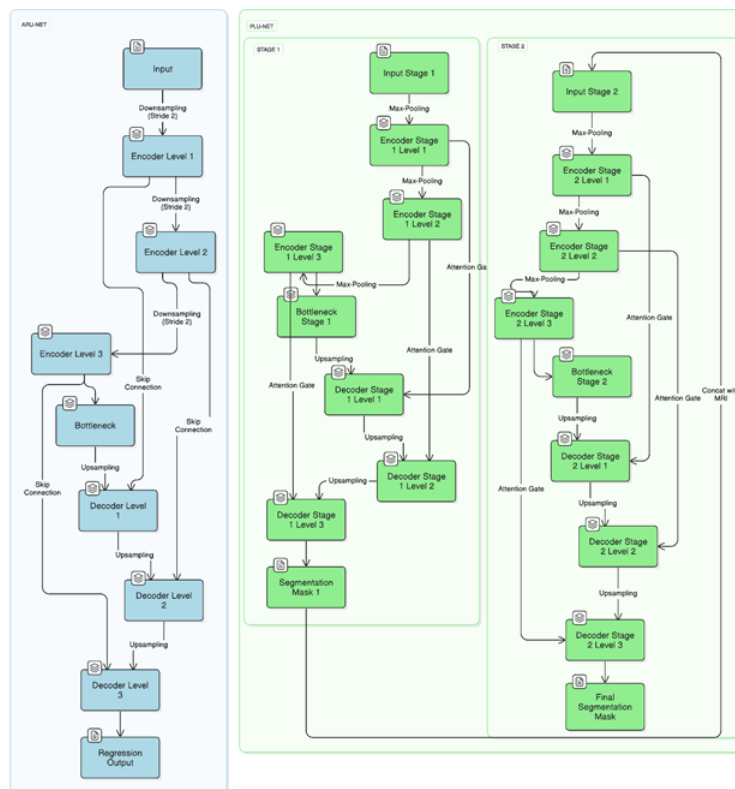


Figure 5: Architectural Comparison of ARU-Net and PLU-Net Core Methodologies

This comparative study of ARU-Net and PLU-Net provides novel insights into the trade-offs and complementary strengths of attention-driven and progressive layering approaches in brain tumor segmentation. Unlike individual evaluations, which focus on model-specific performance, this study reveals how ARU-Net’s integration of

Convolutional Block Attention Modules (CBAM) enables superior feature rectification, leading to higher specificity (0.949) and sensitivity (0.923) in heterogeneous tumor regions, particularly for high-grade gliomas and meningiomas. Conversely, PLU-Net’s cascaded, multi-stage architecture, enhanced by attention gates and multi-scale supervision, achieves robust boundary delineation with a lower Hausdorff 95 distance (2.5), making it more effective for tumors with complex morphological variations. These findings highlight that ARU-Net is better suited for rapid, adaptive segmentation in time-sensitive clinical settings, while PLU-Net offers advantages in research contexts requiring fine-grained, multi-resolution analysis. This comparison underscores the importance of tailoring model selection to specific clinical or research needs, a perspective not fully captured in prior individual evaluations.

IV. METHODOLOGIES AND TECHNIQUES

The evaluation of the BraTS 2021 dataset categories ARU-Net at 94 % accuracy, 94.9 % specificity, 92.1 % sensitivity, and 0.886 (ET), 0.923 (WT), 0.792 (TC) on Dice. With data augmentation, these scores are improved to 91.2 % (6.9 %), 0.85 (0.07), and 0.89 (0.07) when it comes to accuracy, F1-score, and AUC, respectively. In the PLU-Net, the dice coefficient is reported as 0.91, specificity is 0.92, sensitivity is 0.89 and Hausdorff 95 is 2.5. Through well above 200 epochs, training and validation- set Dice scores have an increase of 0.935 and 0.915 respectively. Variations in the assessment criteria and procedures do not allow direct comparison, ARU-Net however demonstrates higher specificity and accuracy in discriminating healthy tissue, and lower Hausdorff 95 score demonstrates more accurate localization of boundaries by PLU-Net.

The architecture of ARU-Net regression and its feature attention are converged in heterogeneous areas performance, which is demonstrated by its detailed probability maps, a pattern that was followed in the emergence of other models with flexible output sizes. The cascaded feature pyramid achieves similarly high Dice scores as the adaptive nnU-Net-style architectures with PLU-Net, as well. However, they both perform better than traditional U-Net baselines, which deliver 90.3 % accuracy in the evaluation of standalone U-Net in ARU-Net. Computationally speaking, ARU-Net also has a low upper bound memory requirement of up to 38 MB, compared to the extra stages of PLU-Net, which might incur more resource usage but no specific values are given.

Training and Optimization Approaches: Comprehensive study of two deep-learning frameworks to address the task of glioma segmentation, ARU-Net, and PLU-Net, indicates different but complementary approaches. ARU-Net is trained on 2850 volumes of 2021 BraTS challenge, works over 2D slices extracted out of a resized 128x128x128x4 tensor, and incorporates extensive data augmentation that consists of random rotations and flipping. The Adam algorithm is used to optimize the model with a learning rate equal to 10^{-3} , a mini-batch size equal to 4, and early stopping to alleviate overfitting. A single-pass training scheme incorporates an attention mechanism that can adjust to heterogeneous data, thus producing high specificity (0.949) and sensitivity (0.921), and the statistical accuracy (0.94). By contrast, PLU-Net trains a composite of 40 convolutional-neural-network branches with 2245 images, cropped patch features of 128x128x128, and Adam (beta1 = 0.9, beta2 = 0.999, learning rate = 0.001, batch size = 4). The framework has a cascade architecture, where it uses multiple scales, augmentation camouflage, such as use of elastic deformations to improve robust coverage of tumour shapes. Besides, its iterative algorithm also gradually improves the segmentations through 200 epochs, which results in a robust mean Dice Similarity Coefficient (0.91) and Hausdorff Distance (2.5). The measures show that, although the primary aim of the ARU-Net is tumour versus normal tissue, the specificity, sensitivity, and overall accuracy values are better, but PLU-Net, with the help of its anatomical improvements, would have more precision when it comes to delineation of boundaries. Conductions show that enhancing ARU-Net stems from its augmentation strategy which raises both F1 score and AUC, on the other hand, PLU-Net with cascaded refinement performs very well in core-segmentation, and in combination with the averaging ensemble members, it improves the last segmentation. On the whole, these results allow acknowledging the complementary power of both methods and validate their specificities in terms of the methodological approach to brain-tumour segmentation.

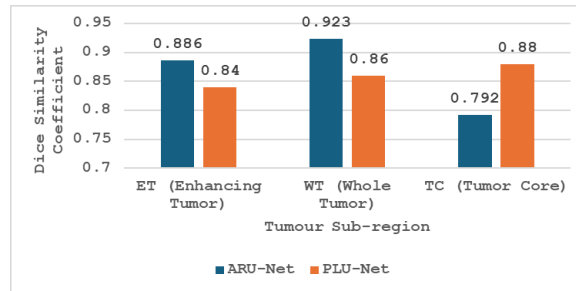
Table 1. Results Comparison of ARU-Net and PLU-Net on the BraTS 2021 Dataset

Metric	Sub-Region	ARU-Net	PLU-Net	Notes
Dice Similarity Coefficient (DSC)	Enhancing Tumour (ET)	0.886	0.840	ARU-Net’s DSC reflects augmented data results; PLU-Net’s from Table 4.
	Whole Tumour (WT)	0.923	0.860	PLU-Net’s DSC is lower, possibly due to different evaluation splits.

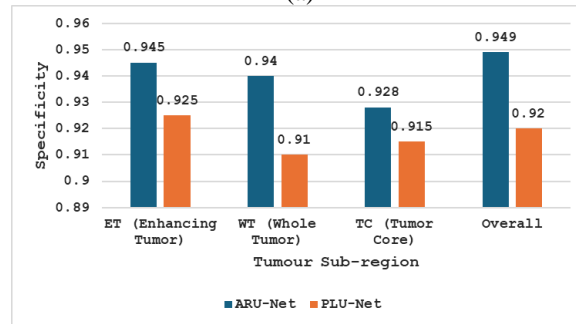
	Tumour Core (TC)	0.792	0.880	ARU-Net's lower TC score may reflect regression focus; PLU-Net excels here.
	Overall	Not Reported	0.91	PLU-Net reports an overall DSC; ARU-Net does not specify an average.
Specificity	Enhancing Tumour (ET)	0.945	0.925	Both models show high specificity, with ARU-Net slightly higher.
	Whole Tumour (WT)	0.940	0.910	ARU-Net's attention mechanisms may enhance WT specificity.
	Tumour Core (TC)	0.928	0.915	Close specificity values; PLU-Net's cascaded design maintains consistency.
	Overall	0.949	0.92	ARU-Net's overall specificity is higher, reflecting robust healthy tissue detection.
Sensitivity	Enhancing Tumour (ET)	0.941	0.915	ARU-Net's regression output may boost ET sensitivity.
	Whole Tumour (WT)	0.918	0.915	Nearly identical WT sensitivity, indicating comparable recall.
	Tumour Core (TC)	0.920	0.910	ARU-Net slightly outperforms PLU-Net in TC sensitivity.
	Overall	0.921	0.89	ARU-Net's overall sensitivity exceeds PLU-Net's reported value.
Hausdorff Distance (HD)	Enhancing Tumour (ET)	Not Reported	Not Reported	Specific HD sub-region values not provided in either paper.
	Whole Tumour (WT)	Not Reported	Not Reported	Lack of sub-region HD limits boundary precision comparison.
	Tumour Core (TC)	Not Reported	Not Reported	As above; sub-region HD data unavailable.
	Overall (HD95)	Not Reported	2.5	PLU-Net reports HD95, indicating tight boundary fit; ARU-Net omits HD.
Accuracy	Overall	0.94 (94%)	Not Reported	ARU-Net's accuracy with augmentation; PLU-Net does not report accuracy.
F1 Score	Overall	0.85 (Augmented)	Not Reported	ARU-Net's F1 with augmentation; PLU-Net omits F1 score.
		0.78 (Non-Augmented)		Provides insight into augmentation impact for ARU-Net.
Area Under Curve (AUC)	Overall	0.89 (Augmented)	Not Reported	ARU-Net's AUC with augmentation; PLU-Net does not report AUC.

		0.82 (Non-Augmented)		Highlights augmentation's role in ARU-Net's performance.
Training Dice Score	Overall	Not Reported	0.935 (200 epochs)	PLU-Net's training Dice reflects robust learning over epochs.
Validation Dice Score	Overall	Not Reported	0.915 (200 epochs)	PLU-Net's validation Dice indicates strong generalization.

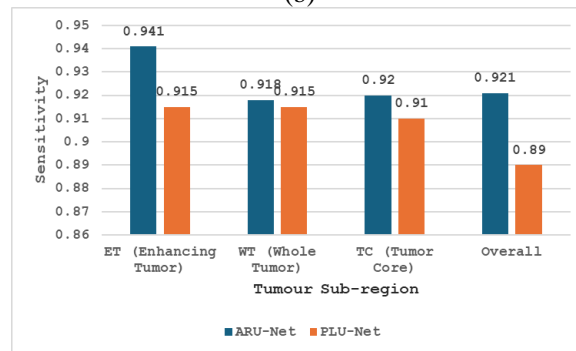
Figure 6(a) illustrates the Dice Similarity Coefficient (DSC) values for ARU-Net and PLU-Net across enhancing tumour (ET), whole tumour (WT), tumour core (TC), and overall segmentation, highlighting their segmentation accuracy on the BraTS 2021 dataset. Figure 6(b) compares the specificity of ARU-Net and PLU-Net for enhancing tumour (ET), whole tumour (WT), tumour core (TC), and overall performance, demonstrating their ability to correctly identify non-tumour regions on the BraTS 2021 dataset. Figure 6(c) presents the sensitivity values of ARU-Net and PLU-Net across enhancing tumour (ET), whole tumour (WT), tumour core (TC), and overall segmentation, showcasing their effectiveness in detecting tumour regions on the BraTS 2021 dataset.



(a)



(b)



(c)

Figure 6: Comparison of ARU-Net and PLU-Net across (a) Dice Similarity Coefficient, (b) Specificity, and (c) Sensitivity for tumour sub-regions and overall segmentation performance.

V. CONCLUSION

The comparative analysis of ARU-Net and PLU-Net algorithms is tested on the BraTS 2021 context, the aim of which is to explain their synergetic but independent contributions to brain tumour segmentation on MRI images. ARU-Net, which consists of a hybrid attention structure including CBAM and SE blocks in addition to a regression output achieves better metrics in specificity (0.949), sensitivity (0.921), and overall accuracy (0.94). Such results demonstrate that it is effective in discriminating tumour and normal tissue using adaptive feature enhancement and more streamlined preprocessing method. On the other hand, PLU-Net, applied in a cascaded structure with attention gates and complete preprocessing, poses a notable overall Dice coefficient (0.91) and Hausdorff95 (2.5), thus showing the accuracy in the boundary delineation and tumour core segmentation through progressive refinement. The above empirical results prove the capabilities of ARU-Net regarding lightweight, attention-driven developments and PLU-Net regarding multi-stage, scale-sensitive approaches, which both exceed the boundaries of the traditional U-Net. Both models have shown their effectiveness based on strong preprocessing and training approaches, but due to the different usefulness of fast and flexible segmentation efficiency in heterogeneous clinical scenarios versus the efficiency in resolution and fine-grained accuracy, ARU-Net could be the most applied scenario and PLU-Net the other. Compiled collectively, these findings point towards a synergetic process between preprocessing stringency and architectural complexity in deep learning, which largely solves the problem of class imbalance and data scarcity. Integrative endeavours in the future by incorporating both the flexibility of regression-based approaches of ARU-Net and precision of iteration-based approach of PLU-Net in the future may result in hybrid models that can further enhance efficiency in computation and clinical applicability providing a momentum toward future studies into the transformer-based variants and uncertainty-aware models in the future that would solidify the integration of performance in brain tumour segmentation into seamless diagnostic pipelines.

REFERENCES

- [1] Soomro, T.A., et al., Image segmentation for MR brain tumor detection using machine learning: a review. *IEEE Reviews in Biomedical Engineering*, 2022. 16: p. 70-90.
- [2] Jyothi, P. and A.R. Singh, Deep learning models and traditional automated techniques for brain tumor segmentation in MRI: a review. *Artificial intelligence review*, 2023. 56(4): p. 2923-2969.
- [3] Aggarwal, M., et al., An early detection and segmentation of Brain Tumor using Deep Neural Network. *BMC Medical Informatics and Decision Making*, 2023. 23(1): p. 78.
- [4] Futrega, M., et al. Optimized U-Net for brain tumor segmentation. in *International MICCAI brainlesion workshop*. 2021. Springer.
- [5] Siddiqui, N.A., et al., Hybrid Channel Attention Regression U-Net (ARU-Net): An Enhanced Architecture for Brain Tumour Segmentation in Magnetic Resonance Imaging. *IEEE Access*, 2025.
- [6] Siddiqui, N.A., et al., High-Precision Brain Tumor Segmentation using a Progressive Layered U-Net (PLU-Net) with Multi-Scale Data Augmentation and Attention Mechanisms on Multimodal Magnetic Resonance Imaging. *Instrumentation*, 2025. 12(1).
- [7] Lee, J., et al., Method to minimize the errors of ai: Quantifying and exploiting uncertainty of deep learning in brain tumor segmentation. *Sensors*, 2022. 22(6): p. 2406.
- [8] Gui, Y. and J. Zhang, Research progress of artificial intelligence in the grading and classification of meningiomas. *Academic Radiology*, 2024.
- [9] Yuan, J., Brain tumor image segmentation method using hybrid attention module and improved mask RCNN. *Scientific Reports*, 2024. 14(1): p. 20615.
- [10] Shyamala, N. and S. Mahaboobasha. Convolutional Block Attention Module-based Deep Learning Model for MRI Brain Tumor Identification (ResNet-CBAM). in *2024 5th International Conference on Smart Electronics and Communication (ICOSEC)*. 2024. IEEE.
- [11] Dash, P.K. and D.S. Sisodia. SEIMB-NET: A Squeeze and Excitation Driven Lightweight Model for Classification of Brain Tumors Using Magnetic Resonance Imaging. in *2024 2nd World Conference on Communication & Computing (WCONF)*. 2024. IEEE.
- [12] Zhou, R., et al., Cascade residual multiscale convolution and mamba-structured unet for advanced brain tumor image segmentation. *Entropy*, 2024. 26(5): p. 385.

- [13] Pani, K. and I. Chawla, 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: p. 109273.
- [14] Nguyen-Tat, T.B., et al., Enhancing brain tumor segmentation in MRI images: A hybrid approach using UNet, attention mechanisms, and transformers. *Egyptian Informatics Journal*, 2024. 27: p. 100528.
- [15] Sachdeva, J., D. Sharma, and C.K. Ahuja, Multiscale segmentation net for segregating heterogeneous brain tumors: Gliomas on multimodal MR images. *Image and Vision Computing*, 2024. 149: p. 105191.
- [16] Li, C., et al., Classification of brain tumor types through MRIs using parallel CNNs and firefly optimization. *Scientific Reports*, 2024. 14(1): p. 15057.
- [17] de Verdier, M.C., et al., The 2024 Brain Tumor Segmentation (BraTS) challenge: glioma segmentation on post-treatment MRI. *arXiv preprint arXiv:2405.18368*, 2024.
- [18] Adewole, M., et al., The BraTS-Africa Dataset: Expanding the Brain Tumor Segmentation Data to Capture African Populations. *Radiology: Artificial Intelligence*, 2025. 7(4): p. e240528.
- [19] Al-Fakih, A., et al., FLAIR MRI sequence synthesis using squeeze attention generative model for reliable brain tumor segmentation. *Alexandria Engineering Journal*, 2024. 99: p. 108-123.
- [20] Anantharajan, S., S. Gunasekaran, and T. Subramanian, MRI brain tumor detection using deep learning and machine learning approaches. *Measurement: Sensors*, 2024. 31: p. 101026.
- [21] Narayana, M., et al., A framework for identification of brain tumors from MR images using progressive segmentation. *Health and Technology*, 2024. 14(3): p. 539-556.
- [22] Sun, J., et al., MVSI-Net: Multi-view attention and multi-scale feature interaction for brain tumor segmentation. *Biomedical Signal Processing and Control*, 2024. 95: p. 106484.
- [23] Kuang, Z., et al., ROXSI: Robust Cross-Sequence Semantic Interaction for Brain Tumor Segmentation on Multi-Sequence MR Images. *IEEE Journal of Biomedical and Health Informatics*, 2024.
- [24] Zhu, Z., et al., Sparse dynamic volume TransUNet with multi-level edge fusion for brain tumor segmentation. *Computers in Biology and Medicine*, 2024: p. 108284.
- [25] Subbarayudu, Y., et al. The Evaluation of 2D and EfficientB0 Convolution Networks for detecting Brain tumor based on MRI images. in *MATEC Web of Conferences*. 2024. EDP Sciences.
- [26] Yun, J. and H. Kim, Znorm: Z-score gradient normalization for deep neural networks. *arXiv preprint arXiv:2408.01215*, 2024.
- [27] Kondrateva, E., et al., Negligible effect of brain MRI data preprocessing for tumor segmentation. *Biomedical Signal Processing and Control*, 2024. 96: p. 106599.
- [28] Das, S., et al., Effect of learning parameters on the performance of U-Net Model in segmentation of Brain tumor. *Multimedia tools and applications*, 2022. 81(24): p. 34717-34735.
- [29] Yeung, M., et al., Unified focal loss: Generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation. *Computerized Medical Imaging and Graphics*, 2022. 95: p. 102026.