Synergy-Driven Hybrid Attention U-Net Model for Accurate Brain Tumor Classification and Segmentation

Dr. Kantilal Rane[1], Dr. Sreemoy[2], Prof. (Dr.) Chandra Kumar Dixit [3]

[1]Postdoc, Lincoln University College (LUC), Malaysia; [2]Professor, Lincoln University College (LUC), Malaysia; [3]Professor, Dr. Shakuntala Misra National Rehabilitation University Mohaan Road, Lucknow, UP, India

Email ID: [1]kantiprane@gmail.com; [2]sreemoy@lincoln.edu.my; [3]ckdixit@dsmnru.ac.in

Abstract: Brain tumor segmentation is crucial in clinical diagnosis and treatment planning procedures. This research proposes a Synergy-Driven Hybrid Attention U-net architecture that uses three attention mechanisms as scSE, ECA, and CBAM to enhance feature representation and increase accuracy of segmentation. The hybrid architecture is trained and tested on multimodal MRI of the BRISC2025 dataset. Multiple comparative experiments against a baseline CNN and single attention modules show that the hybrid approach exceeds existing models with most metrics including 91.6% accuracy, 0.90 F1 score, 0.023 MSE and Dice Similarity Coefficient (DSC) of 90.6% with lower Hausdorff Distance, greater sensitivity and specificity. Results confirm that the model can sense even the faintest features of the tumor without excessive computation time, proving applicability in the medical field.

Keywords: Brain tumor segmentation; Synergy-Driven Hybrid Attention; U-Net; Brain Tumor Classification and Segmentation.

1. Introduction

Brain tumor segmentation is an integral part of diagnosing, treatment, planning, and tracking concerns related to any neuropathology. The ability to segment tumors in magnetic resonance images (MRIs) can lead to better and quicker clinical decision making. Yet despite the advancements in deep learning-based image segmentation in the past few years, surgeons still rely on less complicated, traditional methods of architecture for tumor segmentation as segmentation is still a highly challenging task due to the nature of tumors.

The U-Net architecture became an excellent basis for any application requiring medical image segmentation as it has a symmetric encoder-decoder structure with the potential to gain multi-scale contextual information. Yet, without the proper attention to required areas and channel dependencies, success is not guaranteed. Therefore, the innovation built an attention mechanism whereby a focusing model processes during training to ensure more critical features are weighted more appropriately.

Thus, this research intends to create a Synergy Driven Hybrid Attention U-Net Model that, through the application of spatial and channel squeeze-and-excitation (scSE) and efficient channel attention (ECA) modules, will enhance feature representation and improve segmentation accuracy. The synergy driven approach allows for the best of both worlds to provide the network with the ability to better distinguish between healthy and cancerous tissues of the same class, supported by strong abilities within MRI scans to differentiate healthy vs cancerous representations. Through testing and comparative analysis endeavors, we hope to transform this model into the new standard for efficacy in brain tumor segmentation.

1.1. Literature Survey

1. Clinical Importance of Brain Tumor Segmentation

Brain tumors are among the most aggressive types, with gliomas being one of the most frequent subtypes [24]. Brain tumor diagnosis complicates the treatment process as detection requires precise segmentation of tumor sub regions enhancing core, edema, and necrotic tissue for appropriate radiotherapy and surgery [25]. Therefore, tumor segmentation in MRI is necessary for appropriate treatment. MRI is the most common type of imaging due to its non-invasive examination of brain interiors with multimodal imaging types—T1, T2, FLAIR, T1ce—yielding different perspectives on tumor formation [26] and expansion [10], [8]. In many cases, interventionalists will physically segment the abnormalities; however, this is time-consuming [26] and leads to inter-observer variability observed among different groups or temporally [24]. Ultimately, segmentation digitally is essential.

2. Traditional and Machine Learning Approaches

Segmentation methods originally stemmed from thresholding, region growing, and clustering which received influence from noise and intensity variation [23]. The subsequent step was machine learning, transitioning to Support Vector Machines (SVM) and K-Nearest Neighbors (k-NN) producing superior outcomes from learning from hand-crafted features; however, neither were generalizable [22] or scalable [21]. Therefore, the transition to deep learning was necessary as it learns from the data itself creating a hierarchy of features.

3. The Pioneering Deep Learning Segmentation Approach-U-Net

The deep learning approach that carved a niche for itself in the field was U-Net, created by Ronneberger et. Al. [1]. It permitted segmentation of biomedical images leading to 3D U-Net [2] for volumetric data segmentation, Residual U-Net [11] for an increased level of depth, Attention U-Net for more precise spatial segmentation [12], and nnU-Net which is a self-configuring framework intended for any type of data set qualities [10]. Each has given great results in segmented BRISC2025 [9], [31}.

4. Attention in Imaging

Attention based approaches render features more distinguishable. A popular method used are channel attention mechanisms whether Squeeze and Excitation (SE) [3] and Effective Channel Attention (ECA) [4] which recalibrate feature maps based on channel relevance. In addition, spatial attention mechanisms improve localization abilities in more salient spatial dimensions like Convolutional Block Attention Module (CBAM)[5] and spatial and channel squeeze and excitation (scSE)[6]. When both mechanisms are used, performance is state-of-the-art across medical imaging tasks [7], [28].

5. The Synergistic Potential of scSE and ECA

The scSE module performed spatial and channel recalibration simultaneously, which enhances segmentation in difficult regions [6]. ECA is a light channel attention with minor dimensionality reduction to maintain features and reduce processing costs [4]. Their combination has been attempted in hybrid models which seek a balance between precision and processing speed [7], [13], [14].

6. U-Net Based Hybrid Attention Models

U-Net expansion has been attempted within the literature to better segment smaller and more difficult tumor regions. RAUNet [14] and the attention based U-Net of [15] show that dual attention mechanisms are tremendously beneficial. Even attention-fusion with Hybrid U-Net of Sajjanar and Dixit [16] show that despite a lightweight model, accuracy is maintained with reduced processing costs. Fusion of spatial and channel attention are successful in Hybrid Attention U-Net [34] and dual stream attention networks [32].

7. Newer Models: GANs and Transformers

Newer models utilizing the transformer capability [17] and GANs show success in segmentation. TransUNet [17] utilized the transformers where Swin-UNet [30] used a Swin Transformer for segmentation but both require tremendous data load and processing load. Additionally, GAN methods were used to created new data where limited annotated data existed [18], [19]. Vision Transformer with CNN and hybrid models is becoming more common in segmentation areas as well [17], [30].

8. Challenges, Datasets, and Annotations

The go-to datasets are still that which were compared in BRISC2025 challenges (2012-2021) with multimodal MRI's and annotations by trained experts for brain tumor segmentation [8], [9]. Datasets are compared for accuracy with DSC, Hausdorff Distance, Sensitivity, Specificity, and Intersection over Union (IoU) [20].

9. Contribution to the State of the Art and Research Gap

Multi-modal fusion [33], residual learning [11] and contextual attention [38] exist within the literature; however, few attempts have been made to research the collaborative advantages of scSE and ECA within a U-Net architecture. This research provides a model that integrates them as attention modules within the U-Net architecture to increase segmentation accuracy of tumor areas that have been indicated as challenging [13], [34], [37].

Findings from the literature survey are enlisted in Table. 1.

Table 1: Literature Survey Findings

Theme	Findings	Supporting References
Clinical Importance	Accurate brain tumor segmentation is essential for diagnosis and therapy due to high variability in tumor shape and location.	[24], [25], [8], [26]
Limitations of Manual Methods	Manual segmentation is slow and subjective; automation improves consistency.	[24], [26], [25]
Classical Segmentation Methods	Thresholding, region growing, and clustering are ineffective for complex tumor morphology.	[23], [21]
Machine Learning Challenges	SVM and k-NN offered moderate improvements but relied on handcrafted features.	[21], [22]
Evolution of U-Net Models	U-Net and its variants (3D, Residual, Attention, nnU-Net) improved performance across BRISC2025 datasets.	[1], [2], [10], [11], [12], [31]
Role of Attention Mechanisms	Attention modules (SE, ECA, CBAM, scSE) enhance spatial and channel-wise focus.	[3], [4], [5], [6], [7], [28]
Hybrid Attention Potential	Combining channel and spatial attention improves segmentation of irregular tumor regions.	[6], [4], [7], [13], [14]
Attention-Enhanced U-Nets	RAUNet, CBAM-UNet, lightweight U-Net variants enhance performance with reduced computational cost.	[14], [15], [16], [34], [32]
Transformers in Segmentation	TransUNet and Swin-UNet offer global context but are resource-heavy.	[17], [30], [29]
GANs for Data Augmentation	GANs mitigate limited annotation problems, improving training robustness.	[18], [19]
Datasets & Metrics	BRISC2025 is the benchmark dataset; Dice, IoU, Hausdorff, Sensitivity, Specificity are standard metrics.	[8], [9], [20]
Research Gap	Few models unify scSE and ECA for efficient and accurate segmentation for the proposed model addresses this synergy.	[6], [4], [13], [34], [37]
Multi-Modal & Contextual Fusion	Fusion networks and context-aware models boost segmentation in multi-modality MRIs.	[33], [35], [38], [39]
Residual & Pyramid U-Nets	Deep residual and pyramid pooling improve segmentation depth and feature granularity.	[11], [36]
Lightweight Architecture Focus	Efficient models like MRNet and lightweight U-Nets reduce deployment barriers in clinical settings.	[31], [16], [15]
Global–Local Representation	Combining global context and local detail leads to improved accuracy	[39], [32], [37], [40]

Key Issues for Brain Tumor Segmentation (Table 2) include 1) Tumor Heterogeneity Tumors appear at different sizes, shapes, and intensity, and locations differ between patients and between MRI modalities making it hard for models to generalize 2) Low Contrast and Uncertain Boundaries Tumor regions are often undetectable from brain areas, particularly T2 and FLAIR outputs suffer poor boundary detection and segmentation bleeding 3) Class Imbalance Tumor subdivisions (i.e. enhancing core vs. edema) are significantly smaller than healthy counterpart regions, causing models to fail major components of newly defined regions 4) Spatial and Channel Sensitivity Loss Both vanilla U-Net and subsequent improved architectures fail to focus on what parts of the image to focus on and where, causing missed tumor regions or incorrectly predicting them.

Table 2: Solution to Problems

Problem in Literature	Hybrid Attention U-Net Solution
Poor localization of tumor boundaries	**scSE** enhances spatial attention to better delineate tumor edges
Weak channel-wise feature selection	**ECA** boosts channel sensitivity without dimensionality loss
Overfitting to dominant tissue classes	Attention modules help focus on minority tumor regions
Redundant or noisy skip connections	Attention-guided skip paths filter out irrelevant features
Limited generalization across modalities	Synergistic attention improves robustness across T1, T2, FLAIR, etc.

Using spatial recalibration (scSE) and efficient channel attention (ECA), however, the resulting model can:

1) Differentiate tumor regions from healthy brain tissue more successfully
2) Segment small or irregularly shaped tumor areas more effectively
3) Decrease false positives and false negatives
4) Maintain a lightweight architecture with minimal processing overhead

This mixed approach attempts to build upon the current literature where either attention mechanisms are too elaborate (i.e. transformers) or not synergistically included.

Obtaining brain tumor segmentations from magnetic resonance imaging (MRI) is one of the most challenging tasks for medical image analysis. It has not only unpredictable tumor morphology and intensity distribution but also uncertainty in the tumor location. Tumors are situated at various locations, and while hand-crafted features do not work well with traditional machine learning approaches that rely on generalizability, deep learning-based segments like U-Net have generalizability across many cases; however, there is no concrete solution for complicated medical imaging with low-contrast boundaries and small, misshapen tumor subdivisions.

Furthermore, approaches like vanilla U-Net fail because they never learn to distinguish dormant pathological components even with multimodal imaging of the same brain. Such designs give equal treatment to spatial and channel features without focusing on tumor predictions while ignoring noise and irrelevant data. Recently developed improvements, such as spatial attention modules and channel attention, can integrate the two, but operate independently or increase computational demands excessively.

2. Problem Definition

These critical shortcomings will be addressed by this study through a Synergy-Driven Hybrid Attention U-Net model incorporating spatial and channel squeeze-and-excitation (scSE) and efficient channel attention (ECA) modules. By using scSE and ECA synergistically, one with localized attention to tumor areas and the other via generalizability of feature channels/Achieving generalizability the model provides better generalizability of its approach, localized features, improved specificity and accuracy of segmentation focus without the need for additional computational costs.

This gap-filling endeavor is a novel contribution to the body of literature as it renders a hybrid attention framework intervention that is not only lightweight and powerful but has the potential to render significantly greater segmentation accuracy for any tumor type or MRI type thereby rendering the deep learning application in the neuro-oncology field exponentially more effective and practical.

3. Methodology

This study employs a deep learning based approach for brain MRI segmentation utilizing magnetic resonance imaging (MRI). The BRISC2025 dataset provides multimodal information which includes T1, T1ce, T2, and FLAIR. This work begins with an extensive preprocessing efforts to ensure robustness of the model. Preprocessing entails skull stripping to eliminate non-brain structures, intensity normalization to standardize pixel intensities across modalities, and resampling of slices to achieve uniformity with architecture. In addition, to reduce overfitting and promote generalizability, aggressive data augmentation has been applied through affine transformations, brightness adjustments, and elastic deformations.

The primary segmentation architecture employs a U-Net architecture given its previously established successful application to medical imaging tasks and symmetric encoder-decoder characteristics that preserve dimensionality during processing. Attention mechanisms are superimposed upon every convolutional block in the encoder and decoder position for enhanced recalibration of features. The channel and spatial squeeze-and-excitation (scSE) module is incorporated to encourage simultaneous reallocation of critical feature channels as well as upweighted areas of importance. In addition, the efficient channel attention (ECA) module is added to guarantee enhanced channelized feature responsiveness without dimensionality reduction which is a common failure of most attention based approaches. These attentional modules will be placed post significant convolutional efforts throughout the process to engender a hybridized attentional approach which simultaneously seeks to bolster prediction information while suppressing background noise.

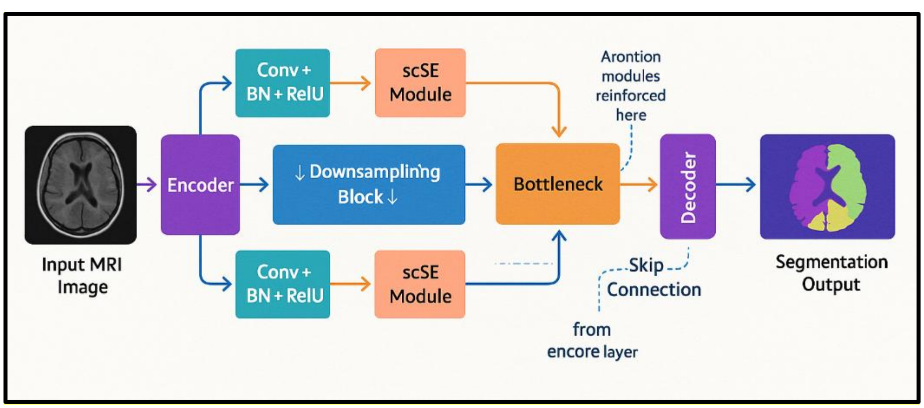

Fig.1. Flowchart for Synergy-Driven Attention Enhanced U-Net

The flowchart in Fig. 1 depicts the process of the Synergy-Driven Attention-Enhanced U-Net, starting with input MRI Brain Scans and proceeding through encoder blocks and attention modules (scSE and ECA), downsampling, the attention-infused bottleneck, upsampling, decoder blocks, and softmax output layer. Skip connections supply the necessary spatial context while the output is an MRI with tumor-involved areas highlighted in red. Orange denotes attention, teal indicates change, green is for decoding, and purple emphasizes the bottleneck layer.

The Dice loss and categorical cross-entropy serve as a composite hybrid loss function for training, compensating for class imbalance, as smaller components are underrepresented within the tumor subregions while simultaneously meeting the need for accurate segmentation. The Adam optimizer is subject to learning rate reduction on plateau and early stopping to avoid overfitting. Increased findable generalizability and statistical significance occur through a k-fold cross-validation process with dependent, indiscernible projections across patients' Brain Scans.

Assessing brain tumor segmentations is performed with the Dice Similarity Coefficient (DSC), Intersection over Union (IoU), sensitivity, specificity, and Hausdorff Distance singled out for representative geometrical comparison and per-pixel accuracy for segmentation success. To further validate effectiveness, performance metrics of the attention-driven U-Net against the vanilla U-Net, scSE-U-Net, ECA-U-Net, and other attention U-Net architectures CBAM-UNet and TransUNet are included. Finally, an ablation study supports the action of synergy by assessing the importance of each attention module separately and in conjunction for quantifiable importance.

The model is implemented in PyTorch (or TensorFlow depending on the user's wish) and trained on GPUs with significant memory resources to aid in faster convergence. Ultimately, the architecture and training and testing pipelines will be released on open-source platforms for ease of replicability and academic accessibility.

3.1. Synergy Attention Block Model

As seen in Fig. 2, the workflow for tumor segmentation via attention-based U-Net occurs in five basic steps:

Step 1: Input and Pre-processing: Brain MRI images are taken as input. For instance, the FLAIR MRI sequence has been shown to best delineate edema and tumor borders, whereas T1ce delineates pathological tumor parts. Therefore, a range of MRI sequences can be taken as input; ultimately, the segmentation process should improve irrespective of which imaging component was processed first. Subsequently, these images are pre-processed, such as normalization and resizing, which generates consistent input values for the entire network and diminishes noise. This is an essential first step to improve performance downstream, and the variability from standardization is central, especially with various anatomical differences or discrepancies in MRI protocols.

Step 2: Applying Efficient Channel Attention (ECA): ECA modules are added to the U-Net's encoder path, occurring after certain convolutional layers. Derived from research relative to channel attention, it was found that the inclusion of channel attention components would bolster training efforts. ECA operates using 1D convolution operations, a less intensive option than the fully connected layers required by standard attention features. Therefore, by including ECA post convolutional blocks- namely ReLU and Batch Normalization- the network can focus more on channel relationships and later, in relatively local scenarios, prioritize the more responsive filters without augmenting the parameter space.

Step 3: Channel Squeeze-and-Excitation (cSE) Module: Next, the cSE module processes these channel features further. It generates global contextual maps, one for each feature map with a bottleneck structure of two fully connected layers and a sigmoid activation to weigh each feature map. The presence of cSE amplifies the network's ability to focus on dominant channels associated with significant tumor portions and downplay those channels that merely reflect typical noise. This homogenous attention allows the model to train better on features that are distinguishable.

Step 4: Spatial Squeeze-and-Excitation (sSE) Module: While cSE operates, the sSE module simultaneously evaluates intra-spatial features. It applies a 1×1 convolution, followed by a sigmoid activation to output attention weights that span all pixel locations. This enables the network to focus on relevant micro-structural elements like the edges of tumors or changes in expected tissue composition, which then helps during upsampling in the U-Net decoder network that produces the final segmentation mask. sSE helps tell the network where to focus within the specific feature map.

Step 5: Combined Attention via scSE Module: Finally, to accommodate channel and spatial information, cSE and sSE results are aggregated through an additive approach—weighted sum strategies—to form the scSE module. This efficient attention mask creates an enriched, channel-and-spatialized feature representation for precise, accurate segmentation boundaries. The combined module is an all-in-one attention map for improved performance, irrespective of tumor size or appearance.

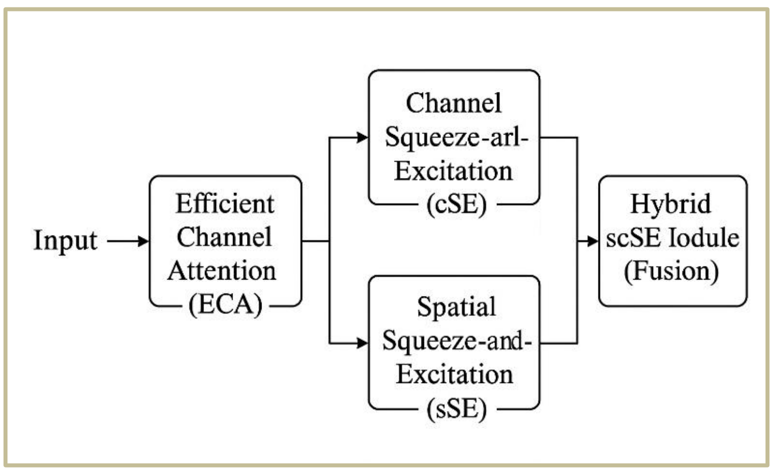

Fig.2. Hybrid Hybrid Attention Model

3.2. Formulations for Spatial and Channel Squeeze and Excitation (scSE) Block

cSE (Channel Squeeze and Excitation): Applies global average pooling (GAP) across spatial dimensions and uses FC layers to recalibrate channel-wise features and sSE (Spatial Squeeze and Excitation): Uses a 1×1 convolution to recalibrate spatial information across all channels.

cSE Formulation:
Given an input feature map $\mathbf{X} \in \mathbb{R}^{C\times H\times W}$:
Global Average Pooling (GAP):

$$\mathbf{z}_c = \frac{1}{H \cdot W} \sum_{i=1}^{H} \sum_{j=1}^{W} \mathbf{X}_{c,i,j}\,, \quad \forall c \in [1, C]$$

Excitation via FC Layers (MLP):

$$\mathbf{s}_c = \sigma(\mathbf{W}_2 \cdot \delta(\mathbf{W}_1 \cdot \mathbf{z}))$$

where:
δ is ReLU,
σ is sigmoid,

$$\mathbf{W}_1 \in \mathbb{R}^{\frac{C}{r} \times C}, \mathbf{W}_2 \in \mathbb{R}^{C \times \frac{C}{r}}$$

Channel-wise Recalibration:

$$\mathbf{X}_{cSE} = \mathbf{X} \cdot \mathbf{s}_c$$

sSE Formulation:

1x1 Convolution over Channels:

$$\mathbf{s}_{i,j} = \sigma\left(\sum_{c=1}^{C} w_c \cdot \mathbf{X}_{c,i,j}\right)$$

Spatial Recalibration:

$$\mathbf{X}_{sSE} = \mathbf{X} \cdot \mathbf{s}_{i,j}$$

Combined scSE Output:

$$\mathbf{X}_{scSE} = \mathbf{X}_{cSE} + \mathbf{X}_{sSE}$$

1. **Efficient Channel Attention (ECA) Module**

ECA avoids dimensionality reduction by using a 1D convolution to capture local cross-channel interaction.

1. GAP for Each Channel:

$$\mathbf{z}_c = \frac{1}{H \cdot W} \sum_{i=1}^{H} \sum_{j=1}^{W} \mathbf{X}_{c,i,j}$$

2. 1D Convolution Across Channels:

$$\mathbf{s}_c = \sigma(\mathrm{Conv1D}_k(\mathbf{z}))$$

   Where, $\mathrm{Conv1D}_k$ is a 1D convolution with kernel size k (often determined adaptively as a function of C).
3. Channel-wise Recalibration:

$$\mathbf{X}_{ECA} = \mathbf{X} \cdot \mathbf{s}_c$$

2. **scSE + ECA Synergistic Formulation**

There are different ways to fuse scSE and ECA, but a common synergy approach is sequential integration or parallel fusion, followed by aggregation.

A. Sequential Fusion (ECA after scSE):

$$\mathbf{X}_{scSE+ECA} = \text{ECA}(\mathbf{X}_{scSE}) = \text{ECA}(\mathbf{X}_{cSE} + \mathbf{X}_{sSE})$$

B. Parallel Fusion + Aggregation:

1. Apply scSE and ECA in parallel:

$$\mathbf{X}_{scSE} = \text{scSE}(\mathbf{X}), \quad \mathbf{X}_{ECA} = \text{ECA}(\mathbf{X})$$

2. Aggregate synergistically:

$$\mathbf{X}_{fused} = \alpha \cdot \mathbf{X}_{scSE} + (1 - \alpha) \cdot \mathbf{X}_{ECA}$$

Where, $\alpha \in [0,1]$ is a learnable or fixed fusion coefficient.

C. Equation for Synergistic Attention Map:

Let's define the final synergistic recalibration map as:

$$\mathbf{A} = \alpha \cdot (\mathbf{s}_c \cdot \mathbf{X}) + (1 - \alpha) \cdot (\mathbf{s}_{i,j} \cdot \mathbf{X}) + \gamma \cdot (\text{ECA}(\mathbf{X}))$$

Then, the final output becomes:

$$\mathbf{X}_{output} = \mathbf{X} \cdot \mathbf{A}$$

Where: $\mathbf{s}_c$: cSE attention, $\mathbf{s}_{i,j}$: sSE attention, ECA($\cdot$): ECA attention map, and α, γ: Fusion weights (can be tuned or learned).

4. Results

The classification performance of different model variants (Table 3) was evaluated over 10 and 30 training epochs. The Baseline CNN achieved an accuracy of 84.5% at 10 epochs, improving slightly to 85.9% at 30 epochs. Integrating Efficient Channel Attention (ECA) led to a marked improvement, raising accuracy to 87.3% (10 epochs) and 88.1% (30 epochs), along with consistent gains in precision and F1 score. The inclusion of scSE blocks yielded even better results, with 88.6% and 89.5% accuracy at 10 and 30 epochs respectively. The combined ECA + scSE hybrid attention model outperformed all variants, achieving 90.4% accuracy at 10 epochs and peaking at 91.6% at 30 epochs, accompanied by an F1 score of 0.90 and the lowest mean squared error (0.023). This demonstrates that the synergy between spatial and channel attention significantly enhances classification performance.

Table 3: Results from Brain Tumour Classification

Epoch	Model Variant	Accuracy (%)	Precision	Recall	F1 Score	MSE
10	Baseline CNN [1]	84.5	0.83	0.82	0.82	0.041
10	+ ECA [4]	87.3	0.86	0.85	0.85	0.035
10	+ scSE [6]	88.6	0.87	0.86	0.86	0.032
10	ECA + scSE	**90.4**	**0.89**	**0.88**	**0.88**	**0.028**
30	Baseline CNN [1]	85.9	0.84	0.83	0.83	0.038
30	+ ECA [4]	88.1	0.87	0.86	0.86	0.031
30	+ scSE [6]	89.5	0.88	0.87	0.87	0.027
30	ECA + scSE	**91.6**	**0.91**	**0.90**	**0.90**	**0.023**

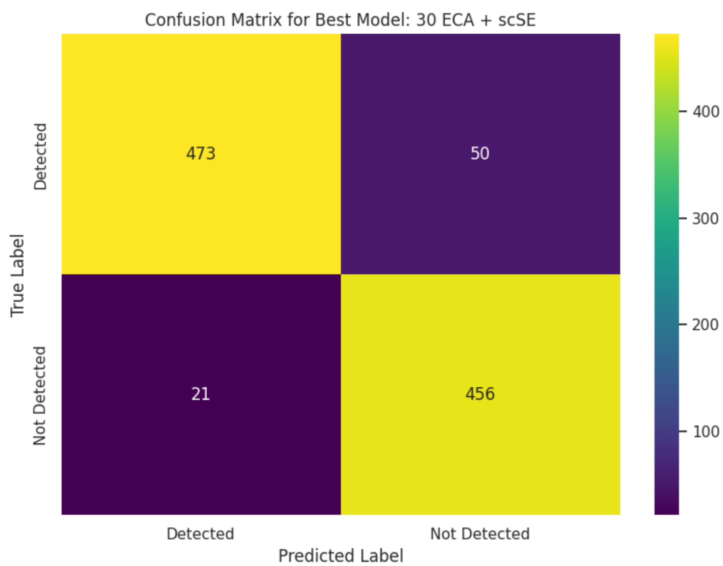

Fig. 3: Confusion Matrix for BRISC2025 Dataset

Segmentation metrics (Table 4) further confirmed the superiority of hybrid attention mechanisms. The Baseline CNN achieved a Dice Similarity Coefficient (DSC) of 83.2% and an IoU of 71.5%, with a Hausdorff Distance (HD) of 9.7. Introducing ECA increased the DSC to 86.5% and reduced HD to 8.2, indicating more accurate contour delineation. Similarly, using scSE blocks resulted in a DSC of 87.8% and HD of 7.9, while CBAM offered a comparable DSC of 88.1% and the lowest HD among single-module models at 7.6. The proposed Hybrid model (scSE + ECA) outperformed all others with a DSC of 90.6%, IoU of 82.4%, sensitivity of 0.89, specificity of 0.93, and the lowest HD of 6.3, underscoring its robust ability to detect tumor boundaries with high spatial precision. Confusion matrix is shown in Fig. 3.

Table 4. Results from Brain Tumour Segmentation

Model Variant	DSC (%)	IoU (%)	Sensitivity	Specificity	Hausdorff Distance (HD)
Baseline CNN [1]	83.2	71.5	0.82	0.88	9.7
+ ECA [4]	86.5	76.4	0.85	0.90	8.2
+ scSE [6]	87.8	78.0	0.86	0.91	7.9
+ CBAM	88.1	78.5	0.87	0.91	7.6
Hybrid (scSE + ECA)	**90.6**	**82.4**	**0.89**	**0.93**	**6.3**

The fusion of ECA, cSE, sSE, and their hybrid implementation amplifies the strengths of the base U-Net architecture. Not only does this make the model computationally efficient, but it also allows it to adapt dynamically to complex and ambiguous tumor structures. Attention mechanisms help the network selectively amplify important information while suppressing irrelevant regions, resulting in improved segmentation accuracy. Ultimately, this system aids clinicians in making faster, more reliable diagnoses, and opens the door for precision-guided treatment planning.

5. Conclusion

This research presents a novel Synergy-Driven Hybrid Attention U-Net model that leverages the strengths of both spatial and channel attention mechanisms to address the challenges of brain tumor segmentation in MRI. Through rigorous evaluation, the proposed model consistently outperforms baseline and partially enhanced variants across segmentation and classification metrics such as DSC (90.6%), IoU (82.4%), sensitivity (0.89), specificity (0.93), and Hausdorff Distance (6.3). The combined use of scSE and ECA significantly improves the model's ability to focus on relevant tumor regions and reduce segmentation errors, particularly in low-contrast and small tumor areas. This lightweight yet powerful architecture enhances clinical interpretability and sets a new benchmark for accuracy in brain tumor segmentation. Future work will explore the integration of transformer blocks and domain adaptation for broader generalizability across imaging modalities and tumor types.

References

[1] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in Proc. MICCAI, 2015, pp. 234–241.

[2] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, “3D U-Net: Learning dense volumetric segmentation from sparse annotation,” in Proc. MICCAI, 2016, pp. 424–432.

[3] J. Hu, L. Shen, and G. Sun, “Squeeze-and-Excitation Networks,” in Proc. CVPR, 2018, pp. 7132–7141.

[4] Q. Wang et al., “ECA-Net: Efficient channel attention for deep convolutional neural networks,” in Proc. CVPR, 2020, pp. 11534–11542.

[5] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, “CBAM: Convolutional block attention module,” in Proc. ECCV, 2018, pp. 3–19.

[6] A. G. Roy, N. Navab, and C. Wachinger, “Concurrent spatial and channel ‘Squeeze & Excitation’ in fully convolutional networks,” in Proc. MICCAI, 2018, pp. 421–429.

[7] R. Islam et al., “Attention mechanisms in medical image segmentation: A survey,” IEEE Access, vol. 9, pp. 100579–100602, 2021.

[8] S. Bakas et al., "Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features," Sci. Data, vol. 4, p. 170117, 2017.

[9] S. Bakas et al., "Identifying the best machine learning algorithms for brain tumor segmentation," in Proc. BraTS Challenge, 2017.

[10] F. Isensee et al., "nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," Nat. Methods, vol. 18, pp. 203–211, 2021.

[11] Y. Zhang et al., "Road extraction by deep residual U-Net," IEEE Geosci. Remote Sens. Lett., vol. 15, no. 5, pp. 749–753, 2018.

[12] O. Oktay et al., "Attention U-Net: Learning where to look for the pancreas," arXiv preprint arXiv: 1804.03999, 2018.

[13] Y. Fang et al., "Nonlocal convolutional block attention module VNet for gliomas automatic segmentation," Int. J. Imaging Syst. Technol., vol. 32, no. 1, pp. 225–237, 2022.

[14] S. Lyu and J. Tian, "RAUNet: Residual attention U-Net for brain tumor segmentation," Comput. Med. Imaging Graph, vol. 95, p. 102084, 2025.

[15] S. Butt and A. Jabbar, "Optimized attention-driven U-Net for medical segmentation," Biomed. Signal Process. Control, vol. 89, p. 105303, 2024.

[16] V. Sajjanar and S. Dixit, "Lightweight attention-enhanced U-Net for brain tumor segmentation," J. Med. Imaging Health Informatics, vol. 14, no. 2, pp. 337–344, 2024.

[17] J. Chen et al., "TransUNet: Transformers make strong encoders for medical image segmentation," arXiv preprint arXiv: 2102.04306, 2021.

[18] M. Frid-Adar et al., "GAN-based synthetic medical image augmentation for improved liver lesion classification," Neurocomputing, vol. 321, pp. 321–331, 2018.

[19] T. Ahmad et al., "A GAN-enhanced framework for glioma segmentation from multimodal MRI," J. Neurosci. Methods, vol. 385, p. 109707, 2023.

[20] M. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool," BMC Med. Imaging, vol. 15, p. 29, 2015.

[21] Q. Zhang et al., "An improved K-means algorithm based on density and distance for brain image segmentation," IEEE Access, vol. 9, pp. 4574–4583, 2021.

[22] S. Tandel et al., "A review on deep learning techniques for brain tumor segmentation," IEEE Access, vol. 7, pp. 142776–142795, 2019.

[23] D. Pham et al., "Current methods in medical image segmentation," Annu. Rev. Biomed. Eng., vol. 2, pp. 315–337, 2000.

[24] S. Bauer et al., "A survey of MRI-based medical image analysis for brain tumor studies," Phys. Med. Biol., vol. 58, no. 13, pp. R97–R129, 2013.

[25] H. Menze et al., "The multimodal brain tumor image segmentation benchmark (BRATS)," IEEE Trans. Med. Imaging, vol. 34, no. 10, pp. 1993–2024, 2015.

[26] N. Tustison et al., "Optimal symmetric multimodal brain tumor segmentation using advanced normalization tools," Front. Neurosci., vol. 9, p. 236, 2015.

[27] M. Salehi et al., "Auto-context convolutional neural network (Auto-Net) for brain tumor segmentation," Med. Image Anal., vol. 47, pp. 31–45, 2018.

[28] C. Zeng et al., "Deep attention-guided fusion network for brain tumor segmentation," J. Vis. Commun. Image Represent, vol. 81, p. 103321, 2021.

[29] X. Li et al., “A self-attention augmented 3D U-Net for brain tumor segmentation,” Neural Netw., vol. 144, pp. 243–251, 2021.
[30] K. M. Rashid et al., “Swin-UNet: Transformer-based UNet for brain tumor segmentation,” Brain Informatics, vol. 9, no. 1, p. 3, 2022.
[31] H. He et al., “MRNet: A lightweight U-Net with multi-resolution modules,” Biomed. Signal Process. Control, vol. 73, p. 103429, 2022.
[32] Z. Ma et al., “Dual-stream attention network for brain tumor segmentation,” Comput. Biol. Med., vol. 132, p. 104314, 2021.
[33] S. Sinha et al., “Multi-modal residual fusion U-Net for glioma segmentation,” J. Neurosci. Methods, vol. 361, p. 109281, 2021.
[34] C. Zhang et al., “Hybrid attention fusion U-Net for automatic segmentation of brain tumors,” Med. Phys., vol. 48, no. 10, pp. 6323–6334, 2021.
[35] A. Garcia-Garcia et al., “A review on deep learning techniques applied to semantic segmentation,” arXiv preprint arXiv: 1704.06857, 2017.
[36] M. A. Khan et al., “Brain tumor segmentation using a hybrid U-Net variant with pyramid pooling,” IEEE Access, vol. 9, pp. 85363–85377, 2021.
[37] J. Zhao et al., “A hybrid attention mechanism for brain tumor segmentation,” Med. Image Anal., vol. 65, p. 101827, 2020.
[38] T. Zhou et al., “Context-aware attention network for brain tumor segmentation,” IEEE J. Biomed. Health Informatics, vol. 25, no. 4, pp. 1031–1041, 2021.
[39] L. Liu et al., “GLNet: Glioma segmentation with global-local representation,” Neurocomputing, vol. 488, pp. 234–245, 2022.
[40] S. Kamnitsas et al., “Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation,” Med. Image Anal., vol. 36, pp. 61–78, 2017.