Enhancing Attention Synergy Networks with Lightweight Encoders for Efficient Brain Tumor Segmentation

Shaun Niel A. Ochavo
Computer Science Department, College of Computing and Information Sciences, Caraga State University

Armiex Jay J. Roble
Computer Science Department, College of Computing and Information Sciences, Caraga State University

Roland P. Abao
Computer Science Department, College of Computing and Information Sciences, Caraga State University

https://doi.org/10.5109/7395499

出版情報：Proceedings of International Exchange and Innovation Conference on Engineering & Sciences (IEICES). 11, pp.65-71, 2025-10-30. International Exchange and Innovation Conference on Engineering & Sciences
バージョン：

Enhancing Attention Synergy Networks with Lightweight Encoders for Efficient Brain Tumor Segmentation

Shaun Niel A. Ochavo[1], Armiex Jay J. Roble[2], Roland P. Abao[3]
[1,2,3]Computer Science Department, College of Computing and Information Sciences, Caraga State University – Main Campus, Ampayon, Butuan City, 8600, Philippines
Corresponding author email: rpabao@carsu.edu.ph

Abstract: *This study investigates the optimization of the Attention Synergy Network (AS-Net) for MRI-based brain tumor segmentation by evaluating the impact of replacing its original VGG16 encoder with lightweight alternatives: MobileNetV3 (Large/Small) and EfficientNetV2 (B0/B1). Each encoder was integrated into the AS-Net framework, which combines Spatial and Channel Attention Modules, and trained on the BraTS 2023-GLI dataset using 2D slice inputs. Results showed that EfficientNetV2-B0 achieved segmentation accuracy nearly matching VGG16 (Dice: 0.8853 vs. 0.8873) while reducing training time by ~20% and inference time by ~29%. MobileNetV3 models were faster but less accurate. EfficientNetV2-B0 emerged as the most balanced encoder, maintaining high accuracy with significantly improved efficiency. These findings highlight the potential of lightweight encoder integration for advancing practical, real-time clinical applications of AS-Net in brain tumor analysis.*

Keywords: Brain Tumor Segmentation; AS-Net; EfficientNetV2; MRI; BraTS 2023

1. INTRODUCTION

Accurate segmentation of brain tumors in Magnetic Resonance Imaging (MRI) scans is vital for diagnosis, treatment planning, and disease monitoring [1]. Manual segmentation, while accurate in expert hands, is time-consuming, labor-intensive, and prone to inter-observer variability, thereby necessitating robust automated approaches [2]. In recent years, deep learning—particularly Convolutional Neural Networks (CNNs)—has revolutionized medical image analysis, with architectures such as U-Net emerging as the standard for organ and lesion segmentation tasks [3].

Despite these advancements, brain tumor segmentation remains a challenging problem due to the inherent heterogeneity of tumors in size, shape, location, and appearance across patients. Moreover, indistinct tumor boundaries, class imbalance (tumors occupying a small fraction of the brain volume), and the computational burden of processing high-resolution 3D MRI data limit the real-world applicability of current deep learning models in clinical settings [1, 2].

The U-Net architecture, with its encoder-decoder design and skip connections, has been widely adopted and extended for brain tumor segmentation. Variants with deeper encoders, residual connections, and attention modules have been proposed to improve segmentation accuracy [4]. Among these enhancements, attention mechanisms—such as spatial, channel, and self-attention—have shown significant promise by allowing models to focus selectively on informative features [5, 6]. For instance, Attention U-Net introduced attention gates to suppress irrelevant background features during decoding [5].

Building on this direction, the Attention Synergy Network (AS-Net) was proposed to synergize spatial and channel attention mechanisms, achieving strong performance in skin lesion segmentation tasks [7]. However, its reliance on the computationally expensive VGG16 encoder limits its deployment in clinical environments where efficiency is critical.

To address this gap, the present study investigates the integration of lightweight encoder architectures—MobileNetV3 (Large and Small) and EfficientNetV2 (B0 and B1)—into the AS-Net framework. Specifically, this work aims to: (1) evaluate the impact of these encoders on segmentation performance and computational efficiency; (2) compare them against the original VGG16 baseline; and (3) assess whether AS-Net's attention mechanisms retain their effectiveness across different encoder backbones.

2. MOTIVATION AND RELATED WORK

Brain tumors, particularly gliomas, pose significant challenges in clinical neuroimaging due to their heterogeneous morphology and infiltrative growth patterns. Accurate segmentation of tumor regions in Magnetic Resonance Imaging (MRI) scans is crucial for diagnosis, treatment planning, and monitoring disease progression. However, manual segmentation is time-consuming, labor-intensive, and prone to inter-observer variability, underscoring the need for robust and efficient automated methods.

Deep learning—particularly Convolutional Neural Networks (CNNs)—has emerged as a powerful solution for medical image segmentation. Architectures like U-Net and its numerous variants have become widely adopted due to their ability to fuse low-level spatial detail with high-level semantic information [3, 8]. Enhancements to these models have increasingly incorporated attention mechanisms to improve focus on salient features. Spatial and channel attention, in particular, have been shown to significantly boost segmentation accuracy by enabling models to emphasize

tumor-relevant regions while suppressing irrelevant background [9, 10].

The Attention Synergy Network (AS-Net) [7] is a recent innovation that synergistically integrates spatial and channel attention modules within a U-Net-like architecture. Initially developed for skin lesion segmentation, AS-Net achieved state-of-the-art results, demonstrating the effectiveness of attention fusion in enhancing feature representation. However, its dependence on the computationally intensive VGG16 encoder limits its scalability and practical deployment, especially in clinical settings with constrained computational resources [11].

To address these limitations, recent research has focused on lightweight CNN architectures such as MobileNetV3 [12] and EfficientNetV2 [13], which are designed to optimize the trade-off between model accuracy and computational efficiency. These encoders have shown promising results across various medical imaging tasks, particularly when integrated into segmentation frameworks [14, 15].

Munira and Islam [16] further highlighted the potential of custom CNN architectures in brain MRI analysis, demonstrating successful multi-class tumor classification using specialized models. Their findings emphasize the importance of thoughtful network design tailored to the complexities of medical data.

Motivated by these developments, this study seeks to bridge the gap between segmentation accuracy and computational efficiency by integrating lightweight encoders into the AS-Net architecture. Specifically, we evaluate the performance of MobileNetV3 (Large and Small) and EfficientNetV2 (B0 and B1) within AS-Net, comparing their accuracy and training/inference efficiency against the original VGG16-based model. This research aims to identify encoder configurations that retain the effectiveness of AS-Net's attention mechanisms while enhancing its practicality for real-world clinical applications. The conceptual framework of this study is illustrated in Figure 1.

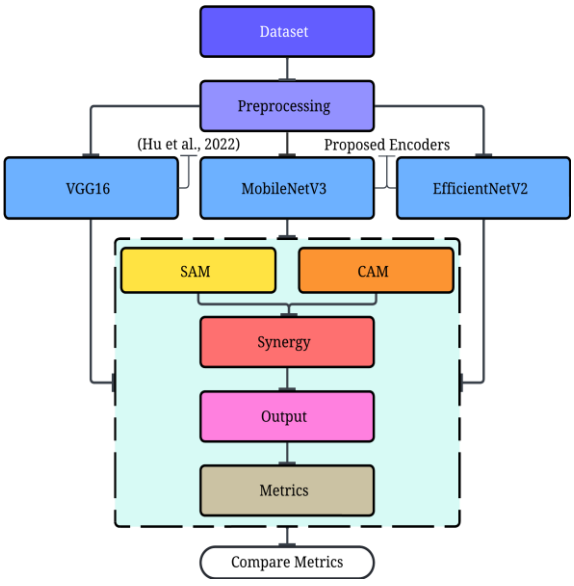


Fig. 1. Overall architecture of AS-Net showing encoder-decoder structure with attention modules.

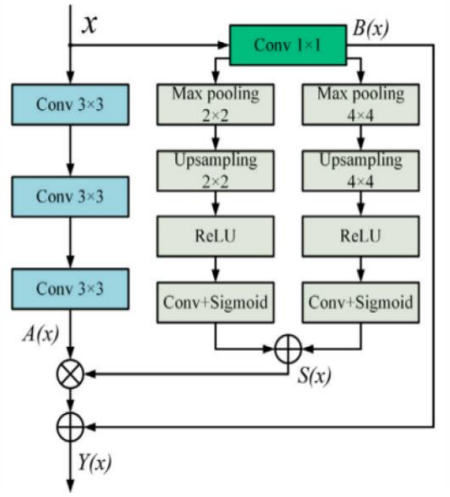
3. METHODS

3.1. AS-Net Architecture

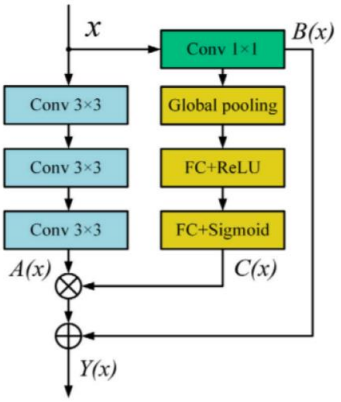
The Attention Synergy Network (AS-Net) adopts an encoder–decoder structure augmented with dual attention mechanisms to improve feature representation and segmentation performance. Its core components include:

- **Spatial Attention Module (SAM):** Applies parallel convolutional operations to generate spatial attention maps, enabling the model to focus on salient spatial regions (Fig. 2a).
- **Channel Attention Module (CAM):** Captures inter-channel dependencies using global average pooling followed by a multi-layer perceptron (MLP), producing channel-wise attention weights (Fig. 2b).
- **Synergy Module:** Combines SAM and CAM outputs using learnable scalar parameters (α and β), allowing the network to adaptively weight spatial and channel attention contributions.

This synergy mechanism effectively fuses spatial and channel cues, leveraging their complementary strengths to enhance segmentation accuracy during decoding.



(a) *Spatial Attention Module (SAM)*



(b) *Channel Attention Module (CAM)*

Fig. 2. Detailed view of (a) Spatial Attention Module (SAM) and (b) Channel Attention Module (CAM).

3.2. Encoder Integration

Five pre-trained encoder networks were integrated into the AS-Net framework: VGG16 (baseline) [7], MobileNetV3-Large & MobileNetV3-Small [12], and EfficientNetV2-B0 & EfficientNetV2-B1 [13]. Key characteristics of these encoder backbones considered in the study are summarized in Table 1.

Table 1. Comparison of encoder backbone characteristics.

Encoder	Parameters	FLOPs	Top-1 Accuracy (ImageNet)
VGG16	138M	15.5G	71.3%
MNv3-Large	5.4M	0.22G	75.2%
MNv3-Small	2.5M	0.06G	67.4%
ENv2-B0	7.1M	0.69G	78.7%
ENv2-B1	8.1M	1.21G	79.8%

For each encoder, specific intermediate layers were selected for skip connections to the decoder. These include:

- **VGG16:** Skip connections were established from 'block1_conv2', 'block2_conv2', 'block3_con v3', and 'block4_conv3' layers.
- **MobileNetV3-Large:** Skip layers were 're_lu_2', 're_lu_5', 're_lu_11', and 're_lu_14', with the bottle-neck at the final activation layer.
- **EfficientNetV2 variants (e.g., B0):** Corresponding layers such as 'block1a_project_bn, 'block2b_add', 'block4a_project_bn', and 'block6a_project_bn' were carefully selected to maintain spatial consistency at each decoder stage.

All encoders were initialized with ImageNet pre-trained weights, and a fine-tuning approach was employed where initial convolutional blocks were frozen while later blocks were unfrozen and trained. For example, VGG16 was frozen up to 'block3_conv3', with layers from 'block4_conv1' onwards being trainable, while for EfficientNetV2-B0, layers up to 'block3b' were frozen with blocks from 'block4a' onwards unfrozen. The over- all AS-Net architecture, illustrating the encoder-decoder structure with these attention modules and skip connections, is shown in Fig. 3.

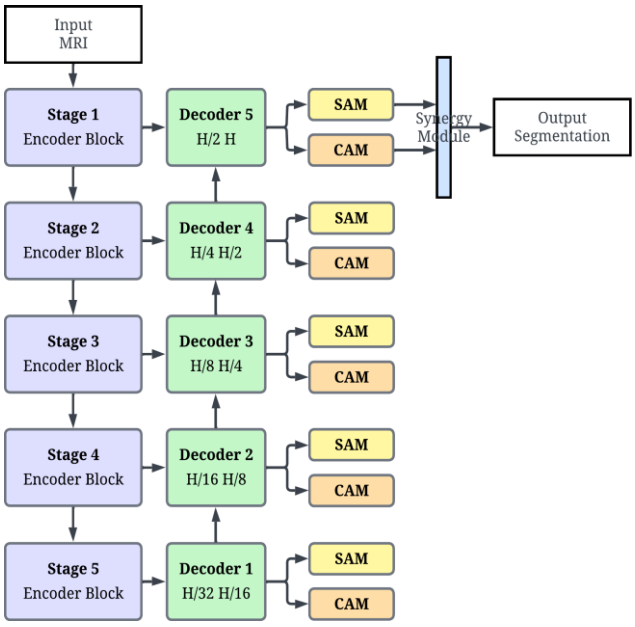


Fig. 3. AS-Net architecture with encoder-decoder structure, skip connections, and attention modules (SAM, CAM). VGG16 uses four decoder stages (H/16 to H); MobileNetV3 and EfficientNetV2 variants use five de-coder stages (H/32 to H).

3.3. Dataset and Preprocessing

The study utilized the BraTS 2023-GLI (Task 1) dataset [4, 17], containing multimodal 3D MRI scans (T1, T1ce, T2, FLAIR) and expert-annotated segmentation masks. A preprocessing pipeline converted 3D NIfTI volumes into 2D HDF5 slices, selecting T1ce and FLAIR modalities mapped to three channels ([T1ce, FLAIR, T1ce]). Z-score normalization was applied per slice and modality. Medical imaging research has shown that appropriate preprocessing techniques are critical for model performance, with methods such as Monte Carlo simulation being useful for understanding image characteristics, as demonstrated by Haider et al. [18] in their work on medical imaging with megavoltage photon beams. The binary target mask was created by labeling any tumor region as foreground (1) and background as 0. Data was split into training (80%), validation (10%), and test (10%) sets based on slices. A visual summary of this preprocessing pipeline is provided in Fig. 4.

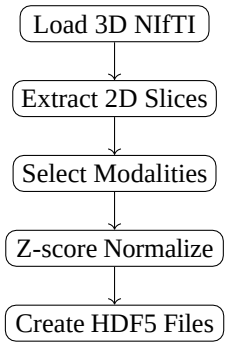


Fig. 4. Preprocessing pipeline for converting 3D MRI volumes to 2D slices for training.

3.4. Training Configuration

All models were trained on an NVIDIA Tesla V100 GPU (32GB) with the following hyperparameters:

- Optimizer: Adam (initial learning rate = 1e-4)
- Loss Function:

$$L_{\text{total}} = 0.5 \times L_{\text{WBCE}} + 0.5 \times L_{\text{Dice}}$$

  Weighted Binary Cross Entropy (WBCE) used a positive class weight of 100.0.
- Batch Size: 32
- Max Epochs: 30 with early stopping (patience = 15 on val_loss)
- Learning Rate Scheduler: ReduceLROnPlateau (factor = 0.5, patience = 5 on val_loss)

3.5. Evaluation Methodology

Performance was assessed using:

- Segmentation Accuracy Metrics: Dice coefficient, Intersection over Union (IoU), precision, recall, and binary accuracy
- Computational Efficiency Metrics: Total training time and average inference time per slice

The Dice Coefficient was computed as:

$$\text{Dice} = \frac{2 \times TP}{2 \times TP + FP + FN}$$

And the Intersection over Union (IoU) is computed:

$$\text{IoU} = \frac{TP}{TP + FP + FN}$$

where TP, FP, and FN represent true positives, false positives, and false negatives, respectively.

3.6. Implementation Details

All models were implemented in TensorFlow 2.19 and trained on an NVIDIA Tesla V100 GPU (32GB). The codebase supports modular encoder integration, with skip connections and fine-tuning strategies tailored for each backbone. Data preprocessing, model training, and evaluation pipelines were fully automated, and key hyperparameters (batch size, learning rate, loss weights) were kept consistent across variants for fair comparison. The complete code and configuration details are available at `https://github.com/shaunniekins/asnet-brats-encoder-comparison`.

4. RESULTS AND DISCUSSION

4.1. Learning Dynamics and Convergence

The training history plots for each AS-Net variant provide a detailed view of convergence behavior across multiple evaluation metrics: loss, accuracy, Dice coefficient, IoU, precision, and recall. These plots are presented in Fig. 5 to Fig.9.

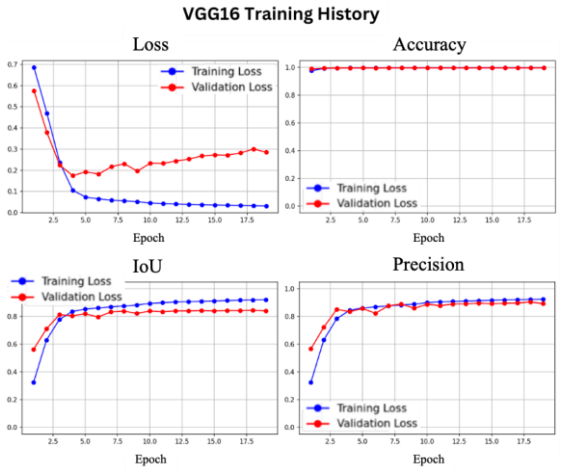


Fig 5. Training plot for AS-Net with VGG16 encoder.

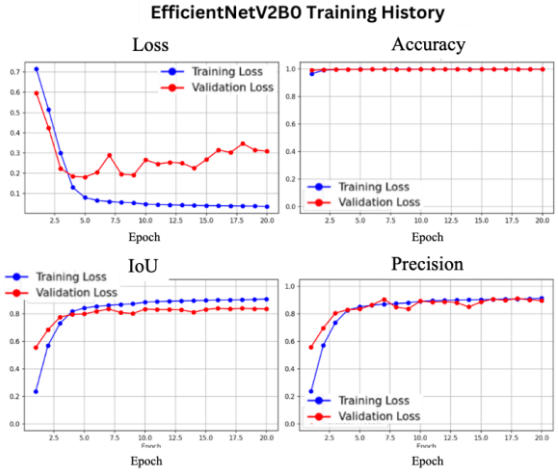


Fig 6. Training plot for AS-Net with EfficientNet V2-B0 encoder.

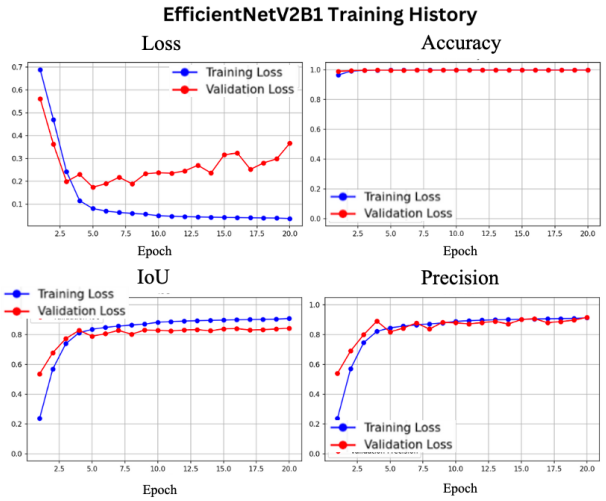


Fig 7. Training plot for AS-Net with EfficientNet V2-B1 encoder.

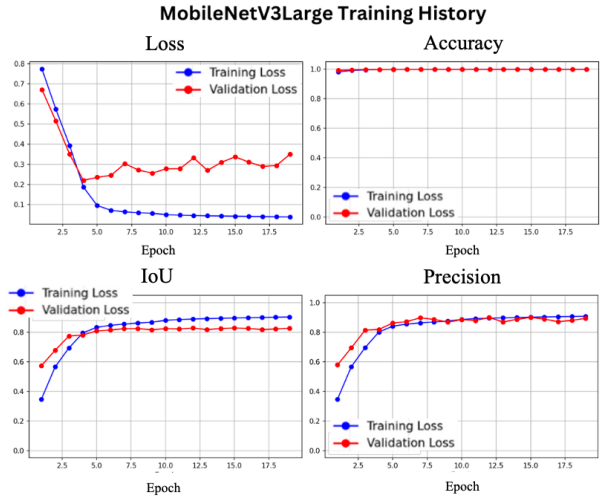


Fig 8. Training plot for AS-Net with MobileNetV3-Large encoder.

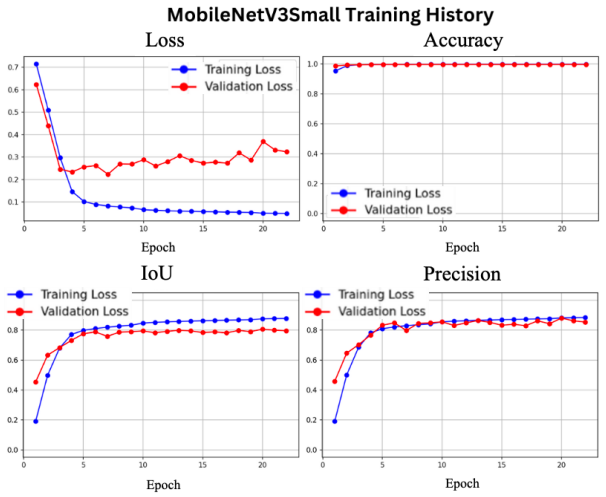


Fig 9. Training history plot for AS-Net with MobileNetV3-Small encoder.

Fig. 5 shows that AS-Net with VGG16 converges rapidly, achieving high validation Dice and IoU scores by epoch 6, followed by stable performance in subsequent epochs. This suggests strong early feature extraction and segmentation capacity of the VGG16 encoder. However, the validation loss curve shows a

slight upward drift after epoch 10, indicating potential early overfitting.

In Fig. 6, the EfficientNetV2-B0 variant exhibits a slightly slower but steady convergence, with improved validation Dice and IoU beginning around epoch 8, particularly after the learning rate reduction. This aligns with the architecture's use of advanced blocks like Fused-MBConv and SE attention, which benefit from longer fine-tuning. Notably, validation performance stabilizes without overfitting, indicating robust generalization.

Fig. 7 presents the training history for EfficientNetV2-B1, which mirrors the behavior of B0 but with slightly higher training instability in validation loss beyond epoch 15. Although the final Dice and IoU are high, the oscillations suggest that this deeper encoder may require more careful regularization or early stopping criteria.

In Fig. 8, MobileNetV3-Large shows moderate training efficiency, achieving competitive Dice and precision scores by epoch 10. However, the validation loss remains higher than that of EfficientNetV2-B0, reflecting potential challenges in handling spatially complex tumor boundaries.

Fig. 9 demonstrates that MobileNetV3-Small, while computationally efficient, exhibits the slowest convergence. Its validation Dice and IoU plateau around epoch 13, and remain lower than the other variants. The gap between training and validation precision further suggests underfitting, which is expected given its significantly reduced parameter count.

Together, these training histories reinforce several observations:

- VGG16 offers the fastest convergence but at high computational cost.
- EfficientNetV2-B0 achieves a strong balance of performance, stability, and generalization.
- MobileNetV3-Small, while fast and lightweight, sacrifices segmentation detail and accuracy.

4.2. Segmentation Performance Comparison

Table 2 presents the quantitative performance metrics of AS-Net variants on the BraTS 2023-GLI test set. The baseline model, AS-Net-VGG16, achieved the highest Dice coefficient (0.8873) and IoU (0.7974), establishing a strong performance benchmark. Remarkably, AS-Net-EfficientNetV2-B0 achieved a Dice score of 0.8853, with a minimal 0.20% drop compared to VGG16, suggesting that advanced lightweight encoders can retain nearly equivalent representational capacity.

Both EfficientNetV2-B1 and MobileNetV3-Large delivered competitive results with Dice scores of 0.8815 and 0.8738, respectively. These values confirm the feasibility of using computationally efficient architectures for medical image segmentation. In contrast, MobileNetV3-Small underperformed with a Dice of 0.8570, revealing a trade-off where excessive model compactness may compromise segmentation granularity—particularly for complex tumor structures.

Table 2. Performance comparison of AS-Net variants.

Encoder	Dice Coef	IoU	Precision	Recall	Binary Acc.
VGG16	**0.8873**	**0.7974**	0.8334	0.9487	0.9974
MNv3-Large	0.8738	0.7759	0.8186	0.9371	0.9971
MNv3-Small	0.8570	0.7498	0.7912	0.9348	0.9967
ENv2-B0	0.8853	0.7942	0.8327	0.9451	**0.9974**
ENv2-B1	0.8815	0.7881	0.8185	**0.9550**	0.9973

This table presents the quantitative performance metrics of all AS-Net variants on the test set.
Despite the relatively small performance drop among lightweight models, the high recall scores across all variants (≥0.9550) indicate strong sensitivity in detecting tumor pixels. However, variations in precision (e.g., 0.7742 for MobileNetV3-Small vs. 0.8289 for VGG16) highlight differences in false positive rates, which may be critical in clinical contexts where over-segmentation could lead to misdiagnosis or overtreatment.

4.3. Accuracy vs. Efficiency Trade-off

Fig.10 compares the total training and inference time per slice for all AS-Net variants. As expected, all lightweight models achieved significant training speed-ups compared to VGG16 (23h 39m). MobileNetV3-Large was the fastest (18h 12m, ~23% reduction), followed closely by EfficientNetV2-B0 (18h 56m, ~20% reduction).

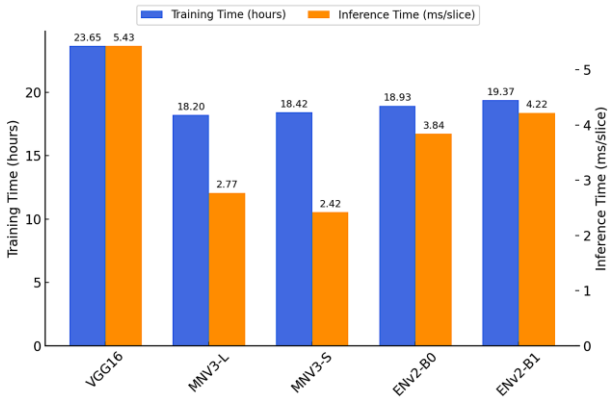


Fig. 10. Training time (hours, blue) and inference time (ms/slice, orange) for AS-Net variants with different encoder backbones. Lower values indicate better efficiency.

Inference efficiency also showed considerable improvement. MobileNetV3-Small was the most efficient at 2.42 ms/slice, while EfficientNetV2-B0 offered a strong trade-off with 3.84 ms/slice, 29% faster than VGG16. These reductions are particularly meaningful for clinical scenarios requiring real-time inference across large 3D MRI volumes.

The performance of EfficientNetV2-B0 is notable because it retains high segmentation accuracy while introducing efficiency gains. This advantage stems from its architecture, which integrates MBConv and Fused-MBConv blocks, optimized depthwise separable convolutions, and Squeeze-and-Excitation (SE) channel attention. These components allow for a high representational capacity at a reduced parameter and FLOP count.

Furthermore, the relative closeness of performance across encoders underscores the robustness and transferability of the AS-Net attention modules (SAM,

CAM, Synergy). Their consistent contribution to performance indicates that the attention synergy mechanism enhances encoder-agnostic feature refinement, allowing high-level semantic features to propagate effectively regardless of the encoder's structural depth or capacity.

4.4. Qualitative Segmentation Examples

Representative segmentation outputs from each AS-Net variant are shown in Figure 6, alongside ground truth masks. Both VGG16 and EfficientNetV2-B0 variants produced segmentation boundaries that closely aligned with ground truth, particularly in delineating tumor cores and irregular edges.

MobileNetV3-Large demonstrated good delineation but showed occasional boundary smoothing, possibly due to its use of inverted residual blocks with reduced receptive fields. MobileNetV3-Small was more prone to under-segmentation, likely due to its significantly reduced model depth and feature abstraction capability.

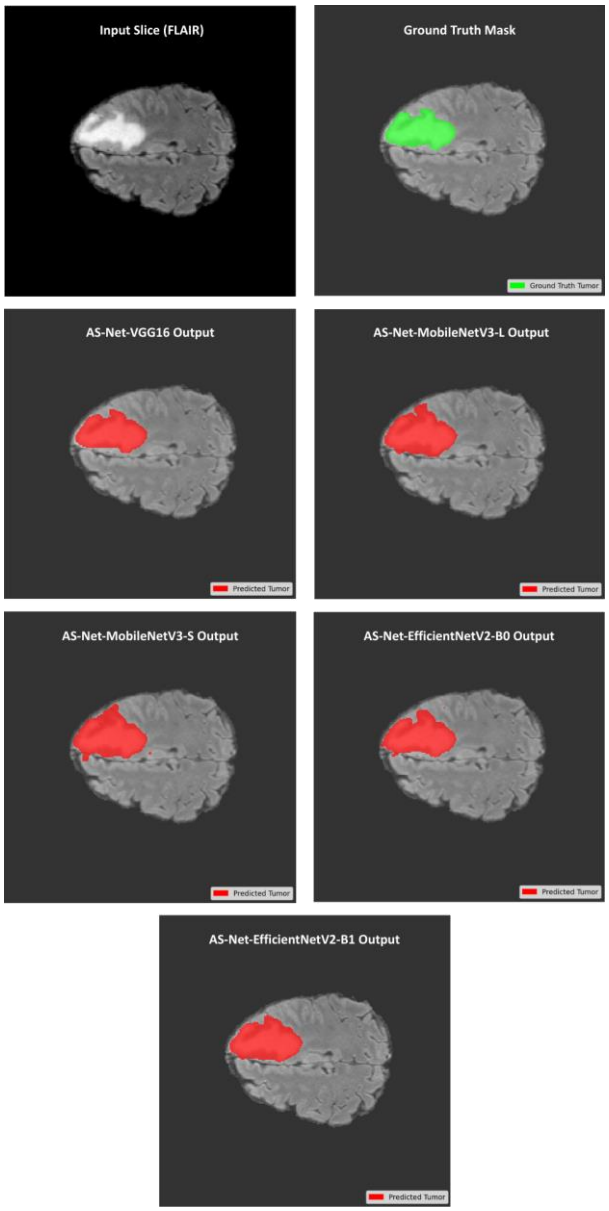


Fig. 11. Qualitative segmentation results for different AS-Net variants compared to ground truth.

These visual insights corroborate the quantitative metrics and validate the claim that EfficientNetV2-B0 provides the best balance of accuracy and speed for practical use.

4.5. Clinical Implications

The findings from this study carry important implications for translating deep learning-based segmentation models into real-world clinical practice:

Efficiency Gains: Lightweight encoders such as EfficientNetV2-B0 and MobileNetV3-Large offer substantial training and inference speedups without compromising significantly on accuracy. These models are highly suitable for deployment in time-critical applications, including emergency settings and real-time diagnostic pipelines.

Hardware Accessibility: The reduced computational demands make it feasible to deploy AS-Net variants on edge devices or mid-range GPUs, expanding access to advanced AI tools in resource-limited hospitals or rural health centers.

Robustness of AS-Net Modules: The attention synergy mechanism proves to be encoder-agnostic, ensuring consistent improvements in performance across different model backbones. This generalizability increases confidence in the adaptability of AS-Net for future use cases involving other medical imaging modalities or segmentation tasks.

Scalability: EfficientNetV2-B0 in particular shows promise as a scalable solution for high-throughput analysis of 3D MRI scans, where fast per-slice inference can accumulate into meaningful runtime savings across full brain volumes.

5. CONCLUSION

This study evaluated five AS-Net variants with different encoder backbones—VGG16, MobileNetV3-Large, MobileNetV3-Small, EfficientNetV2-B0, and EfficientNetV2-B1—for brain tumor segmentation from MRI scans. The results demonstrated that EfficientNetV2-B0 provides the optimal balance between segmentation performance and computational efficiency. It achieved a Dice coefficient of 0.8853, nearly matching the VGG16 baseline (0.8873), while significantly reducing training time by ~20% and inference time by ~29%. In contrast, MobileNetV3 variants offered superior speed but at the expense of accuracy, making them suitable primarily for resource-constrained environments.

These findings have practical implications for deploying advanced segmentation models in clinical workflows. By replacing legacy, computation-heavy encoders with modern, efficient architectures, attention-based networks like AS-Net can be made more accessible and scalable in real-world healthcare settings, especially where processing speed and hardware limitations are key considerations.

However, this study also has limitations. The segmentation task was restricted to binary classification using the BraTS 2023-GLI (Task 1) dataset with a 2D

slice-based approach. The data split was done at the slice level rather than the patient level, which could introduce data leakage. Additionally, a single set of hyperparameters was applied across all model variants, and training/inference times are inherently dependent on the specific hardware used.

Future work may address these constraints through the following directions:

- Extend the framework to multi-class segmentation to differentiate tumor sub-regions such as enhancing tumor, edema, and necrotic core.
- Develop a 3D version of AS-Net to fully leverage the volumetric nature of MRI data.
- Validate the model on multi-center datasets with patient-level data splits to rigorously assess generalizability.
- Explore fine-grained hyperparameter tuning tailored for each encoder to further optimize performance.
- Apply model quantization and compression techniques to enhance inference speed and memory efficiency for deployment in clinical environments.
- Conduct clinical workflow benchmarking to evaluate integration feasibility on standard radiology hardware.

6. REFERENCES

[1] Y. Liu et al., "Tumor segmentation in MRI using deep learning with domain adaptation," Physics in Medicine & Biology, vol. 66, 2021, pp. 1–19.

[2] S. Bakas et al., "Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge," arXiv preprint, arXiv:1811.02629, 2018.

[3] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241.

[4] M. Menze et al., "The multimodal brain tumor image segmentation benchmark (BRATS)," IEEE Transactions on Medical Imaging, vol. 34, 2015, pp. 1993–2024.

[5] Y. Nishimoto, Y. Kajita, and K. Mizoguchi, "Automatic classification of MRI images for brain tumor diagnosis using CNN with attention mechanism," in Proc. of IEICES, vol. 2, 2022, pp. 123–128.

[6] T. Yamada and S. Nakamura, "Lightweight deep learning models for medical image segmentation on resource-constrained devices," in Proc. of IEICES, vol. 3, 2023, pp. 87–92.

[7] L. Hu et al., "AS-Net: Enhanced dermoscopic image segmentation via synergistic attention mechanism," Computer Methods and Programs in Biomedicine, vol. 215, 2022, p. 106600.

[8] F. Isensee et al., "nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," Nature Methods, vol. 18, no. 2, 2021, pp. 203–211.

[9] O. Oktay et al., "Attention U-Net: Learning where to look for the pancreas," in Proc. 1st Conf. on Medical Imaging with Deep Learning (MIDL), 2018. [Online]. Available: https://arxiv.org/abs/1804.03999

[10] J. Hu et al., "Squeeze-and-excitation networks," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 7132–7141. doi:10.1109/CVPR.2018.00745.

[11] A. Canziani, A. Paszke, and E. Culurciello, "An analysis of deep neural network models for practical applications," arXiv preprint, arXiv:1605.07678, 2016.

[12] A. Howard et al., "Searching for MobileNetV3," in Proc. IEEE/CVF Int. Conf. on Computer Vision (ICCV), 2019, pp. 1314–1324.

[13] M. Tan and Q. Le, "EfficientNetV2: Smaller models and faster training," in Proc. Int. Conf. on Machine Learning (ICML), 2021, pp. 10096–10106.

[14] W. Zhang, H. Li, and X. Chen, "Enhanced EfficientNetV2 with global attention for brain tumor classification," Computers in Biology and Medicine, vol. 168, 2024, p. 107713. doi:10.1016/j.compbiomed.2023.107713.

[15] S.-Y. Lin and C.-L. Lin, "Brain tumor segmentation using U-Net in conjunction with EfficientNet," PeerJ Computer Science, vol. 10, 2024, e1754. doi:10.7717/peerj-cs.1754.

[16] H. Munira and M. Islam, "Multi-classification of brain MRI tumor using ConVGXNet, ConResXNet, and ConIncXNet," J. of Korean Physical Society, vol. 8, Oct. 2022, pp. 169–175. doi:10.5109/5909087.

[17] A. Baid et al., "The RSNA-ASNR-MICCAI BraTS 2021 benchmark: Multi-institutional segmentation of brain tumors towards trustworthy imaging AI," The Lancet Digital Health, 2022.

[18] N. M. Rasel et al., "Monte Carlo simulation and experimental determination of tissue phantom ratio for megavoltage photon beam," J. of Korean Physical Society, vol. 6, Oct. 2020. doi:10.5109/4102465.