ISSN: 3126-2953
E-ISSN: 3126-3097

GI-ENSNet: A HYBRID MODEL FOR SEGMENTING SOME GASTROINTESTINAL ORGANS

HO MINH QUANG[a], PHAN ANH CANG
Vinh Long University of Technology Education, Vietnam
[a]Corresponding author: hmquang252@gmail.com

Reviewed: 05/9/2025; Accepted: 25/12/2025

ABSTRACT

Gastrointestinal cancer is a leading cause of mortality, requiring highly accurate treatment approaches such as radiotherapy. Accurate localization of organs such as the stomach, small intestine and large intestine on MRI scans is crucial to minimize radiation-induced damage to surrounding healthy tissues. However, manual segmentation is often time-consuming and prone to subjective errors. This study proposes a deep learning-based model named GI-ENSNet, which ensembles three architectures: UNet with EfficientNet-B2 as backbone, UNet++ with EfficientNet-B0 optimized using Stochastic Weight Averaging (SWA) and a modified UNet integrated with the Boundary Awareness Module (BAM). The outputs of these models are fused using a Soft Voting ensemble strategy, followed by post-processing with the Segment Anything Model (SAM) to refine segmentation masks. The model is trained on a dataset of 38,496 MRI images provided by the UW-Madison Cancer Center and achieves high segmentation performance with a Dice coefficient of 0.9255 and an IoU of 0.8964. The proposed approach significantly enhances segmentation accuracy and supports clinicians in delivering more precise radiotherapy while reducing processing time in clinical practice.

Keywords: *Gastrointestinal tract, Ensemble Learning, Deep Learning, MRI.*

1. INTRODUCTION

1.1. Problem Statement

Gastrointestinal cancers, particularly colorectal and gastric cancers, are rapidly increasing worldwide. Approximately half of patients require radiotherapy, with each treatment course lasting from 1 to 6 weeks and daily sessions of approximately 10–15 minutes. However, to accurately deliver X-ray radiation to tumors while sparing adjacent healthy organs such as the stomach, small intestine, and large intestine, clinicians must manually segment these organs on each MRI image. This process is time-consuming and can extend a radiotherapy session to nearly one hour, leading to patient fatigue and reduced treatment efficiency.

To address this challenge, this study proposes the application of artificial intelligence, particularly deep learning models, to automate the segmentation of gastrointestinal organs in MRI images. We develop a hybrid deep learning framework that integrates three architectures: UNet [1] with an EfficientNet-B2 encoder [2], UNet++ [3] with EfficientNet-B0 incorporating Stochastic Weight Averaging [4], and a modified UNet enhanced with a Boundary Awareness Module [5]. The outputs from these three models are combined using a soft voting strategy [6], followed by post-processing with the

ISSN: 3126-2953
E-ISSN: 3126-3097

Segment Anything Model [7] to refine segmentation based on the original images, thereby improving overall accuracy.

The proposed model not only reduces the workload of clinicians but also contributes to improving treatment effectiveness for patients with gastrointestinal cancers.

1.2. Related Work

In recent years, medical image segmentation, particularly in the gastrointestinal (GI) domain, has achieved significant progress [8]. UNet, introduced by Ronneberger et al. in 2015 [9], is one of the pioneering models featuring a symmetric encoder–decoder architecture that effectively preserves spatial features. Building upon this foundation, numerous improved models have been proposed. Yuan developed a saliency-based model for polyp segmentation from endoscopic videos [10]. ResUNet++ achieved Dice scores of 81.33% on Kvasir-SEG and 79.55% on CVC-612 datasets [11]. A GAN-based approach by Poorneshwaran reported a Dice score of 88.48% [12]. TMD-UNet (2.5D) incorporates depth information to enhance segmentation performance [13]. Ghosh improved brain tumor segmentation accuracy by integrating VGG-16 into the UNet architecture [14].

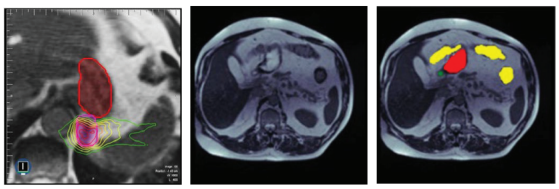
In 2022, Punn et al. conducted a comprehensive survey of UNet variants, highlighting that although effective, these models require substantial computational resources and are highly dependent on input data quality [15]. In the same year, a study by Alina Chou at Stanford University demonstrated that Mask R-CNN outperformed UNet in gastrointestinal MRI segmentation, achieving a Dice coefficient of 0.7266 compared to 0.5134 [16].

In early 2023, Zhou et al. proposed a 2.5D UNet with a Se-ResNet50 backbone, achieving a Dice score of 0.848 on the UW-Madison dataset [17]. Meanwhile, Sharma combined EfficientNet-B0 with a Feature Pyramid Network (FPN), attaining a Dice score of 0.8975 and an IoU of 0.8832 [18]. These results indicate that current research trends focus on enhancing encoder architectures, leveraging multi-dimensional data, and incorporating attention mechanisms to improve GI segmentation performance, with a target Dice score approaching 0.90.

2. RELATED BACKGROUND

2.1. Gastrointestinal Organs and Gastrointestinal Cancers

The stomach, small intestine, and large intestine are the three primary organs of the gastrointestinal system, responsible for ingestion, digestion, and waste elimination. In the study and treatment of gastrointestinal cancers, accurately identifying the location and morphology of these organs in MRI images plays a crucial role in diagnosis, disease monitoring, and the development of personalized treatment plans.



(a) Tumor image (b) Abdominal image (c) Segmentation overlay

Figure 1. Representative MRI images from the UW-Madison gastrointestinal dataset

Figure 1(a) illustrates a tumor (outlined in magenta) located near the stomach (outlined in red), where a high radiation dose is concentrated on the tumor while sparing the stomach. The dose distribution is visualized using a rainbow color scale, ranging from red (high) to green (low). Figure 1(b) presents an MRI image from

ISSN: 3126-2953
E-ISSN: 3126-3097

the UW-Madison dataset. Figure 1(c) shows three segmentation masks overlaid on the original image: the stomach (red), large intestine (yellow), and small intestine (green).

2.2. Deep Learning Models and Training Methods

In this study, we propose an ensemble of three advanced deep learning models to address the problem of segmenting gastrointestinal organs in abdominal MRI images. The three models are described as follows:

UNet with EfficientNet-B2 Backbone

We employ EfficientNet-B2 [2] as the encoder for the UNet architecture to enhance feature extraction capability while optimizing computational efficiency. EfficientNet-B2, developed by Google Brain, is based on Neural Architecture Search and Compound Scaling [19], achieving strong performance in both accuracy and computational efficiency.

UNet++ with EfficientNet-B0 Backbone and SWA

The second model is UNet++ [3], an extended variant of UNet featuring densely connected skip pathways between the encoder and decoder, which improves feature representation and segmentation performance. We utilize EfficientNet-B0 as the encoder and apply Stochastic Weight Averaging (SWA) [4] to improve training stability, reduce overfitting, and enhance overall performance.

Improved UNet with Boundary-Aware Module (BAM)

The third model is a modified UNet integrated with a Boundary-Aware Module (BAM) [5], which consists of two attention branches: Channel Attention, emphasizing important feature channels, and Spatial Attention, focusing on spatial regions with well-defined organ boundaries (see Figure 2). This two-stage attention mechanism enables the UNet-BAM model to learn more informative features and reduce noise in complex anatomical structures in MRI images.

Ensemble Learning with Soft Voting

The segmentation outputs from the three models are combined using a soft voting strategy [6], in which the predicted probabilities at each pixel are averaged to generate a unified probability map.

Specifically, for each pixel xxx, the three segmentation models produce logits $z_1(x), z_2(x), z_3(x)$. After applying the sigmoid function, the corresponding probabilities $p_i(x)$ are obtained:

$$p_i(x) = \sigma\left(z_i(x)\right) = \frac{1}{1 + e^{-z_i(x)}} \quad , i = \{1, 2, 3\}$$

For unweighted soft voting, the aggregated probability $\bar{p}_x$ is computed as the average:

$$\bar{p}_x = \frac{p_1(x) + p_2(x) + p_3(x)}{3}$$

Post-processing and Accuracy Enhancement with the Segment Anything Model (SAM)

After model ensembling, we apply post-processing using the Segment Anything Model (SAM) [7] to refine the segmentation results based on the input images. SAM enables flexible boundary detection and captures fine-grained details, thereby improving the final segmentation accuracy of the proposed model.

2.3. Loss Functions and Evaluation Metrics

Loss Function

ISSN: 3126-2953
E-ISSN: 3126-3097

Binary Cross Entropy (BCE), as defined in Equation (1), measures the discrepancy between the predicted probabilities and the ground truth labels. Let N denote the total number of pixels (or samples) in a batch. y_i the ground truth label of the *i-th* pixel. y_i the predicted probability that pixel *i* belongs to the positive class.

$$\text{BCE} = -\frac{1}{\text{N}}\sum_{i=1}^{N}[y_i.\log(\hat{y_i}) + (1 - y_i).\log(1 - \hat{y_i})] \quad (1)$$

Tversky Loss (as defined in Equation (2)) is a generalized loss function derived from Dice Loss and Intersection over Union (IoU) Loss.

$$\text{Tversky Loss} = 1 - \frac{|Y \cap \hat{Y}| + \epsilon}{|Y \cap \hat{Y}| + \alpha|Y - \hat{Y}| + \beta|\hat{Y} - Y| + \epsilon} \quad (2)$$

Where:

- Y denotes the ground truth mask, in which pixels belonging to the positive class (i.e., the target organs to be segmented) are labeled.
- $\hat{Y}$ denotes the predicted mask, where pixels are classified by the model as belonging to the positive class.
- α is a weighting factor that controls the penalty for false negatives (False Negatives).
- β is a weighting factor that controls the penalty for false positives (False Positives).
- ϵ is a small constant added to both the numerator and denominator to prevent division by 0.

The final loss function is a linear combination of the two aforementioned losses, as defined in Equation (3):

$$\mathcal{L}_{\text{total}} = 0.5 \cdot \mathcal{L}_{\text{BCE}} + 0.5 \cdot \mathcal{L}_{\text{Tversky}} \quad (3)$$

Evaluation Metrics

To evaluate the performance of the segmentation model, this study employs two widely used metrics: the Dice Coefficient and Intersection over Union (IoU). The Dice coefficient measures the degree of overlap between the predicted and ground truth regions, while IoU is computed based on the ratio of the intersection to the union of the two regions.

3. PROPOSED METHOD

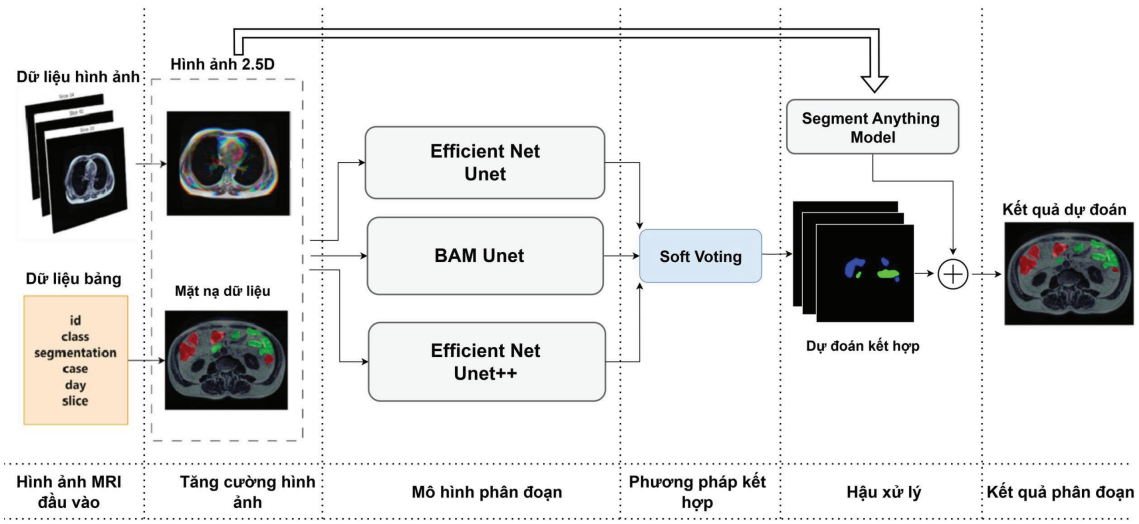


Figure 2. The proposed GI-ENSNet framework

ISSN: 3126-2953
E-ISSN: 3126-3097

Figure 2 illustrates the overall workflow of the proposed method for gastrointestinal organ segmentation. The process is divided into four main stages: image augmentation, segmentation models, ensembling method, and post-processing. Each stage is described in detail as follows.

3.1. Image Augmentation

The proposed segmentation method employs a 2.5D technique that combines the advantages of both 2D and 3D segmentation, enabling the exploitation of spatial and depth information without requiring excessive computational resources. The data augmentation process includes several key techniques, such as horizontal flipping, translation, scaling, rotation, geometric distortions (e.g., GridDistortion and ElasticTransform), noise injection, and blurring (e.g., GaussNoise and Blur).

In this study, the input data are constructed in a 2.5D format by stacking adjacent slices into multiple channels, allowing a 2D model to capture contextual information along the z-axis (depth). Specifically, with three channels and a stride of 2, each input sample consists of the current slice ttt, along with two neighboring slices at offsets of 2 and 4 steps (t+2, t+4). These three slices are combined to form a three-channel image, preserving the computational efficiency of 2D approaches while incorporating spatial context from adjacent slices, thereby improving segmentation accuracy.

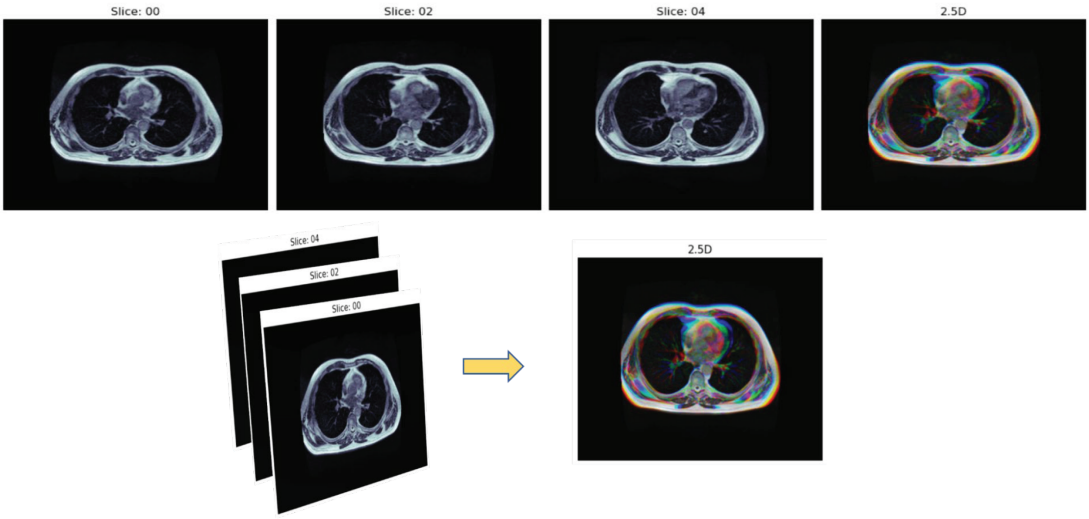


Figure 3. Illustration of the transformation from 2D slices to a 2.5D representation

3.2. Segmentation Models

After preprocessing, the MRI images are fed into three enhanced deep learning models: UNet with EfficientNet-B2, UNet++ combined with EfficientNet-B0 and SWA, and a UNet integrated with the BAM module. Each model learns distinct feature representations, enabling multi-dimensional exploitation of information from medical images. The training process results in three independent segmentation models, each capable of capturing different characteristics from the data.

3.3. Ensembling and Post-processing Stage

ISSN: 3126-2953
E-ISSN: 3126-3097

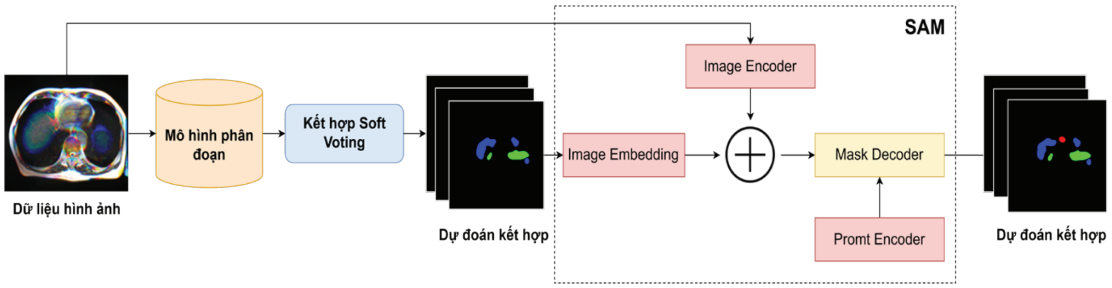


Figure 4. Illustration of the ensembling and post-processing stage

Figure 4 illustrates the ensembling process, in which the outputs from the three trained models are integrated using a soft voting strategy. The test image is fed into each model, producing pixel-wise probability maps for segmentation. These probabilities are then averaged to generate a combined segmentation result. Subsequently, the Segment Anything Model (SAM) is applied to further refine the output based on the original image, thereby improving the final segmentation accuracy.

4. Experimental Results

4.1. Implementation Environment and Dataset

Implementation Environment

The model was trained using Python in a Kaggle Notebook environment, leveraging a multi-core CPU, 29.5 GB of RAM, and an NVIDIA P100 GPU with 15 GB of VRAM, along with the necessary supporting libraries.

Experimental Dataset

This study utilizes an abdominal magnetic resonance imaging (MRI) dataset collected from patients undergoing and having undergone treatment at the UW-Madison Carbone Cancer Center [8], where MR Linac technology has been employed since 2015 to integrate radiotherapy with MRI for more precise cancer treatment. The dataset consists of 38,496 MRI images in 16-bit PNG format, collected from 85 patients across multiple scanning days and slices. The data have been anonymized and publicly released on the Kaggle platform (https://www.kaggle.com/competitions/uw-madison-gi-tract-image-segmentation).

4.2. Training Scenarios

Table 1. Summary of training scenarios

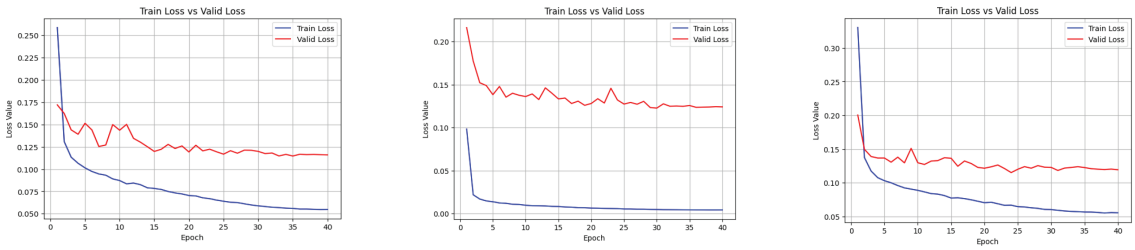
ISSN: 3126-2953
E-ISSN: 3126-3097

Scenario	Training model	Image size	Batch size	Num classes	Learning rate	Epochs	Optimizer
1	UNet - EfficientNetB2	512 x 512	64	3	1e-6	40	AdamW
2	UNet++ - SWA	512 x 512	64	3	1e-6	40	AdamW
3	UNet - BAM	512 x 512	64	3	1e-6	40	AdamW
4	Proposed method (GI-ENSNet, ensemble of Scenarios 1-2-3 with post-processing)	512 x 512	-	3	-	-	-

Table 1 presents the training scenarios corresponding to different models, including Scenario 1 (UNet with EfficientNet-B2), Scenario 2 (UNet++-SWA), Scenario 3 (UNet- BAM), and the proposed method in Scenario 4 (GI-ENSNet). Experiments were conducted on the dataset using training parameters such as batch size, learning rate, and number of epochs. Three classes were considered for segmentation: the stomach, large intestine, and small intestine.

4.3. Experimental Results

Comparison of loss values across different training scenarios



ISSN: 3126-2953
E-ISSN: 3126-3097

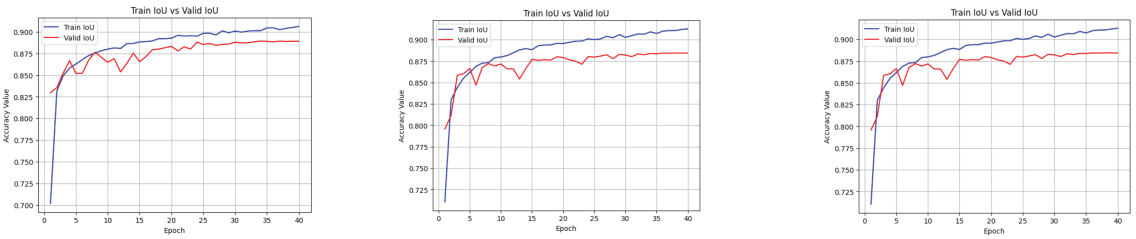
Scenario 1 Scenario 2 Scenario 3

Figure 5. Comparison of loss values across different training scenarios

Figure 5 presents the loss values of the three training scenarios on both the training and test sets. Scenario 1 achieves a loss of 0.0547 on the training set and 0.1146 on the test set. Scenario 2, designed to improve stability, achieves a loss of 0.0042 on the training set and 0.1227 on the test set. Scenario 3 exhibits higher variability, with a loss of 0.0551 on the training set and 0.1150 on the test set.

Comparison of accuracy across training scenarios

Scenario 1 *Scenario 2* *Scenario 3*



Scenario 1 *Scenario 2* *Scenario 3*

Figure 6. Comparison of segmentation performance measured by Dice (top row) and IoU (bottom row)

Figure 6 presents the segmentation performance of the three training scenarios based on the Dice and IoU metrics on both the training and test sets. Scenario 1 achieves Dice scores of 0.9328 (train) and 0.9164 (test), and IoU scores of 0.9062 (train) and 0.8893 (test). Scenario 2 demonstrates good stability, with Dice scores of 0.9387 (train) and 0.9120 (test), and IoU scores of 0.9126 (train) and 0.8845 (test). Scenario 3 achieves Dice scores of 0.9319 (train) and 0.9127 (test), and IoU scores of 0.9051 (train) and 0.8854 (test).

Training time across different scenarios

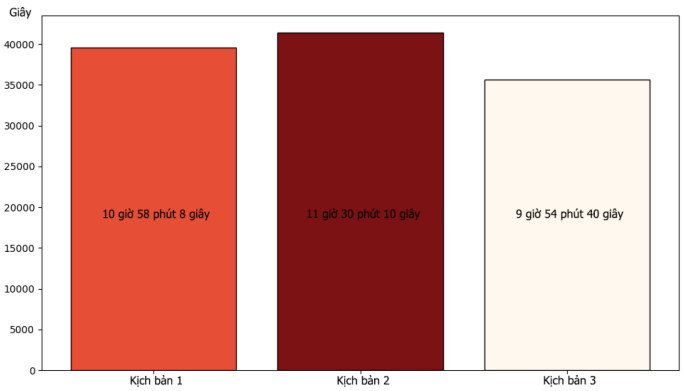


Figure 7. Comparison of training time across different scenarios

ISSN: 3126-2953
E-ISSN: 3126-3097

Figure 7 illustrates the training time of the different scenarios on the UW-Madison dataset. Scenario 1 requires 10 hours and 59 minutes (39,548 seconds), Scenario 2 takes 11 hours and 30 minutes, while Scenario 3 achieves the shortest training time of 9 hours and 54 minutes.

Trade-off between segmentation accuracy and processing time across different scenarios

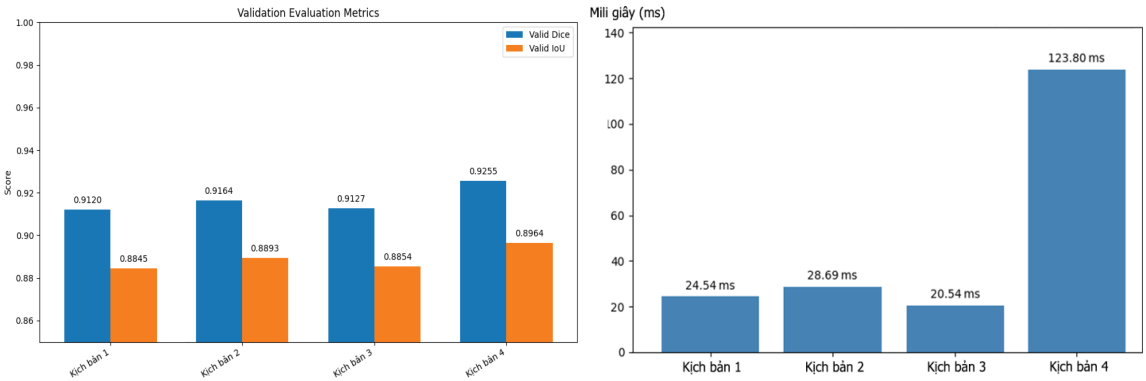


Figure 8. Trade-off between segmentation accuracy and processing time across different scenarios

Figure 8 shows that Scenario 4 achieves the highest segmentation accuracy on the test set, with a Dice score of 0.9255 and an IoU of 0.8964, outperforming the individual models. However, this scenario also incurs the longest processing time, at 123.8 ms per image, due to the need to load and run all three models simultaneously for soft voting, followed by SAM-based post-processing.

Comparison of Experimental Results

Table 2. Summary of experimental results

Authors	Methods	Dice	IoU
Ye, R et al [20]	SIA-Unet	-	0.65
Chou, A. et al [21]	UNet + Mask R-CNN	0.73	-
Zhou, H. et al [17]	Se-ResNet50	0.848	-
Sharma, N. et al [18]	UNet Decoder FPN	0.8975	0.8832
Qiu, Y. [22]	Swin Transformer with EfficientNet B4	0.8682	-
Jiang, X. [23]	BiFTransNet transformer-based model	0.8951	-
John, S.V. [24]	UNet with EfficientNet B7	0.8991	0.8693
Sharma, Neha, et al. [25]	DeepLab V3+ Model with ResNet 50 Encoder	0.9082	0.8796
Proposed method	GI-ENSNet	0.9255	0.8964

Figure 8 shows that Scenario 4 achieves the highest segmentation accuracy on

ISSN: 3126-2953
E-ISSN: 3126-3097

the test set, with a Dice score of 0.9255 and an IoU of 0.8964, outperforming the individual models. However, this scenario also incurs the longest processing time, at 123.8 ms per image, due to the need to load and run all three models simultaneously for soft voting, followed by SAM-based post-processing.

5. CONCLUSION

This study proposes a method for gastrointestinal organ segmentation in MRI images by integrating three deep learning models: UNet with EfficientNet-B2, UNet++ with EfficientNet-B0 combined with Stochastic Weight Averaging (SWA), and UNet enhanced with a Boundary-Aware Module (BAM). The models are integrated using a soft voting strategy, followed by post-processing with the Segment Anything Model (SAM).

Experimental results demonstrate that the proposed method significantly improves segmentation accuracy compared to individual models, achieving a Dice score of 0.9698 on the training set and 0.9255 on the test set, and IoU scores of 0.9388 and 0.8964, respectively. These improvements confirm the effectiveness of combining multiple architectures with post-processing techniques in enhancing both accuracy and stability.

However, the use of Dice and IoU metrics to determine the application of SAM increases computational time. Therefore, future work will focus on automating and optimizing this process to reduce processing time. In addition, this study has not explored larger backbone models such as EfficientNet-B7 or newer variants of EfficientNet. Similarly, the SAM2 model has not been investigated and represents a potential direction for future research.

Furthermore, the current study focuses only on three organs (stomach, small intestine, and large intestine), which limits its generalizability. Future work should extend the approach to additional organs to broaden its applicability.

REFERENCE

[1] Ronneberger, O., Fischer, P., & Brox, T. (2015). Unet: Convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 (pp. 234-241). Springer international publishing.

[2] Tan, M., & Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning (pp. 6105-6114). PMLR.

[3] Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2018). Unet++: A nested Unet architecture for medical image segmentation. In Deep learning in medical image analysis and multimodal learning for clinical decision support: 4th international workshop, DLMIA 2018, and 8th international workshop, ML-CDS 2018, held in conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, proceedings 4 (pp. 3-11). Springer International Publishing.

ISSN: 3126-2953
E-ISSN: 3126-3097

[4] Guo, H., Jin, J., & Liu, B. (2023). Stochastic weight averaging revisited. Applied Sciences, 13(5), 2935.

[5] Sun, X., Shi, A., Huang, H., & Mayer, H. (2020). BAS $^{4} $ Net: Boundary-aware semi-supervised semantic segmentation network for very high resolution remote sensing images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 5398-5413.

[6] Kumari, S., Kumar, D., & Mittal, M. (2021). An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier. International Journal of Cognitive Computing in Engineering, 2, 40-46.

[7] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., ... & Girshick, R. (2023). Segment anything. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 4015-4026).

[8] Pogorelov, K., Randel, K. R., Griwodz, C., Eskeland, S. L., de Lange, T., Johansen, D., ... & Halvorsen, P. (2017, June). Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection. In Proceedings of the 8th ACM on Multimedia Systems Conference (pp. 164-169).

[9] Ronneberger, O., Fischer, P. and Brox, T., 2015. Unet: Convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 (pp. 234-241). Springer International Publishing.

[10] Yuan, Y., Li, D. and Meng, M.Q.H., 2017. Automatic polyp detection via a novel unified bottom-up and top-down saliency approach. IEEE journal of biomedical and health informatics, 22(4), pp.1250-1260.

[11] Jha, D., Smedsrud, P.H., Riegler, M.A., Johansen, D., De Lange, T., Halvorsen, P. and Johansen, H.D., 2019, December. Resunet++: An advanced architecture for medical image segmentation. In 2019 IEEE international symposium on multimedia (ISM) (pp. 225-2255). IEEE.

[12] Poorneshwaran, J.M., Kumar, S.S., Ram, K., Joseph, J. and Sivaprakasam, M., 2019, July. Polyp segmentation using generative adversarial network. In 2019 41St annual international conference of the IEEE engineering in medicine and biology society (EMBC) (pp. 7201-7204). IEEE.

[13] Tran, S.T., Cheng, C.H., Nguyen, T.T., Le, M.H. and Liu, D.G., 2021, January. Tmd-unet: Triple-unet with multi-scale input features and dense skip connection for medical image segmentation. In Healthcare (Vol. 9, No. 1, p. 54). MDPI.

[14] Ghosh, S., Chaki, A. and Santosh, K.C., 2021. Improved UNet architecture with VGG-16 for brain tumor segmentation. Physical and Engineering Sciences in Medicine, 44(3), pp.703-712.

[15] Punn, N.S. and Agarwal, S., 2022. Modality specific UNet variants for biomedical

ISSN: 3126-2953
E-ISSN: 3126-3097

image segmentation: a survey. Artificial Intelligence Review, 55(7), pp.5845-5889.

[16] Chou, A., Li, W. and Roman, E., 2022. GI tract image segmentation with UNet and mask R-CNN. Image Segmentation with UNet and Mask R-CNN. Available online: http://cs231n. stanford. edu/reports/2022/pdfs/164. pdf (accessed on 4 June 2023).

[17] Zhou, H., Lou, Y., Xiong, J., Wang, Y. and Liu, Y., 2023. Improvement of Deep Learning Model for Gastrointestinal Tract Segmentation Surgery. Frontiers in Computing and Intelligent Systems, 6(1), pp.103-106.

[18] Sharma, N., Gupta, S., Reshan, M.S.A., Sulaiman, A., Alshahrani, H. and Shaikh, A., 2023. EfficientNetB0 cum FPN based semantic segmentation of gastrointestinal tract organs in MRI scans. Diagnostics, 13(14), p.2399.

[19] Elsken, T., Metzen, J. H., & Hutter, F. (2019). Neural architecture search: A survey. Journal of Machine Learning Research, 20(55), 1-21.

[20] Ye, R.; Wang, R.; Guo, Y.; Chen, L. SIA-Unet: A Unet with Sequence Information for Gastrointestinal Tract Segmentation. In Proceedings of the Pacific Rim International Conference on Artificial Intelligence, Shanghai, China, 10–13 November 2022; Springer: Cham, Switzerland, 2022; pp. 316–326

[21] Chou, A.; Li, W.; Roman, E. GI Tract Image Segmentation with UNet and Mask R-CNN. CS231n: Deep Learning for Computer Vision, Stanford University. 2022. Available online: https://cs231n.stanford.edu/reports/2022/pdfs/164.pdf (accessed on 1 February 2025).

[22] Qiu, Y. Upernet-Based Deep Learning Method for The Segmentation of Gastrointestinal Tract Images. In Proceedings of the 2023 8th International Conference on Multimedia and Image Processing, Tianjin, China, 21–23 April 2023; pp. 34–39.

[23] Jiang, X.; Ding, Y.; Liu, M.; Wang, Y.; Li, Y.; Wu, Z. BiFTransNet: A unified and simultaneous segmentation network for gastrointestinal images of CT & MRI. Comput. Biol. Med. 2023, 165, 107326

[24] John, S.V.; Benifa, B. Automated Segmentation of Tracking Healthy Organs from Gastrointestinal Tumor Images. In Proceedings of the International Conference on Frontiers of Intelligent Computing: Theory and Applications, Cardiff, UK, 11–12 April 2023; Springer Nature: Singapore, 2023; pp. 363–373.

[25] Sharma, Neha, et al. "Encoder–Decoder Variant Analysis for Semantic Segmentation of Gastrointestinal Tract Using UW-Madison Dataset." Bioengineering 12.3 (2025): 309.